Estimating the limit state space of quasi-nonlinear Fuzzy Cognitive Maps

Peer-reviewed author version

CONCEPCION PEREZ, Leonardo; NAPOLES RUIZ, Gonzalo; Jastrzebska, Agnieszka; Grau, Isel & Salgueiro, Yamisleydi (2025) Estimating the limit state space of quasi-nonlinear Fuzzy Cognitive Maps. In: Applied soft computing, 169 (Art N° 112604).

DOI: 10.1016/j.asoc.2024.112604
Handle: http://hdl.handle.net/1942/45182

Estimating the Limit State Space of Quasi-Nonlinear Fuzzy Cognitive Maps

Leonardo Concepción[a,b], Gonzalo Nápoles[c,a], Agnieszka Jastrzebska[1], Isel Grau[e], Yamisleydi Salgueiro[f,*]

[a] *UHasselt, Business Informatics Research Group, Hasselt, Belgium*
[b] *Universidad Central de Las Villas, Santa Clara, Cuba*
[c] *Department of Cognitive Science & Artificial Intelligence, Tilburg University, The Netherlands.*
[d] *Faculty of Mathematics and Information Science, Warsaw University of Technology, Poland.*
[e] *Information Systems Group, Eindhoven University of Technology, The Netherlands.*
[f] *Department of Industrial Engineering, Faculty of Engineering, Universidad de Talca, Campus Curicó, Chile.*

Abstract

Quasi-Nonlinear Fuzzy Cognitive Maps (q-FCMs) generalize the classic Fuzzy Cognitive Maps (FCMs) by incorporating a nonlinearity coefficient that is related to the model's convergence. While q-FCMs can be configured to avoid unique fixed-point attractors, there is still limited knowledge of their dynamic behavior. In this paper, we propose two iterative, mathematically-driven algorithms that allow estimating the limit state space of any q-FCM model. These algorithms produce accurate lower and upper bounds for the activation values of neural concepts in each iteration without using any information about the initial conditions. As a result, we can determine which activation values will never be produced by a neural concept regardless of the initial conditions used to perform the simulations. In addition, these algorithms could help determine whether a classic FCM model will converge to a unique fixed-point attractor. As a second contribution, we demonstrate that the covering of neural concepts decreases as the nonlinearity coefficient approaches its maximal value. However, large covering values do not necessarily translate into better approximation capabilities, especially in the case of nonlinear problems. This finding points to a trade-off between the model's nonlinearity and the number of reachable states.

Keywords: Fuzzy Cognitive Maps, Recurrent Neural Networks, Modeling

and Simulation, Convergence Analysis.

1. Introduction

Fuzzy Cognitive Maps (FCMs) [1] are recurrent learning systems composed of well-defined neural concepts and causal relationships. They have gained substantial attention due to their effectiveness in data processing tasks such as scenario simulation [2, 3, 4], the modeling of control systems [5, 6], pattern classification [7, 8, 9, 10], multi-output regression [11, 12], time series forecasting [13, 14, 15], and federated learning [16, 17, 18, 19]. Compared to other recurrent neural systems, FCMs are preferred when expert knowledge can be integrated into the model in the form of causal relationships or constraints. Such a feature not only enables building models that better represent the problem domain but also supports hybrid reasoning. The intrinsic interpretability offered by these cognitive networks [20, 21, 22] also accounts for the popularity of these models among practitioners.

During reasoning, FCMs update the neurons' activation values using a recurrent approach, where the values produced in the current iteration are used to compute the system's output in the next iteration. After performing a fixed number of iterations, one of the following dynamic behaviors will be observed [23]: (i) activation values remain constant; (ii) activation values exhibit a limit cycle; and (iii) activation values show no regular pattern. Stabilized output signals are used as the basis for prediction, which can be utilized to make decisions or assign class labels to a given problem instance. As Bottero et al. [24] and Groumpos [25] underlined, reaching an equilibrium point is essential for a valid model interpretation and correct execution of decision-making and simulation tasks in most cases. However, unique fixed-point attractors are the downfall of FCM models used for prediction and scenario simulation, rendering them invalid. A more explicit explanation is that an FCM model governed by a single fixed point will always produce the same output, regardless of the input used to start reasoning. This behavior contradicts the fundamental principle of any prediction task where the model's response cannot remain invariant.

*Corresponding author
Email address: ysalgueiro@utalca.cl (Yamisleydi Salgueiro)

The literature reports a range of theoretical studies investigating the conditions under which an FCM model would converge to fixed-point attractors. The most influential work in this area has been conducted by Harmati, Hatwagner, and Kóczy [26, 27, 28, 29]. Their studies focused on detecting the existence and uniqueness of fixed-point attractors in FCM-based models using several mathematical tools. Their research also covered FCM extensions, including FCMs implemented with rough sets [30] and interval sets [31]. In another paper from the same research group [32], they performed a sensitivity analysis to highlight the relationship between weight values and the occurrence of a unique fixed-point attractor. Despite the remarkable progress in understanding the dynamic of FCM models, the challenge of how to avoid unique fixed-point attractors has remained an open issue.

The Quasi-Nonlinear Fuzzy Cognitive Maps (q-FCMs) were introduced by Nápoles et al. [33] to resolve the convergence issues of FCMs, particularly those caused by unique fixed-point attractors. This model includes a nonlinearity coefficient that regulates the amount of information resulting from the reasoning process used to update the neurons' activation values. Therefore, instead of dealing with the aftermaths of potentially malfunctioning models, q-FCMs allow for configuring the model through the nonlinearity coefficient to ensure that unique fixed points are never produced.

Although q-FCMs hold significant promise in the context of scenario analysis and machine learning tasks, their dynamic properties have not been extensively studied. Specifically, there is limited understanding of how their state space behaves during reasoning under the influence of the nonlinearity coefficient. More importantly, to our knowledge, no algorithmic method has been developed to estimate bounds on the neurons' reachable activation values in q-FCM models without knowing the initial conditions. In practice, domain experts run the reasoning process for a limited set of initial activator vectors and draw conclusions about the behavior of concepts' activation values. The drawback of this approach is that the observed behavior is likely to change when using a different set of initial conditions, which might not even be available in some application domains.

This paper brings two contributions to the theoretical analysis of q-FCM models and their dynamics. Firstly, we propose two iterative algorithms to estimate the *limit state space* of any q-FCM model, supported by two theorems and several definitions extended from the FCM theory. These algorithms produce accurate lower and upper bounds for the neurons' activation values at each iteration without using any information about the initial conditions.

Such a tool will enable domain experts to determine which activation values will never be produced by a neural concept regardless of the initial conditions used to perform the simulations. In addition, these algorithms could indicate whether a classic FCM model will converge to a unique fixed point, which must be avoided at all costs in most application domains. Secondly, we prove that the covering values of neural concepts in q-FCM models become smaller as the nonlinearity coefficient approaches its maximal value. However, it should be stressed that large covering values do not necessarily translate into better approximation capabilities. This finding indicates a trade-off between the model's nonlinearity and the number of reachable states, which helps explain why classic FCM models sometimes perform poorly when solving complex simulation problems. The numerical simulations using both real-world and synthetically generated problems provide empirical evidence for the accuracy and usability of the estimated bounds.

In the following sections of this article, we introduce the theoretical background of q-FCMs (Section 2) and the framework for state space analysis, including shrink functions and their extension to q-FCMs (Section 3). Next, Section 4 demonstrates the shrinking state space theorems for q-FCMs. Section 5 addresses the impact of the nonlinearity coefficient on the estimated state spaces. Section 6 presents an empirical study concerning the proposed approach. Finally, Section 7 concludes the study.

2. Quasi-Nonlinear Fuzzy Cognitive Maps

FCMs were introduced in [1] as a tool for modeling causal relationships between concepts within a complex system. These recurrent neural networks are represented as directed and weighted graphs supporting feedback loops, with nodes symbolizing neural concepts and edges denoting causal relationships. In these neural systems, the weight $w_{ij} \in [-1, 1]$ characterizes the impact of neural concept C_i on C_j, which can be positive, negative, or neutral. The weight matrix $\mathbf{W_{M\times M}}$ gathers the causal relationships in the cognitive network. Originally, domain experts were tasked with building the FCM topology, but modern approaches have resorted to supervised learning algorithms to derive the network structure from historical data.

Aiming at improving FCMs' convergence features, Nápoles et al. [9] introduced the q-FCM model as a generalization of the classic FCM formalism. The key difference between these two approaches is that q-FCMs use a *quasi-nonlinear reasoning rule*. This rule incorporates a nonlinearity coefficient $\phi \in [0, 1]$ to control the degree to which the model considers the value produced by the activation function over the initial activation value of the neuron. Equation (1) illustrates this reasoning rule:

$$a_i^{(t+1)} = \underbrace{\phi\, f_i \left(\sum_{j=1}^{M} w_{ji} a_j^{(t)} \right)}_{\text{nonlinear part}} + (1-\phi) \underbrace{a_i^{(0)}}_{\text{linear part}}, i \neq j \tag{1}$$

where the activation value of the j-th neuron at the t-th iteration ($t \in \mathbb{N}$) is represented by $a_j^{(t)}$, while $f_i(.)$ is the activation function used to constrain the activation values of the i-th neuron to a desired interval. The nonlinear part from Equation (1) can be also expressed as $f_i\left(w_i \mathbf{A^{(t)}}\right)$, where $\mathbf{A^{(t)}} = (a_1^{(t)}, a_2^{(t)}, \ldots, a_M^{(t)})$ is the activation vector and $w_i = (w_{1i}, w_{2i}, \ldots, w_{i-1i}, 0, w_{i+1i}, \ldots, w_{Mi})$, $\forall i \in \{1, 2, \ldots, M\}$. Therefore, we can rewrite the reasoning rule using vector-like notation as follows:

$$\mathbf{A^{(t+1)}} = \phi f\left(\mathbf{A^{(t)}W}\right) + (1-\phi)\mathbf{A^{(0)}} \tag{2}$$

where $f\left(\mathbf{A^{(t)}W}\right) = (f_1\left(w_i \mathbf{A^{(t)}}\right), \ldots, f_M\left(w_M \mathbf{A^{(t)}}\right))$.

Typically, the reasoning rule is executed until the q-FCM converges to a fixed-point attractor, which can be unique or multiple depending on the ϕ parameter. Conversely, the q-FCM is deemed unstable if it fails to stabilize after a fixed number of iterations. Notably, instability manifests itself in two distinct forms: cyclic or chaotic behavior [9].

When the nonlinearity coefficient ϕ is set to 1.0, the q-FCM narrows down to a classic FCM, where the activation values of neurons depend on the states of connected neurons in the previous iteration. In this setting, unique fixed points are frequent, meaning that the neurons' activation values in each iteration are independent of the initial conditions of the reasoning process. It is straightforward to conclude that any FCM-based model converging to a unique fixed point has limited usability when it comes to scenario analysis. In contrast, when $\phi = 0$, there is no recursion at all, and the model is reduced to an identity relation replicating the initial activation vector. The setting $\phi < 1$ guarantees that the initial conditions are explicitly taken into account when updating the neuron's activation values in each iteration [9]. More importantly, Nápoles et al. [9] demonstrated that q-FCM models using $\phi < 1$ will never converge to unique fixed attractors.

We should mention that selecting the right activation function is a key step when building FCM-based models. As highlighted in [34], the *sigmoid* function and the *hyperbolic tangent* are popular choices among researchers and practitioners alike. Additionally, functions such as *bivalent*, *trivalent*, and *threshold* have been utilized yet to a lesser extent. The former group exhibits continuous open intervals as their image set, while the latter has a discrete image set bounded into closed intervals. In general, any bounded and monotonically increasing function over the set of real numbers can be an activation function, given that the image set of a bounded function belongs to an interval. The *re-scaled* [33] and the *exponential normalized* [4] functions have also been recently introduced and hold promise; however, their performance remains relatively unexplored.

3. Shrink functions for quasi-nonlinear FCMs

In this section, we will first revisit the definitions introduced in [35] for classic FCMs as they will serve as the building blocks of our study. Subsequently, we will analyze the reachable activation values of neural concepts in each iteration of a q-FCM model. Finally, we resort to this theoretical basis to extend the shrink functions theory for the quasi-nonlinear rule, which will be the core of the algorithms proposed in Section 4.

3.1. Preliminaries on the state space estimation

Let $\mathcal{L}$ be the set of all non-negative closed intervals and let $\mathcal{S}^M$ be the set of all M-ary Cartesian products over the elements in $\mathcal{L}$. Formally, $\mathcal{S}^M = \{\mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_M : \mathcal{I}_i \in \mathcal{L}, \forall i \in \{1, 2, \ldots, M\}$. Every element in $\mathcal{S}^M$ is an M-ary Cartesian product of closed intervals.

Definition 1. *Let $\mathcal{I} \in \mathcal{L}$ and $\mathcal{I}' \in \mathcal{L}$. The interval $\mathcal{I}$ contains the interval $\mathcal{I}'$ (denoted by $\mathcal{I} \supseteq \mathcal{I}'$) if $\inf(\mathcal{I}) \leq \inf(\mathcal{I}') \wedge \sup(\mathcal{I}) \geq \sup(\mathcal{I}')$. Analogously, we say that the interval $\mathcal{I}$ strictly contains the interval $\mathcal{I}'$ (denoted by $\mathcal{I} \supset \mathcal{I}'$) if $\inf(\mathcal{I}) < \inf(\mathcal{I}') \wedge \sup(\mathcal{I}) > \sup(\mathcal{I}')$.*

Definition 2. *The closed interval $\mathcal{I}_i$ is the induced activation space for neuron C_i if it is the smallest closed interval containing the interval to which its associated activation function is bounded.*

Remark: It is important to observe that the activation function of the neuron C_i generates values within this interval. For instance, in FCM models using the sigmoid function, the activation values $a_i^{(t)}$ are confined to the $(0, 1)$ interval. Conversely, if a hyperbolic tangent function is adopted, these activation values will fall within the $(-1, 1)$ interval.

Definition 3. *The closed interval $\mathcal{I}_i^{(t)}$ is a feasible activation space at t-th iteration for C_i if the activation values for C_i always lie into $\mathcal{I}_i^{(t)}$ at t-th iteration. Formally, the closed interval $\mathcal{I}_i^{(t)}$ is a feasible activation space at t-th iteration for C_i if $a_i^{(t)} \in \mathcal{I}_i^{(t)}$.*

Remarks:

- The 0-th iteration corresponds to the initial activation values of neurons, representing their input values.
- If $\mathcal{I}_i'$ contains $\mathcal{I}_i^{(t)}$, then $\mathcal{I}_i'$ is also a feasible activation space at the t-th iteration for C_i. It is important to note that the feasible activation space for a given neuron C_i is not unique, suggesting the existence of several feasible activation spaces for the same neuron. Specifically, if $\mathcal{I}_i^{(t)}$ is a feasible activation space at the t-th iteration for neuron C_i, then any closed interval containing $\mathcal{I}_i^{(t)}$ is also a feasible activation space for C_i at the same iteration.

Definition 4. *The induced state space $\mathcal{S}$ is the M-ary Cartesian product over the induced activation spaces of all neurons. It is defined as $\mathcal{S} = \mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_M$, where $\mathcal{I}_i$ is the induced activation space for neuron C_i.*

Definition 5. *A feasible state space $\mathcal{S}^{(t)}$ at the t-th iteration is the M-ary Cartesian product over the feasible activation spaces of all neurons at the t-th iteration. It is defined as $\mathcal{S}^{(t)} = \mathcal{I}_1^{(t)} \times \mathcal{I}_2^{(t)} \times \ldots \times \mathcal{I}_M^{(t)}$, where $\mathcal{I}_i^{(t)}$ is a feasible activation space at the t-th iteration for neuron C_i. Formally, $\mathcal{S}^{(t)}$ is a feasible state space at the t-th iteration if $\mathbf{A^{(t)}} \in \mathcal{S}^{(t)}$.*

Remark: Definition 5 relies on Definition 3 to describe a particular state space of an FCM model. Elements in $\mathcal{S}^{(t)}$ are M-tuples, so we have $\mathcal{S}^{(t)} = \mathcal{I}_1^{(t)} \times \mathcal{I}_2^{(t)} \times \ldots \times \mathcal{I}_M^{(t)}$ and $\mathbf{A^{(t)}} = (a_1^{(t)}, a_2^{(t)}, \ldots, a_M^{(t)})$. Therefore, we can affirm that $\mathcal{S}^{(t)} \in \mathcal{S}^M$. Furthermore, it should be noticed that $\mathbf{A^{(t)}} \in \mathcal{S}^{(t)}$ is equivalent to stating that $a_i^{(t)} \in \mathcal{I}_i, \forall i \in \{1, 2, \ldots, M\}$.

Definition 6. *The state space $\mathcal{S} = \mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_M$ contains the state space $\mathcal{S}' = \mathcal{I}_1' \times \mathcal{I}_2' \times \ldots \times \mathcal{I}_M'$ if $\mathcal{I}_i$ contains $\mathcal{I}_i', \forall i \in \{1, 2, \ldots, M\}$. Formally, $\mathcal{S} \supseteq \mathcal{S}'$*

(state spaces are sets). Analogously, the state space $\mathcal{S} = \mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_M$ strictly contains the state space $\mathcal{S}' = \mathcal{I}'_1 \times \mathcal{I}'_2 \times \ldots \times \mathcal{I}'_M$ if $\mathcal{I}_i$ strictly contains $\mathcal{I}'_i, \forall i \in \{1, 2, \ldots, M\}$. Formally, $\mathcal{S} \supset \mathcal{S}'$.

3.2. Neurons' activation values in q-FCM models

The reasoning rule in Equation (1) is composed of two terms controlled by the nonlinearity coefficient (ϕ). If the q-FCM model uses the sigmoid activation function, then the analysis conducted in the next sections requires that neurons' activation values lie in the $[0, 1]$ interval. To generalize, we will prove for all t that $\mathbf{A^{(t)}} \in (x, y)^M$ or $\mathbf{A^{(t)}} \in [x, y]^M$ when the image set of the activation function is (x, y) or $[x, y]$, respectively. Therefore, the analysis holds for other activation functions as well.

Proof. In the same manner that $\mathbf{A}^{(0)} \in (0, 1)^M$ is considered a restriction in *sigmoid FCMs*, we have that $\mathbf{A}^{(0)} \in (x, y)^M$ or $\mathbf{A}^{(0)} \in [x, y]^M$ as a base case, depending on the activation function. Moreover, let us assume that $\mathbf{A^{(t)}} \in (x, y)^M$ and prove that $\mathbf{A^{(t+1)}} \in (x, y)^M$.

According to Equation (1), for every neuron C_i, its $(t + 1)$-th activation value is $a_i^{(t+1)} = \phi f_i\left(\sum_{j=1}^{M} w_{ji} a_j^{(t)}\right) + (1 - \phi) a_i^{(0)}$. Since the image set of $f_i(.)$ is (x, y), $a_i^{(0)} \in (x, y)$, and $0 \leq \phi \leq 1$ we have that:

$$\phi x < \phi f_i\left(\sum_{j=1}^{M} w_{ji} a_j^{(t)}\right) < \phi y,$$

and

$$(1 - \phi)x < (1 - \phi)a_i^{(0)} < (1 - \phi)y.$$

Combining both inequalities, we obtain:

$$\phi x + (1 - \phi)x < a_i^{(t+1)} < \phi y + (1 - \phi)y.$$

Finally, we can establish the following:

$$x < a_i^{(t+1)} < y. \tag{3}$$

As it holds for every neuron C_i, it is proved that $\mathbf{A^{(t+1)}} \in (x, y)^M$. In contrast, if the image set is closed, we can proceed analogously. The only changes needed are substituting open intervals with closed ones and strict inequalities ($<$ and $>$) with non-strict ones ($\leq$ and $\geq$). Then, assuming that $\mathbf{A^{(t)}} \in [x, y]^M$ we analogously prove that $\mathbf{A^{(t+1)}} \in [x, y]^M$.

□

3.3. Adjusted shrink functions for q-FCM models

The shrink functions were introduced in [35] as tools to estimate the feasible state spaces in FCMs. In that paper, the activation function could be seen as a real function whose input is the raw activation value at each iteration. Therefore, F is defined as the set of all monotonically increasing functions bounded into non-negative intervals. Hereinafter, we will refer to an F-function as any function belonging to F such that $F^0 \subset F$ and $F' \subset F$ are the subsets bounded into open and closed intervals, respectively. Moreover, $f_i \in F$ is the activation function used in the activation process of neuron C_i (i.e., every neuron has its activation function). This means that f_i is bounded into a non-negative interval (open or closed).

Let $\mathcal{G}$ be the set of all q-FCMs. In addition, let us consider two functions, $\mathcal{H}_W : \mathcal{G} \times \mathcal{S}^M \rightarrow \mathcal{S}^M$ and $\mathcal{H}_T : \mathcal{G} \times \mathcal{S}^M \rightarrow \mathcal{S}^M$, which take a q-FCM with M neurons and an M-ary Cartesian product of non-negative closed intervals and produce another M-ary Cartesian product of non-negative closed intervals. These functions are commonly referred to as *shrink functions*, because of the properties that were unveiled in [35]. Specifically, $\mathcal{H}_W$ and $\mathcal{H}_T$ are defined such that $\mathcal{H}_W(\mathcal{X}, \mathcal{S}^{(t)}) = \mathcal{S}^{(t+1)}$ and $\mathcal{H}_T(\mathcal{X}, \mathcal{S}^{(t)}) = \mathcal{S}^{(t+1)}$, respectively, where $\mathcal{X}$ is a q-FCM and $\mathcal{S}^{(t)} \in \mathcal{S}^M, \forall t$.

Assuming that $\mathcal{S}^{(t)}$ is a feasible state space at the t-th iteration, the functions $\mathcal{H}_W$ and $\mathcal{H}_T$ take an FCM model $\mathcal{X}$ and $\mathcal{S}^{(t)}$ to produce a feasible state space at the $(t+1)$-th iteration, denoted as $\mathcal{S}^{(t+1)}$. The difference between these functions lies in the fact that $\mathcal{H}_W$ uses the weight matrix to calculate $\mathcal{S}^{(t+1)}$, whereas $\mathcal{H}_T$ uses only the information regarding the connections between neurons (the topology of the FCM model). The terms $\inf_j^{(t)}$ and $\sup_j^{(t)}$ represent the bounds (infimum and supremum, respectively) of the closed interval $\mathcal{I}_j^{(t)}$ for neuron C_j. Then, given a feasible state space $\mathcal{S}^{(t)}$ at the t-th iteration, the following inequality holds:

$$\inf{}_j^{(t)} \leq a_j^{(t)} \leq \sup{}_j^{(t)}, \forall j. \tag{4}$$

In Equation (1), the dot product between w_i and $\mathbf{A^{(t)}}$ is of utmost importance to calculate the activation value of C_i. In this regard, it is assumed that every neuron is influenced by at least one other neural processing entity.

The activation values for input neurons are either unchanged or inactive, depending on the FCM implementation.

The quasi-nonlinear rule is a linear combination of an F-function and a constant term for each initial activation value. In this case, there are two inputs: the raw activation values passed through the activation function and the initial activation values. As a result, monotonicity cannot be addressed in the same way as in real-valued functions of a real variable [35], so the theory needs to be adapted. While the function associated with the quasi-nonlinear rule does not fall under the class of F-functions, it shares similar properties related to monotonicity that will be advantageous for the bounds estimation. Then, we assume that $\inf_i^{(0)} \leq a_i^{(0)} \leq \sup_i^{(0)}$ where $[\inf_i^{(0)}, \sup_i^{(0)}]$ is the induced activation space for the i-th neuron. Let us recall that ϕ and $1-\phi$ are non-negative numbers, and $f_i \in F$ is monotone-increasing. In this regard, two cases should be analyzed.

- Case 1: The weight matrix $\mathbf{W}$ is unknown

$$\begin{aligned} \inf{}_i^{(t+1)} &= \phi f_i\left(min_T\right) + (1-\phi)\inf{}_i^{(0)} \\ \sup{}_i^{(t+1)} &= \phi f_i\left(max_T\right) + (1-\phi)sup_i^{(0)} \\ \inf{}_i^{(t+1)} \leq a_i^{(t+1)} &\leq \sup{}_i^{(t+1)}, \forall i \end{aligned} \tag{5}$$

  such that the minimum value for the dot product is

$$\min_T \left(w_i \mathbf{A^{(t)}}\right) = -\sum_{j=1}^{M} sup_j^{(t)}, \tag{6}$$

  and the maximum value is

$$\max_T \left(w_i \mathbf{A^{(t)}}\right) = \sum_{j=1}^{M} sup_j^{(t)}. \tag{7}$$

- Case 2: The weight matrix $\mathbf{W}$ is known

$$\begin{aligned} \inf{}_i^{(t+1)} &= \phi f_i\left(min_W\right) + (1-\phi)\inf{}_i^{(0)} \\ \sup{}_i^{(t+1)} &= \phi f_i\left(max_W\right) + (1-\phi)sup_i^{(0)} \end{aligned}$$

$$\inf_i{}^{(t+1)} \le a_i^{(t+1)} \le \sup_i{}^{(t+1)}, \forall i \tag{8}$$

such that the minimum value for the dot product is given by

$$\min_W \left(w_i \mathbf{A^{(t)}}\right) = \sum_{j=1}^{M} \frac{w_{ji}\left(\sup_j^{(t)}(1 - sgn(w_{ji})) + \inf_j^{(t)}(1 + sgn(w_{ji}))\right)}{2}, \tag{9}$$

and the maximum value is given by

$$\max_W \left(w_i \mathbf{A^{(t)}}\right) = \sum_{j=1}^{M} \frac{w_{ji}\left(\sup_j^{(t)}(1 + sgn(w_{ji})) + \inf_j^{(t)}(1 - sgn(w_{ji}))\right)}{2}. \tag{10}$$

Note. The proofs for the maximum and minimum of the dot product between w_i and $\mathbf{A^{(t)}}$ are detailed in [35].

It should be mentioned that $\mathcal{H}_T$ and $\mathcal{H}_W$ return a feasible state space at the $(t+1)$-th iteration for the analyzed q-FCM by computing a feasible activation space $\mathcal{I}_i^{(t+1)} = [\inf_i^{(t+1)}, \sup_i^{(t+1)}]$ for each neuron C_i. The Cartesian product of all these feasible activation spaces allows obtaining a feasible state space $\mathcal{S}^{(t+1)} = \mathcal{I}_1^{(t+1)} \times \mathcal{I}_2^{(t+1)} \times \ldots \times \mathcal{I}_M^{(t+1)}$ at the $(t+1)$-th iteration. Therefore, we can conclude that both $\mathcal{H}_T$ and $\mathcal{H}_W$ functions maintain the feasibility of state spaces over the same q-FCM model.

4. Estimating the state spaces of q-FCM models

In this section, we will present a mathematical formalism to approximate the activation values of any concept in q-FCM models. These approximations consist of intervals defined by the lower and upper bounds of neurons' activation values in each iteration. They are also referred to as the limit state spaces. Such intervals may be either open or closed and will be used to derive the state space of a q-FCM model. To do that, we will consider the following settings: (i) unknown weight matrix and initial conditions, and (ii) known weight matrix and unknown initial conditions. In addition, these settings will be formalized into two algorithms that enable estimating the space state of any q-FCM with reasonable precision.

According to Definition 4, the initial induced state space is given by $\mathcal{S}^{(0)} = [\inf_1^{(0)}, \sup_1^{(0)}] \times [\inf_2^{(0)}, \sup_2^{(0)}] \times \ldots \times [\inf_M^{(0)}, \sup_M^{(0)}]$. In this regard, it can be stated that $g_i : \mathbb{R} \to [\inf_i^{(0)}, \sup_i^{(0)}]$ or $g_i : \mathbb{R} \to (\inf_i^{(0)}, \sup_i^{(0)}), \forall i \in \{1, 2, \ldots, M\}$. Using the shrink functions $\mathcal{H}_T$ and $\mathcal{H}_W$ for a given q-FCM, we can produce feasible activation spaces $\mathcal{S}^{(t+1)}$ from $\mathcal{S}^{(t)}, \forall t \in \mathbb{N}$. If we assume that we have $\mathcal{S}^{(0)} = [\inf_1^{(0)}, \sup_1^{(0)}] \times [\inf_2^{(0)}, \sup_2^{(0)}] \times \ldots \times [\inf_M^{(0)}, \sup_M^{(0)}]$, then $\mathcal{S}^{(1)}, \mathcal{S}^{(2)}, \mathcal{S}^{(3)}, \ldots$ can inductively be obtained.

Theorem 1 (Weak shrinking state space). *In a q-FCM, $\mathcal{S}^{(t)}$ contains $\mathcal{S}^{(t+1)}$, $\forall t \in \mathbb{N}$, when state spaces are iteratively calculated using either shrink function $\mathcal{H}_T$ or $\mathcal{H}_W$ with induced state space $\mathcal{S}^{(0)} = [\inf_1^{(0)}, \sup_1^{(0)}] \times [\inf_2^{(0)}, \sup_2^{(0)}] \times \ldots \times [\inf_M^{(0)}, \sup_M^{(0)}]$ and $f_i \in F', \forall i \in \{1, 2, \ldots, M\}$.*

Remarks:

- If a neuron has no incoming connections, its activation space across all iterations is determined by the induced one.
- $F' \subset F$ is the set of all monotonically increasing functions bounded into closed non-negative intervals.

Proof. Let $\mathcal{S}^{(t-1)} = [\inf_1^{(t-1)}, \sup_1^{(t-1)}] \times \ldots \times [\inf_M^{(t-1)}, \sup_M^{(t-1)}]$, $\mathcal{S}^{(t)} = [\inf_1^{(t)}, \sup_1^{(t)}] \times \ldots \times [\inf_M^{(t)}, \sup_M^{(t)}]$ and $\mathcal{S}^{(t+1)} = [\inf_1^{(t+1)}, \sup_1^{(t+1)}] \times \ldots \times [\inf_M^{(t+1)}, \sup_M^{(t+1)}]$. To prove that $\mathcal{S}^{(t)}$ contains $\mathcal{S}^{(t+1)}$, the fact that $[\inf_i^{(t)}, \sup_i^{(t)}]$ contains $[\inf_i^{(t+1)}, \sup_i^{(t+1)}]$ for every $i \in \{1, 2, \ldots, M\}$ must be demonstrated. Proceeding by induction, let us assume that $\mathcal{S}^{(t-1)}$ contains $\mathcal{S}^{(t)}$ and then prove that $\mathcal{S}^{(t)}$ contains $\mathcal{S}^{(t+1)}$. Therefore, given that $\mathcal{I}_i^{(t-1)}$ contains $\mathcal{I}_i^{(t)}$ for every $i \in \{1, 2, \ldots, M\}$, we will prove that $\mathcal{I}_i^{(t)}$ contains $\mathcal{I}_i^{(t+1)}$.

Since $\mathcal{S}^{(1)}$ is calculated using $\mathcal{S}^{(0)}$, the induced state space $\mathcal{S}^{(0)}$ contains every state space generated by shrink functions because bounds of $\mathcal{S}^{(0)}$ for every neuron match the activation function's bounds for this neuron. This implies that $\mathcal{S}^{(0)} \supseteq \mathcal{S}^{(t)}, \forall t$, and hence $\mathcal{S}^{(0)}$ contains $\mathcal{S}^{(1)}$.

- Case 1: The weight matrix $\mathbf{W}$ is unknown. The bounds for the dot product between w_i and $\mathbf{A^{(t-1)}}$ and between w_i and $\mathbf{A^{(t)}}$ are calculated using Equations (6) and (7), respectively. Therefore, we have that:

$$
\begin{aligned}
\mathcal{I}_i^{(t)} &= [\phi f_i(min_T(w_i \mathbf{A}^{(\mathbf{t-1})})) + (1-\phi)\inf_i^{(0)}, \\
&\qquad \phi f_i(max_T(w_i \mathbf{A}^{(\mathbf{t-1})})) + (1-\phi)\sup_i^{(0)}], \\
\mathcal{I}_i^{(t+1)} &= [\phi f_i(min_T(w_i \mathbf{A}^{(\mathbf{t})})) + (1-\phi)\inf_i^{(0)}, \\
&\qquad \phi f_i(max_T(w_i \mathbf{A}^{(\mathbf{t})})) + (1-\phi)\sup_i^{(0)}].
\end{aligned}
$$

Concepción et al. [35] proved that:

$$
\begin{aligned}
min_T(w_i \mathbf{A}^{(\mathbf{t-1})}) &\leq min_T(w_i \mathbf{A}^{(\mathbf{t})}), \\
max_T(w_i \mathbf{A}^{(\mathbf{t-1})}) &\geq max_T(w_i \mathbf{A}^{(\mathbf{t})}).
\end{aligned}
$$

We must ensure that the lower bound of $\mathcal{I}_i^{(t)}$ is less than or equal to the lower bound of $\mathcal{I}_i^{(t+1)}$, and also that the upper bound of $\mathcal{I}_i^{(t)}$ is greater than or equal to the upper bound of $\mathcal{I}_i^{(t+1)}, \forall i \in \{1, 2, \ldots, M\}$. Such a claim holds from the monotonically increasing property of $f_i \in F$, and the fact that ϕ, $1-\phi$, $\inf_i^{(0)}$ and $\sup_i^{(0)}$ are non-negative numbers.

- Case 2: The weight matrix $\mathbf{W}$ is known. The bounds for the dot product between w_i and $\mathbf{A}^{(\mathbf{t-1})}$ and between w_i and $\mathbf{A}^{(\mathbf{t})}$ are calculated using Equations (9) and (10), respectively.

  Let us define $\mathcal{S}^{(t)} = \mathcal{I}_1^{(t)} \times \mathcal{I}_2^{(t)} \times \ldots \times \mathcal{I}_M^{(t)}$ and $\mathcal{S}^{(t+1)} = \mathcal{I}_1^{(t+1)} \times \mathcal{I}_2^{(t+1)} \times \ldots \times \mathcal{I}_M^{(t+1)}$, such that $\mathcal{I}_i^{(t)}$ and $\mathcal{I}_i^{(t+1)}$ are given by:

$$
\begin{aligned}
\mathcal{I}_i^{(t)} &= [\phi f_i(min_W(w_i \mathbf{A}^{(\mathbf{t-1})})) + (1-\phi)\inf_i^{(0)}, \\
&\qquad \phi f_i(max_W(w_i \mathbf{A}^{(\mathbf{t-1})})) + (1-\phi)\sup_i^{(0)}], \\
\mathcal{I}_i^{(t+1)} &= [\phi f_i(min_W(w_i \mathbf{A}^{(\mathbf{t})})) + (1-\phi)\inf_i^{(0)}, \\
&\qquad \phi f_i(max_W(w_i \mathbf{A}^{(\mathbf{t})})) + (1-\phi)\sup_i^{(0)}].
\end{aligned}
$$

Concepción et al. [35] proved that:

$$min_W(w_i\mathbf{A^{(t-1)}}) \leq min_W(w_i\mathbf{A^{(t)}}),$$
$$max_W(w_i\mathbf{A^{(t-1)}}) \geq max_W(w_i\mathbf{A^{(t)}}).$$

We need to ensure that the lower bound of $\mathcal{I}_i^{(t)}$ is less than or equal to the lower bound of $\mathcal{I}_i^{(t+1)}$, and also that the upper bound of $\mathcal{I}_i^{(t)}$ is greater than or equal to the upper bound of $\mathcal{I}_i^{(t+1)}, \forall i \in \{1, 2, \ldots, M\}$. This holds from the monotonically increasing property of $f_i \in F$, and the fact that ϕ, $1-\phi$, $\inf_i^{(0)}$ and $\sup_i^{(0)}$ are non-negative numbers.

At this point, the induction thesis is proved for both cases (unknown and known weights), and the theorem holds. □

It should be mentioned that Theorem 3 asserts that the state spaces of q-FCM models shrink from one iteration to the next one, although it is possible that $\mathcal{S}^{(t)} = \mathcal{S}^{(t+1)}$, which implies that $\mathcal{S}^{(t)} = \mathcal{S}^{(t+k)}, \forall k \in \mathbb{N}$. If that happens, then the state spaces may not shrink forever.

Theorem 2 (Strong shrinking state space)**.** *In a q-FCM, $\mathcal{S}^{(t)}$ strictly contains $\mathcal{S}^{(t+1)}$, $\forall t \in \mathbb{N}$, when state spaces are iteratively calculated using either shrink function $\mathcal{H}_T$ or $\mathcal{H}_W$ with induced state space $\mathcal{S}^{(0)} = [\inf_1^{(0)}, \sup_1^{(0)}] \times [\inf_2^{(0)}, \sup_2^{(0)}] \times \ldots \times [\inf_M^{(0)}, \sup_M^{(0)}]$ and $f_i \in F^0, \forall i \in \{1, 2, \ldots, M\}$.*

Remarks:

- $F^0 \subset F$ is the set of all monotonically increasing functions bounded into open non-negative intervals.

- Notice that activation functions are now bounded into open intervals, which implies that the activation bounds $\inf_i^{(0)}$ and $\sup_i^{(0)}$ are never reachable at any iteration. This means that $\mathcal{S}^{(t)} \neq \mathcal{S}^{(t+k)}, \forall k \in \mathbb{N}$, and hence, the state spaces will shrink forever.

- To claim that $\mathcal{S}^{(t)}$ strictly contains $\mathcal{S}^{(t+1)}$, only neurons with incoming connections are relevant. Neurons without incoming connections are excluded from the analysis since their activation values are always the same. Depending on the q-FCM implementation, their activation values remain unchanged or inactive.

Proof. To prove that $\mathcal{S}^{(t)}$ strictly contains $\mathcal{S}^{(t+1)}$, we must prove that $[\inf_i^{(t)}, \sup_i^{(t)}]$ strictly contains $[\inf_i^{(t+1)}, \sup_i^{(t+1)}]$ $\forall i \in \{1, 2, \ldots, M\}$.

Let us assume that $\mathcal{S}^{(t-1)}$ strictly contains $\mathcal{S}^{(t)}$ and then prove that $\mathcal{S}^{(t)}$ strictly contains $\mathcal{S}^{(t+1)}$. Given that $\mathcal{I}_i^{(t-1)}$ strictly contains $\mathcal{I}_i^{(t)}$ $\forall i \in \{1, 2, \ldots, M\}$, we must prove that $\mathcal{I}_i^{(t)}$ strictly contains $\mathcal{I}_i^{(t+1)}$.

Since $\mathcal{S}^{(1)}$ is calculated using $\mathcal{S}^{(0)}$, the induction's base case can easily be verified. More explicitly, we can affirm that induced state space $\mathcal{S}^{(0)}$ strictly contains every state space generated by shrink functions because the bounds of $\mathcal{S}^{(0)}$ for every neuron match the activation function's bounds (open intervals), meaning that $\mathcal{S}^{(0)} \supset \mathcal{S}^{(t)}, \forall t$. This happens since, given two intervals with equal bounds, the closed one strictly contains the open one. Consequently, $\mathcal{S}^{(0)}$ strictly contains $\mathcal{S}^{(1)}$.

The proof is analogous to the weak version of the theorem, except that all inequalities are turned into strict ones. This means that every occurrence of the $\leq$ and $\geq$ symbols is replaced with the $<$ and $>$ symbols, respectively. Therefore, the strong version of the theorem is true. □

Motivated by the results in [35], we adapt the definition of limit state space in the context of q-FCM models. We must emphasize that the demonstration is analogous to the one in the original article.

Definition 7. *$\mathcal{S}^{(\infty)} \in \mathcal{S}^M$ is the limit state space of the q-FCM, when state spaces are iteratively calculated using either shrink function $\mathcal{H}_T$ or $\mathcal{H}_W$ and starting with $\mathcal{S}^{(0)}$, such that $\mathcal{S}^{(\infty)} = \lim_{t\to\infty} \mathcal{S}^{(t)}$.*

Algorithm 1 formalizes a deterministic procedure to estimate the limit state space of any q-FCM model. This algorithm requires as input the weight matrix W, the initial induced state space $S^{(0)}$, the maximum number of iterations allowed, and the ϕ value used in the quasi-nonlinear reasoning rule of the q-FCM model under study. Also, $\xi = 1.0\text{e-}5$ is taken as the minimal distance between two consecutive state spaces needed to stop the procedure before reaching the maximum number of iterations.

Algorithm 1 is designed to be independent of whether the `shrink_function` calculation incorporates the weights. Consequently, two supplementary pseudocode procedures describing the `shrink_function` are presented in Algorithms 2 and 3, using the weights and the topology of the q-FCM model, respectively. Observe that although the weight matrix is required as an input of the Algorithm 3, it is only used as an adjacency matrix to infer the topology of the q-FCM model under analysis.

Algorithm 1 Limit State Space estimation using weights or topology

Require: $\mathbf{W}$ $(M \times M)$, $S^{(0)}$ $(2 \times M)$, max_iters (int unsigned/positive), ϕ (float in $[0,1]$)

Ensure: Estimated limit state space (2xM)

  for $t \leftarrow 0$ to max_iters **do**

    $S^{(t+1)} \leftarrow$ shrink_function$(\mathbf{W}, S^{(t)}, S^{(0)}, \phi)$

    if distance$(S^{(t)}, S^{(t+1)}) < \xi$ **then**

      break

    end if

    $S^{(t)} \leftarrow S^{(t+1)}$

  end for

  return $S^{(t)}$

In short, our algorithms produce lower and upper bounds for the activation values of each neuron in a q-FCM model. Since these bounds are independent of the initial conditions, they allow assessing the model's capabilities before performing simulations. For example, these algorithms can help determine which activation values are impossible to produce. When $\phi = 1$, the algorithm may produce a limit state space of zero length, i.e. the lower and upper bounds of a concept are the same, indicating a unique fixed point attractor. If the q-FCM model is devoted to machine learning prediction tasks, such an undesirable state may irreversibly affect the model's prediction capabilities. Conversely, if $0 \leq \phi < 1$, the algorithm will produce a limit state space that does not contain an interval of zero length between the lower and upper bounds. This means that there is no unique fixed-point attractor for all initial activation values of the q-FCM model [33]. In this way, these algorithms may assist experts in understanding the expected model behavior when the input data become available and allow for more informed decisions about the dynamics of the system.

Algorithm 2 Shrink function: next state space calculation using weights

Require: $\mathbf{W}$ $(M \times M)$, $S^{(t)}$ $(2 \times M)$, $S^{(0)}$ $(2 \times M)$, ϕ (float in $[0, 1]$)

Ensure: Estimated next state space $(2 \times M)$

  for $i \leftarrow 0$ to $M - 1$ **do**

    $min\_val \leftarrow 0$

    $max\_val \leftarrow 0$

    for $j \leftarrow 0$ to $M - 1$ **do**

      if $w_{ji} \in \mathbf{W} \geq 0$ **then**

        $min\_val \leftarrow min\_val + w_{ji} \cdot S^{(t)}_{min}[j]$

        $max\_val \leftarrow max\_val + w_{ji} \cdot S^{(t)}_{max}[j]$

      else

        $min\_val \leftarrow min\_val + w_{ji} \cdot S^{(t)}_{max}[j]$

        $max\_val \leftarrow max\_val + w_{ji} \cdot S^{(t)}_{min}[j]$

    end if

  end for

  $S^{(t+1)}_{min}[i] \leftarrow$ quasi-nonlinear$(min_val, S^{(0)}_{min}[i], \phi)$

  $S^{(t+1)}_{max}[i] \leftarrow$ quasi-nonlinear$(max_val, S^{(0)}_{max}[i], \phi)$

end for

return $S^{(t+1)}$

Algorithm 3 Shrink function: next state space calculation using topology

Require: $\mathbf{W}$ $(M \times M)$, $S^{(t)}$ $(2 \times M)$, $S^{(0)}$ $(2 \times M)$, ϕ (float in $[0, 1]$)

Ensure: Estimated next state space $(2 \times M)$

```
for i ← 0 to M − 1 do
    min_val ← 0
    max_val ← 0
    for j ← 0 to M − 1 do
        if w_ij ∈ W ≠ 0 then
            min_val ← min_val + (−1) · S_max^(t)[j]
            max_val ← max_val + S_max^(t)[j]
        end if
    end for
    S_min^(t+1)[i] ← quasi-nonlinear(min_val, S_min^(0)[i], φ)
    S_max^(t+1)[i] ← quasi-nonlinear(max_val, S_max^(0)[i], φ)
end for
return S^(t+1)
```

5. Impact of the ϕ parameter on the state spaces

As discussed, the proposed algorithms devoted to approximating the state spaces of q-FCM models are a powerful tool that domain experts can use to avoid undesired configurations. In the quasi-nonlinear learning rule, the model's nonlinearity is controlled by the ϕ parameter. Since it is deemed a vital aspect of this reasoning rule, this section will study how sensitive the predicted state spaces are to this parameter and its actual influence. In this regard, we will use the concept of covering [35] in our mathematical analysis. Additionally, it must be highlighted that Definitions 8 and 9 are specified for general FCMs, but also applicable to q-FCMs.

Definition 8. *The covering of a feasible activation space at t-th iteration for neural concept C_i is the quotient between the associated interval's length (l_{fas}) and the length of its induced activation space (l_{ias}),*

$$covering(\mathcal{I}_i) = \begin{cases} \frac{l_{fas}}{l_{ias}} & if\ C_i \in \mathcal{N}_{in} \\ 0 & if\ C_i \notin \mathcal{N}_{in}. \end{cases}$$

such that $\mathcal{N}_{in}$ stands for the set of neurons with incoming connections. If the neuron is independent, the covering of every associated feasible activation space at any iteration is assumed to be zero.

This measure quantifies the percentage of the activation space covered by the activation values produced by a neuron. For example, a covering value of 0.3 indicates that the neuron's activation values span a maximum of 30% of the induced activation space. A covering value of zero means that, during the specified iteration, the neuron reaches a constant value (since a zero-length interval contains only a single value) and remains at that value thereafter, regardless of the initial activation value.

Definition 9. *In q-FCMs, the covering of a feasible state space at t-th iteration $\mathcal{S}^{(t)} = \mathcal{I}_1^{(t)} \times \mathcal{I}_2^{(t)} \times \ldots \times \mathcal{I}_M^{(t)}$ is the average covering of all feasible activation spaces at t-th iteration $\mathcal{I}_1^{(t)}, \mathcal{I}_2^{(t)}, \ldots, \mathcal{I}_M^{(t)}$. That is to say:*

$$\theta = \frac{1}{M} \sum_{i=1}^{M} covering(C_i).$$

Similarly, a covering value of 0.3 signifies that, on average, each neuron's activation value reaches at most 30% of its induced activation space. Roughly speaking, we can state that the tuple $\mathbf{A^{(t)}} = (a_1^{(t)}, a_2^{(t)}, \ldots, a_M^{(t)})$ covers, at most, 30% of its induced state space. A covering value of zero implies that the q-FCM will converge to a fixed-point attractor, regardless of the initial activation value used to start the reasoning process.

The concept of covering is related to the universal approximation property of multilayer feed-forward networks [36]. For instance, let us consider a neural network with a single output neuron that produces values in the $[0, 1]$ interval, applied to a prediction problem where the output values are expected to be close to either 0 or 1. In this setting, covering values significantly smaller than 1 suggest inadequate approximations of the given input-output set. Although values close to one do not guarantee accurate approximations, they may indicate improved model performance.

Theorem 3. *Let $\mathcal{S}^{(\phi^+,t)}$ and $\mathcal{S}^{(\phi^-,t)}$ be the state spaces produced by the same q-FCM model at the t-th iteration. It is assumed that the q-FCM's reasoning rule uses activation function $f_i \in F', \forall i \in \{1, 2, \ldots, M\}$, where ϕ^+ and ϕ^- represent the ϕ values used in the estimation of $\mathcal{S}^{(\phi,t)}$. If $\phi^+ > \phi^-$, then $\mathcal{S}^{(\phi^-,t)}$ contains $\mathcal{S}^{(\phi^+,t)}, \forall t \in \mathbb{N}^+$.*

Remarks:

- The state spaces are iteratively calculated using either the shrink function $\mathcal{H}_T$ or $\mathcal{H}_W$ with the induced state space $\mathcal{S}^{(\phi,0)} = [\inf_1^{(\phi,0)}, \sup_1^{(\phi,0)}] \times [\inf_2^{(\phi,0)}, \sup_2^{(\phi,0)}] \times \ldots \times [\inf_M^{(\phi,0)}, \sup_M^{(\phi,0)}]$.
- $\mathcal{I}_i^{(\phi,t)} = [\inf_i^{(\phi,t)}, \sup_i^{(\phi,t)}]$ for every neuron i and iteration t.

Proof. In order to simplify the process and account for varying ϕ values when estimating the dot product boundaries without considering weights, let us define the following equations:

$$min_T(\phi, t) = min_T(w_i \mathbf{A^{(t)}}),$$
$$max_T(\phi, t) = max_T(w_i \mathbf{A^{(t)}}).$$

For the case when we consider weights, we similarly define:

$$min_W(\phi, t) = min_W(w_i \mathbf{A^{(t)}}),$$
$$max_W(\phi, t) = max_W(w_i \mathbf{A^{(t)}}).$$

Additionally, the formulas involving $min_X(\phi, t)$ or $max_X(\phi, t)$ apply to cases where it is deemed irrelevant whether the values were estimated using the weights or not. However, it is important to clarify that in all formulas, these terms refer either to cases where the weights are known or unknown. Both cases cannot apply simultaneously.

The proof in Section 3.2 shows that $\mathbf{A^{(t)}} \in [x, y]^M$ for all t when the image set of the activation function is $[x, y]$ while ϕ can take any possible value. It must be recalled that the image set serves as the induced activation space of a neuron. Therefore, a given neural concept has the same induced activation space regardless of the value of ϕ. Consequently, the feasible activation spaces for the initial activation values are the induced ones, and $\mathcal{I}_i^{(\phi^-,0)} = \mathcal{I}_i^{(\phi^+,0)}, \forall i \in \{1, 2, \ldots, M\}$. Using this equality, let us prove that $\mathcal{I}_i^{(\phi^-,1)} \supseteq \mathcal{I}_i^{(\phi^+,1)}$. Regardless of the knowledge about weights and using Equations (6), (7), (9) and (10), we have the following formulas:

$$\begin{aligned}
\mathcal{I}_i^{(\phi^-,1)} = [&\phi^- f_i(min_X(\phi^-, 0)) + (1 - \phi^-) \inf_i^{(\phi^-,0)}, \\
&\phi^- f_i(max_X(\phi^-, 0)) + (1 - \phi^-) \sup_i^{(\phi^-,0)}], \\
\mathcal{I}_i^{(\phi^+,1)} = [&\phi^+ f_i(min_X(\phi^+, 0)) + (1 - \phi^+) \inf_i^{(\phi^+,0)}, \\
&\phi^+ f_i(max_X(\phi^+, 0)) + (1 - \phi^+) \sup_i^{(\phi^+,0)}]
\end{aligned}$$

The equality $\mathcal{I}_i^{(\phi^-,0)} = \mathcal{I}_i^{(\phi^+,0)}$ is equivalent to $\inf_i^{(\phi^-,0)} = \inf_i^{(\phi^+,0)}$ and $\sup_i^{(\phi^-,0)} = \sup_i^{(\phi^+,0)}$. It also implies that $min_X(\phi^-, 0) = min_X(\phi^+, 0)$ and $max_X(\phi^-, 0) = max_X(\phi^+, 0)$ because the parameter ϕ has no influence on the formulas. Hence, the first step consists of proving that the lower bound of $\mathcal{I}_i^{(\phi^-,1)}$ is less than or equal to the lower bound of $\mathcal{I}_i^{(\phi^+,1)}$ for all i, which holds true if the following inequality is satisfied:

$$
\begin{aligned}
&\phi^- \left(f_i(min_X(\phi^-, 0)) - \inf_i^{(\phi^-,0)} \right) \\
&\leq \phi^+ \left(f_i(min_X(\phi^+, 0)) - \inf_i^{(\phi^+,0)} \right).
\end{aligned} \tag{11}
$$

The terms inside parentheses are equal, and we only need to prove they are non-negative to verify the inequality. Since $f_i \in F^{'}$, this function always produces values greater than or equal to its lower bound $\inf_i^{(\phi^-,0)} = \inf_i^{(\phi^+,0)}$. Therefore, Equation (11) is validated.

Now, let us prove that the upper bound of $\mathcal{I}_i^{(\phi^-,1)}$ is greater than or equal to the upper bound of $\mathcal{I}_i^{(\phi^+,1)}$ for all i, which holds if:

$$
\begin{aligned}
&\phi^+ \left(\sup_i^{(\phi^+,0)} - f_i(max_X(\phi^+, 0)) \right) \\
&\geq \phi^- \left(\sup_i^{(\phi^-,0)} - f_i(max_X(\phi^-, 0)) \right).
\end{aligned} \tag{12}
$$

As stated before, the terms inside parentheses are equal, and we only need to prove they are non-negative to verify the inequality. Since $f_i \in F^{'}$, this function always produces values less than or equal to its upper bound $\sup_i^{(\phi^-,0)} = \sup_i^{(\phi^+,0)}$. Therefore, Equation (12) is validated.

Proceeding by induction, we hypothesize that $\mathcal{S}^{(\phi^-,t)}$ contains $\mathcal{S}^{(\phi^+,t)}, \forall t \in \mathbb{N}^+$, and then prove that $\mathcal{S}^{(\phi^-,t+1)}$ strictly contains $\mathcal{S}^{(\phi^+,t+1)}$. For the t-th iteration, we have the following:

$$
\begin{aligned}
\mathcal{S}^{(\phi^+,t)} &= \mathcal{I}_1^{(\phi^+,t)} \times \ldots \times \mathcal{I}_M^{(\phi^+,t)} \\
\mathcal{S}^{(\phi^-,t)} &= \mathcal{I}_1^{(\phi^-,t)} \times \ldots \times \mathcal{I}_M^{(\phi^-,t)},
\end{aligned}
$$

while for the $(t+1)$-th iteration we have:

$$
\begin{aligned}
\mathcal{S}^{(\phi^+,t+1)} &= \mathcal{I}_1^{(\phi^+,t+1)} \times \ldots \times \mathcal{I}_M^{(\phi^+,t+1)} \\
\mathcal{S}^{(\phi^-,t+1)} &= \mathcal{I}_1^{(\phi^-,t+1)} \times \ldots \times \mathcal{I}_M^{(\phi^-,t+1)}.
\end{aligned}
$$

Subsequently, we will prove that $\mathcal{I}_i^{(\phi^-,t)} \supseteq \mathcal{I}_i^{(\phi^+,t)}$ implies that $\mathcal{I}_i^{(\phi^-,t+1)} \supseteq \mathcal{I}_i^{(\phi^+,t+1)}$, for every $i \in \{1, 2, \ldots, M\}$. Using Equations (6), (7), (9) and (10), and assuming that the information concerning the weight matrix is not available, we can derive the following formulas:

$$\begin{aligned}
\mathcal{I}_i^{(\phi^-,t+1)} &= [\phi^- f_i(min_X(\phi^-, t)) + (1 - \phi^-) \inf_i^{(\phi^-,0)}, \\
&\qquad \phi^- f_i(max_X(\phi^-, t)) + (1 - \phi^-) \sup_i^{(\phi^-,0)}] \\
\mathcal{I}_i^{(\phi^+,t+1)} &= [\phi^+ f_i(min_X(\phi^+, t)) + (1 - \phi^+) \inf_i^{(\phi^+,0)}, \\
&\qquad \phi^+ f_i(max_X(\phi^+, t)) + (1 - \phi^+) \sup_i^{(\phi^+,0)}].
\end{aligned}$$

The relation $\mathcal{I}_i^{(\phi^-,t)} \supseteq \mathcal{I}_i^{(\phi^+,t)}$ is equivalent to the fact that $\inf_i^{(\phi^-,t)} \leq \inf_i^{(\phi^+,t)}$ and $\sup_i^{(\phi^-,t)} \geq \sup_i^{(\phi^+,t)}$. This relation also implies that $min_X(\phi^-, t) \leq min_X(\phi^+, t)$ and $max_X(\phi^-, t) \geq max_X(\phi^+, t)$. Therefore, we need to prove that the lower bound of $\mathcal{I}_i^{(\phi^-,t+1)}$ is less than or equal to the lower bound of $\mathcal{I}_i^{(\phi^+,t+1)}$ for all i, which holds if:

$$\begin{aligned}
&\phi^- f_i(min_X(\phi^-, t)) + (1 - \phi^-) \inf_i^{(\phi^-,0)} \\
&\leq \phi^+ f_i(min_X(\phi^+, t)) + (1 - \phi^+) \inf_i^{(\phi^+,0)}.
\end{aligned} \tag{13}$$

Let us establish that $\inf_i^{(0)} = \inf_i^{(\phi^-,0)} = \inf_i^{(\phi^+,0)}$, since ϕ has no influence in the induced activation values. After simplifying and grouping accordingly, we have the following inequality:

$$\begin{aligned}
&\phi^- \left(f_i(min_X(\phi^-, t)) - \inf_i^{(0)} \right) \\
&\leq \phi^+ \left(f_i(min_X(\phi^+, t)) - \inf_i^{(0)} \right).
\end{aligned} \tag{14}$$

The fact that $f_i(.)$ is monotone increasing and that $min_X(\phi^-, t) \leq min_X$ $(\phi^+, t))$ implies that $f_i(min_X(\phi^-, t)) - \inf_i^{(0)}$ is less than or equal to $f_i(min_X$ $(\phi^+, t)) - \inf_i^{(0)}$. Moreover, these terms are non-negative since $f_i \in F'$ always produces values greater than or equal to $\inf_i^{(0)}$. These factors, combined with the premise $\phi^- < \phi^+$, validate Equation (13).

Next, we will prove that the upper bound of $\mathcal{I}_i^{(\phi^-, t+1)}$ is greater than or equal to the upper bound of $\mathcal{I}_i^{(\phi^+, t+1)}$, which holds if:

$$
\begin{aligned}
&\phi^- f_i(max_X(\phi^-, t)) + (1 - \phi^-) \sup_i^{(\phi^-, 0)} \\
&\geq \phi^+ f_i(max_X(\phi^+, t)) + (1 - \phi^+) \sup_i^{(\phi^+, 0)} .
\end{aligned} \tag{15}
$$

As we did before, let us establish that $\sup_i^{(0)} = \sup_i^{(\phi^-, 0)} = \sup_i^{(\phi^+, 0)}$, since ϕ has no influence in the induced activation values. After simplifying and grouping accordingly, we have the following:

$$
\begin{aligned}
&\phi^+ \left(\sup_i^{(0)} - f_i(max_X(\phi^+, t)) \right) \\
&\geq \phi^- \left(\sup_i^{(0)} - f_i(max_X(\phi^-, t)) \right) .
\end{aligned} \tag{16}
$$

Since $f_i(.)$ is monotone increasing and $max_X(\phi^-, t) \geq max_X(\phi^+, t)$, it holds that $\sup_i^{(0)} - f_i(max_X(\phi^+, t))$ is greater than or equal to $\sup_i^{(0)} - f_i(max_X$ $(\phi^-, t))$. These terms are also non-negative since $f_i \in F'$ produces values less than or equal to its upper bound $\sup_i^{(0)}$. Combining these remarks with the fact that $\phi^+ > \phi^-$ allows us to validate Equation (15). As such, the thesis is proved by induction and the theorem holds. $\square$

Theorem 4. *Let $\mathcal{S}^{(\phi^+, t)}$ and $\mathcal{S}^{(\phi^-, t)}$ be state spaces produced by the same q-FCM model at the t-th iteration. It is assumed that the q-FCM's reasoning rule uses activation functions $f_i \in F^0, \forall i \in \{1, 2, \ldots, M\}$, where ϕ^+ and ϕ^- are the ϕ values used in the estimation of $\mathcal{S}^{(\phi, t)}$. If $\phi^+ > \phi^-$, then $\mathcal{S}^{(\phi^-, t)}$ strictly contains $\mathcal{S}^{(\phi^+, t)}, \forall t \in \mathbb{N}^+$.*

Remark: As previously mentioned, the activation spaces of neurons with no incoming connections remain the induced ones throughout all iterations.

Therefore, to assert that $\mathcal{S}^{(\phi^-,t)}$ strictly contains $\mathcal{S}^{(\phi^+,t)}$, only neurons with incoming connections are relevant for the analysis.

Proof. The proof is analogous to Theorem 3, except that all inequalities, subset, and superset relationships are turned into strict ones. This means that every occurrence of the $\leq$, $\geq$, $\subseteq$ and $\supseteq$ symbols is replaced with the $<$, $>$, $\subset$ and $\supset$ symbols, respectively. Additionally, when referring to non-negative terms, we must state that they are positive. The only exception to the previous rule is that the inequality $\phi < \phi^+$ remains unchanged. Therefore, the strong version of the theorem holds true. $\square$

Theorems 3 and 4 indicate that the covering values decrease (or at least do not increase) as ϕ approaches its maximum value. In contrast, smaller ϕ values lead to larger covering values at the expense of harming the model's nonlinearity. This behavior suggests the existence of a trade-off between the number of reachable states and the model's nonlinearity, which is deemed key when solving pattern classification, multi-output regression, or time series forecasting problems. Overall, these insights highlight the complex relationship between ϕ and the contraction of the state space.

6. Numerical simulations

In this section, we will conduct a two-fold experimental study to assess the correctness of our algorithms. The first experiment uses real-world case studies to exemplify the impact of the ϕ parameter on the state spaces of q-FCM models. More importantly, we will illustrate the effectiveness of the proposed algorithms to estimate the limit state space of q-FCM models even when no data is available. The second experiment relies on synthetically generated q-FCM to further evaluate the precision of the estimated limit state spaces compared to the actual activation values.

6.1. Simulations using real-world case studies

Aiming to study the relation between the ϕ parameter and the estimated state spaces, we will adopt three real-world case studies. These cognitive networks represent well the structural complexity and network density of FCM-based models designed by domain experts.

The crime and punishment model (see Figure 1), proposed by Mohr [37], has been extensively used to evaluate the correctness of new algorithms and methodologies [38, 39]. This network is devoted to modeling the effects of several coupled social attributes on the prevalence of theft in a given community. The concepts in this model are the *presence of property* ($C1$), *opportunity* ($C2$), *theft* ($C3$), *community intervention* ($C4$), *criminal intention* ($C5$), *punishment* ($C6$), and *police presence* ($C7$).

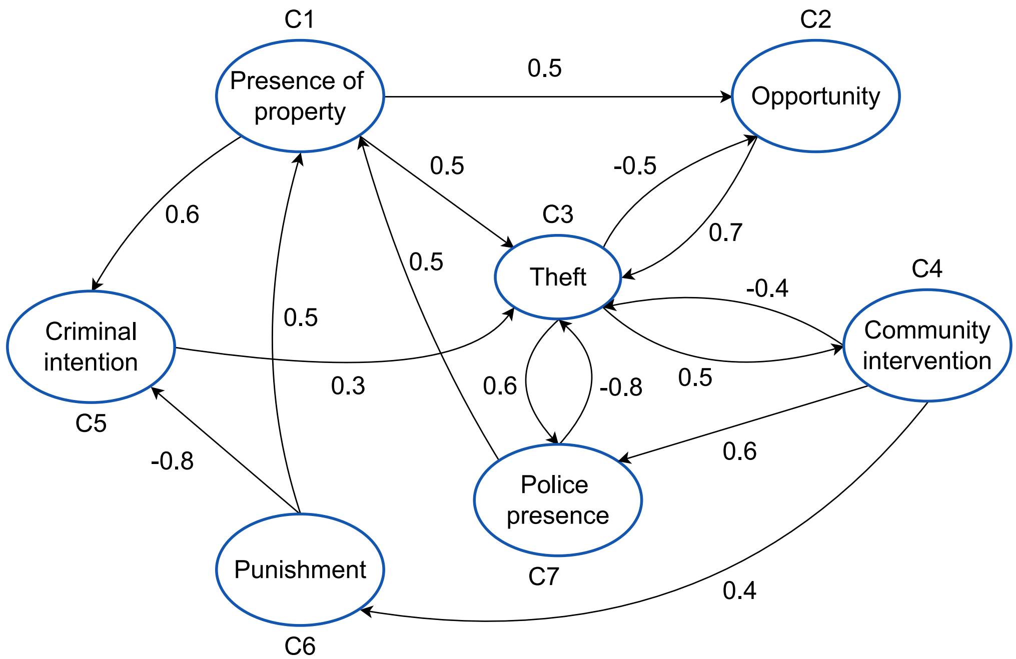


Figure 1: FCM model for the “Crime and Punishment” case study.

The second case study concerns civil engineering and investigates the implications of population growth and urban development on the public health of a city (refer to Figure 2). This network was employed in [40] to assess the inferential capabilities of binary, trivalent, and *sigmoid FCM* models. The concepts in this model are *people in a city* ($C1$), *migration into city* ($C2$), *modernization* ($C3$), *amount of garbage* ($C4$), *sanitation facilities* ($C5$), *diseases per 1000 residents* ($C6$), and *bacteria per area* ($C7$).

The third case study depicts the concepts and causal relationships in a system modeling a car sales company, as sourced from [2] (see Figure 3). The neural concepts describing this system are *high profits* ($C1$), *customer satisfaction* ($C2$), *high sales* ($C3$), *union raises* ($C4$), *safer cars* ($C5$), *foreign competition* ($C6$), and *lower prices* ($C7$).

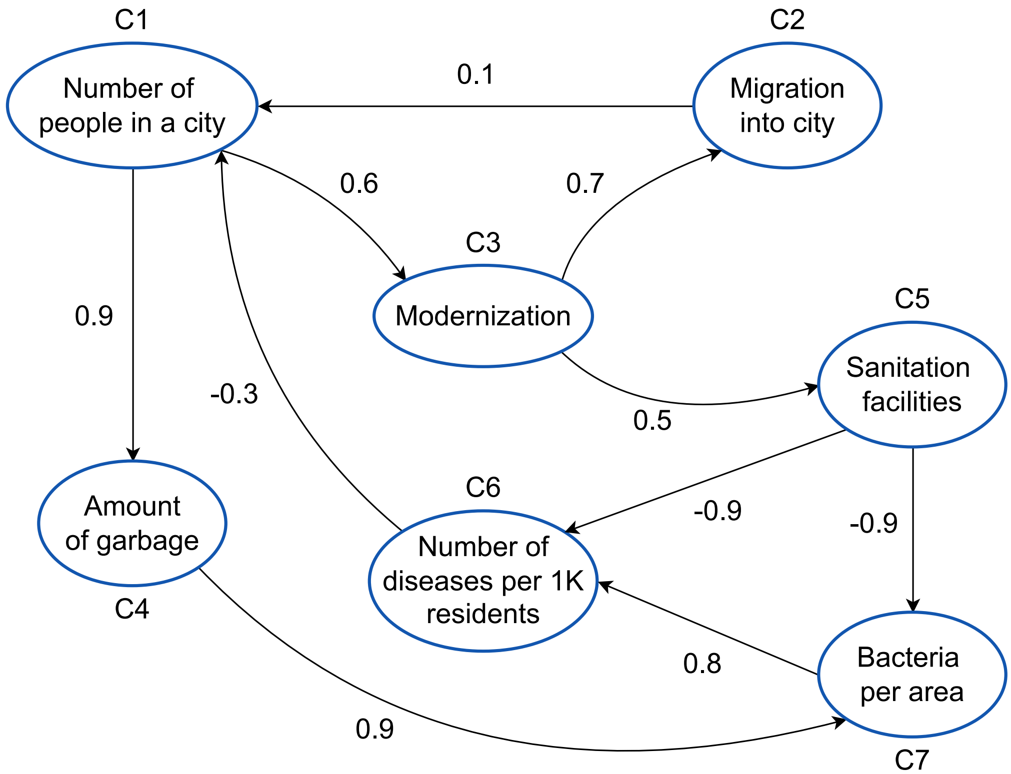


Figure 2: FCM model for the “Public Health System” case study.

Aiming to study the effect of ϕ on the covering values, we randomly generated 1,000 initial activation vectors for each case study. It must be stressed that generated data is used for validation purposes only since the proposed algorithms operate without knowledge of the initial conditions. Moreover, for each problem, we built several q-FCM models resulting from varying ϕ from 0.0 to 1.0, such that $\phi = 1.0$ represents the classic FCM formalism. After activating the neural concepts, the reasoning process is performed for $T = 10$ iterations. In this experiment, the coverage of each neural concept is calculated with and without considering the information about the weight matrix. Figure 4 illustrates how the mean coverage decreases as ϕ increases, whether the weights are included or only the topology is considered. This experiment also demonstrates that incorporating the weights allows for a more precise calculation of the mean coverage.

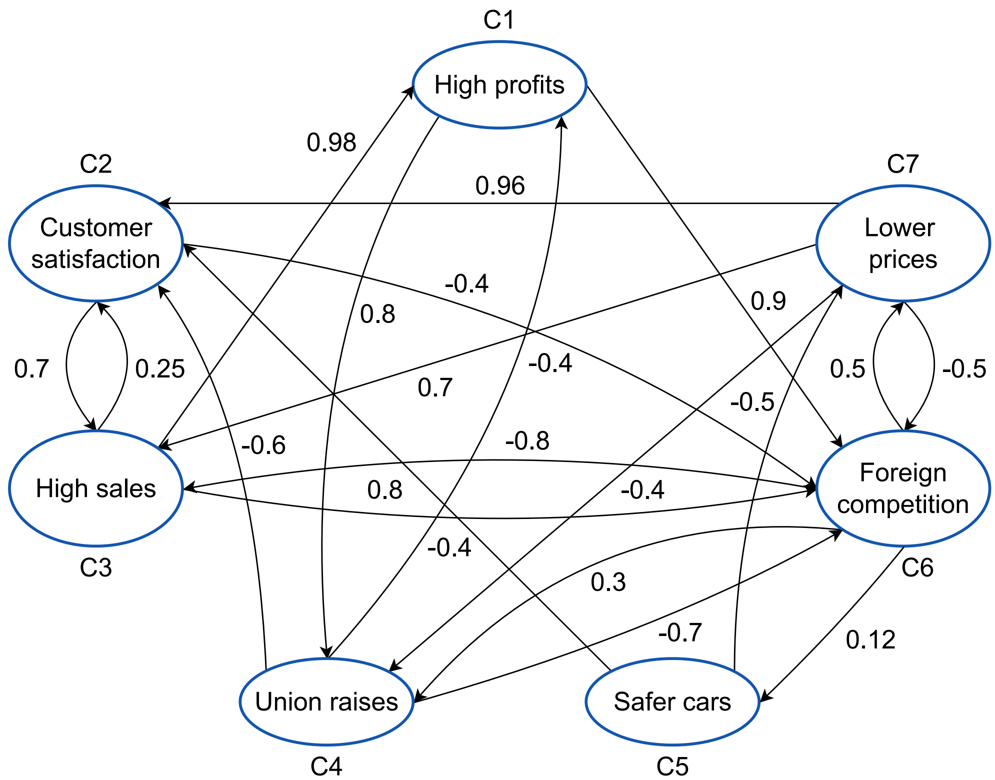
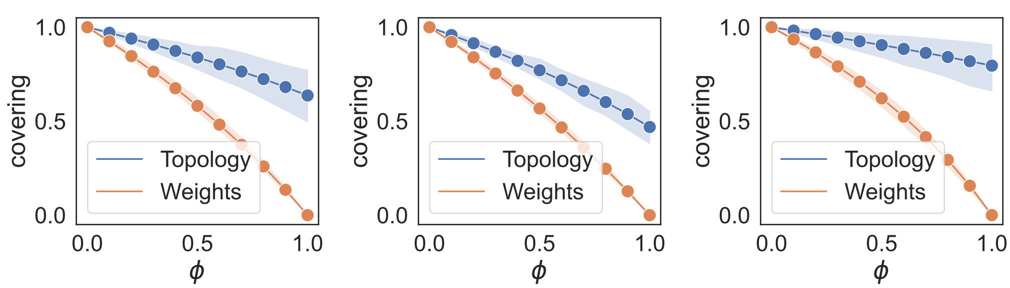


Figure 3: FCM model for the “Car Sales Company” case study.



(a) “Crime and Punishment” case study.

(b) “Public Health System” case study.

(c) “Car Sales Company” case study.

Figure 4: Covering results when varying the ϕ parameter with and without knowing the weight matrix. In both scenarios, the initial conditions are unknown.

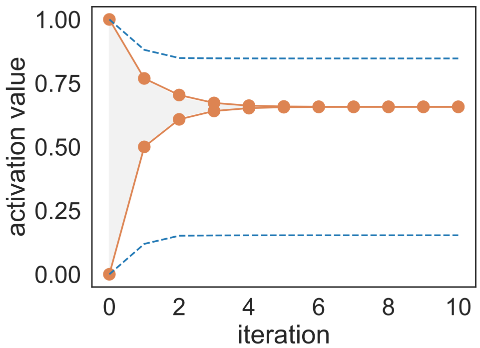
Next, let us compare the activation space of arbitrarily selected neurons throughout the reasoning process using the generated initial activation vectors. For simplicity, we will consider only two models: q-FCM ($\phi = 1.0$) and q-FCM ($\phi = 0.8$). Figure 5 shows the activation values of concept $C7$ for the three case studies. The gray area represents the real activation values at each iteration after using the generated input data to execute the reasoning. The solid orange lines indicate the lower and upper bounds of that concept for each iteration produced by our algorithm when using the weight information. The dashed blue lines indicate the lower and upper bounds computed using only information about the network topology. As expected, the gray area remains within the boundaries defined by the solid lines in every case. Note that for a different set of inputs, the gray area could shrink further but will always remain within the boundaries calculated by our algorithms. Additionally, this experiment illustrates how the activation space decreases through iterations, supporting Theorems 1 and 2.

A closer inspection of the FCM models in Figure 5 reveals that the lower and upper bounds using the weight information indicate the presence of a unique fixed-point attractor, regardless of the initial conditions. This implies that in classic FCM models, our algorithms can be a useful tool for predicting a unique fixed-point attractor without knowledge of the initial input. In contrast, this is not the case for q-FCM models, as $\phi < 1.0$ guarantees that there is no unique fixed-point attractor.

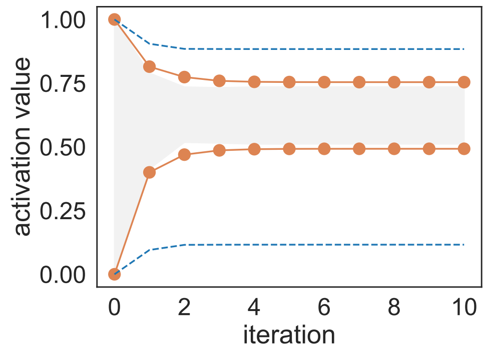
Overall, the experiments using real-world case studies illustrate how the lower and upper bounds computed by our algorithms approximate the actual activation values produced by the neurons. This capability demonstrates that even in the absence of specific input values, we can predict the potential outcomes of the concepts within q-FCM models, reinforcing the robustness of our state space estimation approach.

6.2. Simulations using synthetic q-FCM models

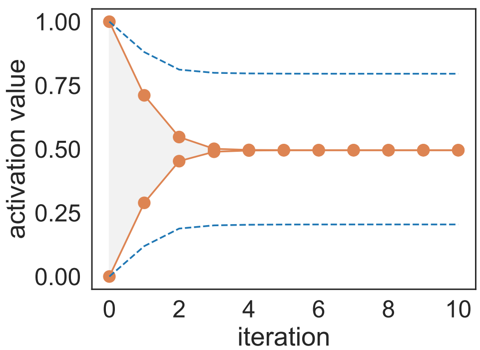
To further analyze the precision of the estimated limit state spaces, let us generate 1,000 synthetic q-FCM models. The number of neural concepts in these models varies from 5 and 20 concepts, while the sigmoid activator is used as the activation function. The connectivity is set to 50%, where connectivity refers to the ratio between the number of non-zero weights and the maximal number of relationships. In our simulations, the ϕ parameter ranges from 0 to 1, and the reasoning rule is executed until the model converges to its limit state space such that $\xi = 1.0\text{e-}5$.



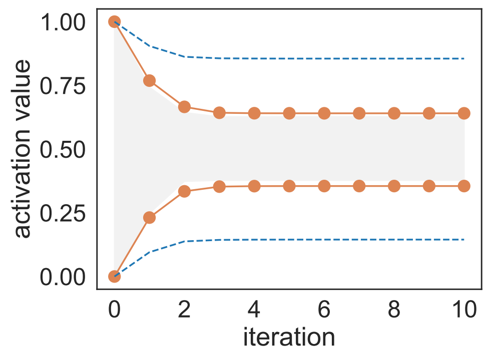
(a) "Crime and Punishment" case study with selected concept "police presence" ($C7$). The nonlinearity coefficient is $\phi = 1$.



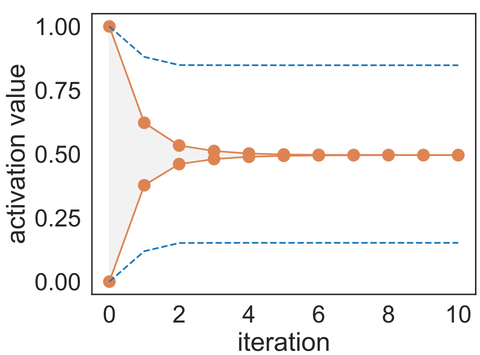
(b) "Crime and Punishment" case study with selected concept "police presence" ($C7$). The nonlinearity coefficient is $\phi = 0.8$.



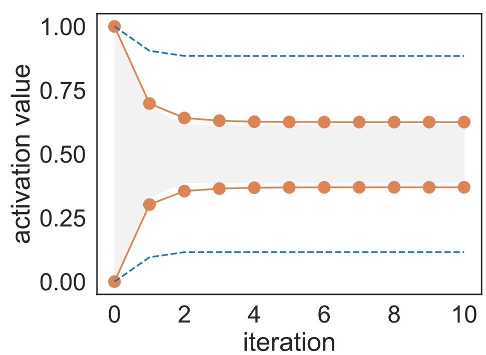
(c) "Public Health System" case study with selected concept "bacteria per area" ($C7$). The nonlinearity coefficient is $\phi = 1$.



(d) "Public Health System" case study with selected concept "bacteria per area" ($C7$). The nonlinearity coefficient is $\phi = 0.8$.



(e) "Car Sales Company" case study with selected concept "lower prices"' ($C7$). The nonlinearity coefficient is $\phi = 1$.



(f) "Car Sales Company" case study with selected concept "lower prices" ($C7$). The nonlinearity coefficient is $\phi = 0.8$.

Figure 5: Estimated feasible activation space of neuron $C7$ for the three real-world case studies with (solid) and without (dashed) using the weight information. The gray area represents the actual activation values.

In this experiment, we introduce a new performance metric called the *relative average gap*, which quantifies how closely the predicted bounds match the actual activation values of the neurons. Let $\max(a_i^{(t)})$ and $\min(a_i^{(t)})$ represent the maximum and minimum values of $a_i^{(t)}$ across all available initial conditions, corresponding to the ideal bounds for the data. Additionally, it is worth recalling that $\sup_i^{(t)}$ and $\inf_i^{(t)}$ denote the upper and lower bounds for the i-th feasible activation space, respectively.

Definition 10. *The relative average gap ($rag_i^{(t)}$) of a feasible activation space $\mathcal{I}_i^{(t)}$ at the t-th iteration for neural concept C_i is the quotient between the average gap around the data extreme values and the length of the induced activation space (l_{ias}). That is to say:*

$$rag_i^{(t)} = \begin{cases} \dfrac{\sup_i^{(t)} - max(a_i^{(t)}) + min(a_i^{(t)}) - \inf_i^{(t)}}{2\ l_{ias}} & \text{if } C_i \in \mathcal{N}_{in} \\ 0 & \text{if } C_i \notin \mathcal{N}_{in} \end{cases}$$

such that $\mathcal{N}_{in}$ stands for the set of neurons with incoming connections. If the neuron is independent, the relative average gap of every associated feasible activation space at any iteration is assumed to be zero.

If $rag_i^{(t)} \approx 0$, the interval length is approximately equal to the data range, indicating an accurate approximation. If $rag_i^{(t)} \gg 0$, the interval is wider than the actual data range, suggesting that the interval has excess space. Notice that $rag_i^{(t)} < 0$ is not possible, since the feasible activation spaces always contain the activation values from the data. The relative average gap can also be computed at the model level.

Definition 11. *In q-FCMs, the relative average gap of a feasible state space at t-th iteration $\mathcal{S}^{(t)} = \mathcal{I}_1^{(t)} \times \ldots \times \mathcal{I}_M^{(t)}$ is the average $rag_i^{(t)}$ of all feasible activation spaces at t-th iteration $\mathcal{I}_1^{(t)}, \ldots, \mathcal{I}_M^{(t)}$. That is to say:*

$$rag^{(t)} = \frac{1}{M} \sum_{i=1}^{M} rag_i^{(t)}.$$

As before, $rag^{(t)} \approx 0$ indicates that the intervals accurately approximate the minimum and maximum activation values produced by all neural concepts on average. At this point, we have the tools to assess the quality of the limit state space estimations for q-FCMs models.

Let us randomly generate 1,000 initial activation vectors, apply the reasoning rule to the generated q-FCM models under study, and compute the relative average gap in the final iteration. This metric is averaged across all initial conditions and topology configurations for each ϕ value. Additionally, the experiment is conducted for state space estimations using either the weight information or only the network topology information. Figure 6 portrays the simulation results, with the standard deviation shown as a shaded area around the average values in each series.

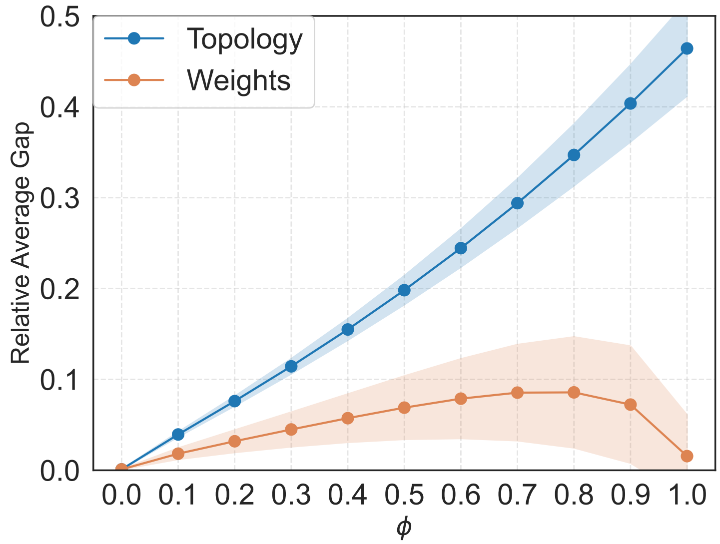


Figure 6: Relative average gap when varying the ϕ parameter, with and without considering the weight information when estimating the bounds.

The relative average gaps in Figure 6 illustrate the substantial impact of using the weight information on the quality of the state space estimations. While the topology information alone provides a reasonable level of precision, integrating the weight information leads to a consistently low relative average gap below 0.15. Since the covering is maximized when $\phi \approx 0$ (see Theorem 4), it is reasonable to conclude that the small relative average gaps stem from the ease of estimating larger state spaces. For the estimations using weights, the relative average gap increases as ϕ approaches 0.8. This behavior indicates that estimating the state spaces becomes more challenging while the covering decreases. However, after this point, the gap begins to decrease until $\phi = 1$, where the q-FCMs align with classic FCMs. When $\phi = 1$, the q-FCM model exhibits the lowest covering values, allowing the algorithm to potentially detect unique fixed-point attractors.

In contrast, when only the network topology is known, the relative average gap increases almost linearly with ϕ, leading to a steady decrease in precision. In this case, while the available information about the q-FCM model is quite limited, we can still gain some insights into the model's behavior, reflected in an overall average gap of 0.21. However, as discussed, using the weight information significantly enhances state space estimation precision, reducing the average gap to 0.05 across all configurations. These findings underscore that the estimated bounds closely match the actual outcomes, especially when the weight information is considered.

7. Concluding remarks

In this paper, we introduced two mathematically grounded algorithms to estimate the limit state spaces of q-FCM models without knowing the input data. Based solely on the structural information of the cognitive network, these algorithms produce reasonably precise upper and lower bounds for the activation values of neural concepts in each iteration. Therefore, the model will never produce activation values outside these bounds, regardless of the initial conditions used to perform reasoning. In addition, the proposed algorithms can detect whether a classic FCM model will converge to a fixed-point attractor. We further demonstrated mathematically that as the nonlinearity coefficient of the q-FCM model approaches one, the covering of neural concepts shrinks. This behavior reveals a trade-off between the model's nonlinearity and the range of achievable states.

In our numerical simulations, we validated the correctness of the proposed algorithms using both real-world and synthetically generated problems. More explicitly, we explored the behavior of q-FCMs across three case studies of different domains, using the weight matrix and only the network topology. The results showed that the covering of the state space consistently decreased as the nonlinearity coefficient increased, regardless of whether the weight matrix or topology was used. This aligns with our theoretical findings, reinforcing the statement that higher ϕ values reduce the number of reachable states. Moreover, the experiments confirmed that when $\phi = 1$, the q-FCM behaves as a classic FCM that often leads to a fixed-point attractor, as evidenced by the activation space bounds shrinking to a single point in some cases. For lower values of ϕ, the q-FCM exhibited more complex dynamics, avoiding the fixed-point attractor and producing a more diverse range of activation values. This was particularly evident in scenarios where weights were considered, as the predicted activation bounds closely matched the actual outcomes. These findings demonstrate the utility of the proposed algorithms in estimating the state space under varying configurations.

Despite the ground-breaking results, the proposed algorithms might suffer from scalability issues. In topologies with many interconnected concepts, the precision of the estimated bounds can decrease, leading to less accurate predictions. However, this limitation is mitigated in practice, as FCM models devoted to real-world scenario simulation typically involve relatively simple structures. Looking forward, a promising direction for future research is to investigate the effectiveness of our algorithms in detecting unique fixed points, particularly by comparing them with other mathematical techniques in the literature. Additionally, we plan to extend our work by developing a learning algorithm based on the proposed weight-driven state space estimation, aiming to bypass the need for input data and offering a new framework for data-free learning in FCM-based models.

Acknowledgments

Y. Salgueiro would like to acknowledge the support provided by ANID Fondecyt Regular 1240293 and the super-computing infrastructure of the NLHPC (ECM-02). L. Concepción was supported by the Special Research Fund (BOF20BL04) from UHasselt, Belgium. Similarly, G. Nápoles would like to acknowledge to support received from the Special Research Fund (BOF24KV18) from UHasselt, Belgium.

References

[1] B. Kosko, Fuzzy cognitive maps, International journal of man-machine studies 24 (1) (1986) 65–75. doi:https://doi.org/10.1016/S0020-7373(86)80040-2.

[2] A. Tsadiras, N. Bassiliades, RuleML representation and simulation of Fuzzy Cognitive Maps, Expert Systems with Applications 40 (5) (2013) 1413–1426. doi:https://doi.org/10.1016/j.eswa.2012.08.035.

[3] V. C. Georgopoulos, C. D. Stylios, Supervisory Fuzzy Cognitive Map Structure for Triage Assessment and Decision Support in the Emergency Department, in: M. S. Obaidat, S. Koziel, J. Kacprzyk, L. Leifsson, T. Ören (Eds.), Simulation and Modeling Methodologies, Technologies and Applications, Springer International Publishing, Cham, 2015, pp. 255–269. doi:https://doi.org/10.1007/978-3-319-11457-6_18.

[4] G. Nápoles, I. Grau, Y. Salgueiro, A revised cognitive mapping methodology for modeling and simulation, Knowledge-Based Systems 299 (2024) 112089. doi:https://doi.org/10.1016/j.knosys.2024.112089.

[5] A. A. Petunin, C. Stylios, Optimization Models of Tool Path Problem for CNC Sheet Metal Cutting Machines, in: IFAC-PapersOnLine, Vol. 49, 2016, pp. 23–28. doi:10.1016/j.ifacol.2016.07.544.

[6] W. Xiaojie, L. Chao, L. Chen, The feedback stabilization of finite-state fuzzy cognitive maps, Transactions of the Institute of Measurement and Control 44 (13) (2022) 2485–2499. doi:https://doi.org/10.1177/01423312221085785.

[7] L. Concepción, G. Nápoles, I. Grau, W. Pedrycz, Fuzzy-Rough Cognitive Networks: Theoretical Analysis and Simpler Models, IEEE Transactions on Cybernetics 52 (2020) 2994–3005. doi:https://doi.org/10.1109/tcyb.2020.3022527.

[8] P. Szwed, Classification and feature transformation with Fuzzy Cognitive Maps, Applied Soft Computing 105 (2021) 107271. doi:https://doi.org/10.1016/j.asoc.2021.107271.

[9] G. Nápoles, Y. Salgueiro, I. Grau, M. L. Espinosa, Recurrence-Aware Long-Term Cognitive Network for Explainable Pattern Classification, IEEE Transactions on Cybernetics 53 (10) (2023) 6083–6094. doi:https://doi.org/10.1109/TCYB.2022.3165104.

[10] M. Tyrovolas, X. S. Liang, C. Stylios, Information flow-based fuzzy cognitive maps with enhanced interpretability, Granular Computing 8 (2023) 2021–2038. doi:https://doi.org/10.1007/s41066-023-00417-7.

[11] G. Nápoles, A. Jastrzebska, I. Grau, Y. Salgueiro, Backpropagation through time learning for recurrence-aware long-term cog-

nitive networks, Knowledge-Based Systems 295 (2024) 111825. doi:https://doi.org/10.1016/J.KNOSYS.2024.111825.

[12] G. Nápoles, F. Vanhoenshoven, K. Vanhoof, Short-term cognitive networks, flexible reasoning and nonsynaptic learning, Neural Networks 115 (2019) 72–81. doi:https://doi.org/10.1016/j.neunet.2019.03.012.

[13] B. Yang, S. Hong, H. Wang, D. Zhao, Complex Dynamic Slow Independent Symbol Aggregation Approximation for Multimodal Temporal Fuzzy Cognitive Map, IEEE Transactions on Industrial Informatics (2024) 1–9doi:https://doi.org/10.1109/TII.2024.3427419.

[14] D. Qin, Z. Peng, L. Wu, Interpretable predictive modeling of non-stationary long time series, Computers & Industrial Engineering 194 (2024) 110412. doi:https://doi.org/10.1016/J.CIE.2024.110412.

[15] G. Nápoles, I. Grau, A. Jastrzębska, Y. Salgueiro, Long short-term cognitive networks, Neural Computing and Applications 34 (19) (2022) 16959–16971. doi:https://doi.org/10.1007/s00521-022-07348-5.

[16] C. Garatejo, W. Hoyos, J. Aguilar, An Approach Based on Fuzzy Cognitive Maps with Federated Learning to Predict Severity in Viral Diseases, in: Proceedings - 2023 49th Latin American Computing Conference, CLEI 2023, 2023. doi:https://doi.org/10.1109/CLEI60451.2023.10346147.

[17] W. Hoyos, J. Aguilar, M. Toro, Federated learning approaches for fuzzy cognitive maps to support clinical decision-making in dengue, Engineering Applications of Artificial Intelligence 123 (2023). doi:https://doi.org/10.1016/j.engappai.2023.106371.

[18] J. L. Salmeron, I. Arévalo, Blind Federated Learning without initial model, Journal of Big Data 11 (1) (2024) 56. doi:https://doi.org/10.1186/s40537-024-00911-y.

[19] J. L. Salmeron, I. Arévalo, Concurrent vertical and horizontal federated learning with fuzzy cognitive maps, Future Generation Computer Systems 162 (2025) 107482. doi:https://doi.org/10.1016/J.FUTURE.2024.107482.

[20] G. Nápoles, N. Ranković, Y. Salgueiro, On the interpretability of Fuzzy Cognitive Maps, Knowledge-Based Systems 281 (2023) 111078. doi:https://doi.org/10.1016/j.knosys.2023.111078.

[21] D. Benito, C. Quintero, J. Aguilar, J. M. Ramírez, A. Fernández-Anta, Explainability analysis: An in-depth comparison between Fuzzy Cognitive Maps and LAMDA, Applied Soft Computing 164 (2024) 111940. doi:https://doi.org/10.1016/j.asoc.2024.111940.

[22] Marios Tyrovolas and Nikolaos D. Kallimanis and Chrysostomos Stylios, Advancing Explainable AI with Causal Analysis in Large-Scale Fuzzy Cognitive Maps (2024). arXiv:2405.09190.

[23] P. J. Giabbanelli, G. Nápoles, Fuzzy Cognitive Maps: Best Practices and Modern Methods, Springer Nature Switzerland, 2024. doi:https://doi.org/10.1007/978-3-031-48963-1.

[24] M. Bottero, G. Datola, R. Monaco, Fuzzy Cognitive Maps: a dynamic approach for urban regeneration processes evaluation, Valori e valutazioni 23 (2019) 77–90.

[25] P. P. Groumpos, Using Fuzzy Cognitive Maps in Analyzing and Studying International Economic and Political Stability, IFAC-PapersOnLine 52 (25) (2019) 23–28. doi:https://doi.org/10.1016/j.ifacol.2019.12.440.

[26] I. Á. Harmati, M. F. Hatwágner, L. T. Kóczy, Global stability of fuzzy cognitive maps, Neural Computing and Applications 35 (2023) 7283–7295. doi:https://doi.org/10.1007/s00521-021-06742-9.

[27] I. Á. Harmati, L. T. Kóczy, Stability of Fuzzy Cognitive Maps with Interval Weights, in: Proceedings of the 11th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2019), Atlantis Press, 2019/08, pp. 756–763. doi:https://doi.org/10.2991/eusflat-19.2019.103.

[28] I. Á. Harmati, L. T. Kóczy, On the Convergence of Input-Output Fuzzy Cognitive Maps, in: Proceedings of International Joint Conference on Rough Sets, 2020, p. 449–461. doi:https://doi.org/10.1007/978-3-030-52705-1_33.

[29] I. Á. Harmati, L. T. Kóczy, Some Dynamical Properties of Higher-Order Fuzzy Cognitive Maps, Springer International Publishing, Cham, 2022, pp. 149–156. doi:https://doi.org/10.1007/978-3-030-74970-5_17.

[30] I. Á. Harmati, L. T. Kóczy, On the Convergence of Sigmoidal Fuzzy Grey Cognitive Maps, International Journal of Applied Mathematics and Computer Science 29 (3) (2019) 453–466. doi:https://doi.org/10.2478/amcs-2019-0033.

[31] I. A. Harmati, L. T. Kóczy, Notes on the Dynamics of Hyperbolic Tangent Fuzzy Cognitive Maps, in: 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2019, pp. 1–6. doi:https://doi.org/10.1109/FUZZ-IEEE.2019.8858950.

[32] M. F. Hatwagner, G. Vastag, V. A. Niskanen, L. T. Kóczy, Improved Behavioral Analysis of Fuzzy Cognitive Map Models, in: L. Rutkowski, R. Scherer, M. Korytkowski, W. Pedrycz, R. Tadeusiewicz, J. M. Zurada (Eds.), Artificial Intelligence and Soft Computing, Springer International Publishing, Cham, 2018, pp. 630–641. doi:https://doi.org/10.1007/978-3-319-91262-2_55.

[33] G. Nápoles, I. Grau, L. Concepción, L. Koutsoviti Koumeri, J. P. Papa, Modeling implicit bias with fuzzy cognitive maps, Neurocomputing 481 (2022) 33–45. doi:https://doi.org/10.1016/j.neucom.2022.01.070.

[34] S. Bueno, J. L. Salmeron, Benchmarking main activation functions in fuzzy cognitive maps, Expert Systems with Applications 36 (3) (2009) 5221–5229. doi:https://doi.org/10.1016/j.eswa.2008.06.072.

[35] L. Concepción, G. Nápoles, R. Falcon, K. Vanhoof, R. Bello, Unveiling the Dynamic Behavior of Fuzzy Cognitive Maps, IEEE Transactions on Fuzzy Systems 29 (5) (2021) 1252–1261. doi:https://doi.org/10.1109/TFUZZ.2020.2973853.

[36] K. Hornik, M. Stinchcombe, H. White, Multilayer Feedforward Networks are Universal Approximators, Neural Networks 2 (5) (1989) 359–366. doi:https://doi.org/10.1016/0893-6080(89)90020-8.

[37] S. Mohr, The Use and Interpretation of Fuzzy Cognitive Maps, Master's thesis, Rensselaer Polytechnic Institute (12 1997). doi:https://doi.org/10.13140/RG.2.2.34905.36961.

[38] J. P. Carvalho, On the semantics and the use of fuzzy cognitive maps and dynamic cognitive maps in social sciences, Fuzzy Sets and Systems 214 (2013) 6–19. doi:https://doi.org/10.1016/j.fss.2011.12.009.

[39] J. T. Rickard, J. Aisbett, R. R. Yager, Computing With Words in Fuzzy Cognitive Maps, in: Proc. World Conference on Soft Computing, 2015, pp. 1–6. doi:https://doi.org/10.1109/NAFIPS-WConSC.2015.7284135.

[40] A. K. Tsadiras, Comparing the Inference Capabilities of Binary, Trivalent and Sigmoid Fuzzy Cognitive Maps, Information Sciences 178 (20) (2008) 3880–3894. doi:https://doi.org/10.1016/j.ins.2008.05.015.