On the Behavior of Fuzzy Grey Cognitive Maps

Leonardo Concepción[1,2]([✉]), Gonzalo Nápoles[2,3], Rafael Bello[1],
and Koen Vanhoof[2]

[1] Department of Computer Science, Universidad Central de Las Villas,
Santa Clara, Cuba
`lcperez@uclv.cu, rbellop@uclv.edu.cu`
[2] Faculty of Business Economics, Hasselt Universiteit, Hasselt, Belgium
`{gonzalo.napoles,koen.vanhoof}@uhasselt.be`
[3] Department of Cognitive Science and Artificial Intelligence, Tilburg University,
Tilburg, The Netherlands

Abstract. Fuzzy Cognitive Maps (FCMs) are recurrent neural networks
made up of well-defined neurons and causal relations. Fuzzy Grey Cognitive Maps (FGCMs) are an extension of FCMs, intended to surpass the
intrinsic uncertainties modeling real-world problems by means of Grey
theory. Despite the rising number of studies about FGCM-based models, little has been investigated with regard to the convergence of such
networks. In this paper, we build a mathematical basis to uncover the
behavior FGCM-based models equipped with transfer F-functions. To
do so, we propose sufficient conditions for the existence and unicity of
fixed-point attractors. Also, the results reported in the literature on the
convergence of FGCMs, are compared with ours. Furthermore, we elucidate the reach and depth of our findings, especially and not exclusive to
the prediction of FCMs' behavior.

Keywords: Fuzzy Cognitive Maps · Fuzzy Grey Cognitive Maps ·
Convergence · Grey theory · Shrinking Grey State Vector · Limit grey
state

1 Introduction

Fuzzy Cognitive Maps (FCMs) are knowledge-based recurrent neural networks for
modeling complex systems [5] and an increasing number of FCM scientific articles have been published in the last few years [4,11–14]. Whether using FCMs or
not, the construction of models to face real-world problems always carry in intrinsic uncertainties. Sometimes the vast information contained in a complex system
cannot be represented only by means of crisp values. Grey numbers [20] emerged
as a way to shape these uncertainties and, given that FCMs are not exempt from
these issues, Fuzzy Grey Cognitive Maps (FGCMs) were proposed by [16]. These
networks are convenient for modeling human knowledge in decision-making process. FGCMs are considered a generalization of FCMs, since the latter is an FGCM

with all the causal relations represented by white numbers. Some FGCMs' theoretical breakthroughs emerged in the last decade, as well as applications on time series forecasting [6], surveillance assets coordination [17], reliability engineering [18] and radiotherapy treatment planning [19].

Convergence analysis is the most discussed topic in theoretical studies on the FCM field [4,10–12]. The possible states of the FCMs' inference process are the same as in FGCMs, since both may reach a fixed point or a limit cycle, or exhibit chaotic behavior. Like in FCMs, the convergence in FGCMs is crucial because cycles and chaos make the network responses to be unstable. Unstable cases occur given that activation values always vary through iterations, while remaining stable when a fixed point is reached. While several theoretical studies as well as applications have been conducted using FGCM-based models, convergence issues have been little analyzed [7,8].

In [7], the authors provided some sufficient conditions for the convergence of FGCMs to a unique fixed point, regardless of the initial values of neurons. These conditions are proven for the cases of the log-sigmoid and the hyperbolic tangent threshold functions and also generalized to arbitrary sigmoidal (S-shaped) threshold functions. The convergence conditions are expressed by the elements of the weight matrix and the maximal value of the derivative of the transfer function. Such conditions are applicable to a particular set of the FGCMs' universe, since a small number of neurons and low values for weights are required. Also, the estimation of the maximal value of the derivative could produce loose bounds.

Our previous research on the theoretical analysis of FCM-based models and their dynamics [4] motivated a similar approach to FGCMs. First, we propose novel simplified formulas to calculate the grey raw activation values in FGCMs. Such formulas are the basis to introduce several definitions, lemmas, theorems and corollaries, that allow studying the dynamic behavior of FGCMs equipped with F-functions. The proposed theorems and corollaries give sufficient conditions for the grey state vector to continuously shrink through the inference process and to converge to a so-called limit grey state. After that, we contrast our findings with the achievements depicted in [7] and the employment of FGCMs as predictors for FCMs is proposed.

The rest of this paper is organized as follows. Section 2 goes over the FGCMs' mathematical underpinnings. Section 3 introduces novel formulas for the grey raw activation values, while Sect. 4 enunciates two theorems and their respective corollaries giving conditions on the convergence of FGCM-based models. Section 5 compares our findings with those in the literature. In Sect. 6, we briefly summarize our achievements in this research.

2   Theoretical Background

As mentioned, FGCMs are an extension of traditional FCMs so that both weight and neurons' activation values are described with interval grey numbers.

2.1   Grey Theory

Let $\boldsymbol{\Omega}$ be the universal set, then a grey set $\psi \subset \boldsymbol{\Omega}$ is defined by two membership functions $\mu_\psi^- (\cdot) \in [0,1]$ and $\mu_\psi^+ (\cdot) \in [0,1]$ denoting the lower and upper membership functions, respectively, such that $\mu_\psi^- (\cdot) \leq \mu_\psi^+ (\cdot)$. Interval grey numbers with upper (g^+) and lower (g^-) limits are denoted as $g^\pm \in [g^-, g^+] \mid g^- \leq g^+$ [20]. The crisp value of a grey number is unknown, but the range in which the crisp value is located is known [9].

If $g^\pm$ only has an upper limit, it is denoted by $g^\pm \in (-\infty, g^+]$, but if $g^\pm$ only has a lower limit it is denoted by $g^\pm \in [g^-, +\infty)$. A black number with both unknown limits is denoted as $g^\pm \in (-\infty, +\infty)$ and it becomes a white number when both limits have the same value $g^- = g^+$. Although the length of a grey number with only one limit known ($g^\pm \in [g^-, +\infty)$ or $g^\pm \in (-\infty, g^+]$) is infinite, the grey number is not necessarily a black number because it is possible to know one of these limits.

Equations (1a)–(1d) show the grey arithmetic operations according to the common interval algebra:

$$g_a^\pm + g_b^\pm \in \left[g_a^- + g_b^-, g_a^+ + g_b^+ \right] \tag{1a}$$

$$g_a^\pm - g_b^\pm \in \left[g_a^- - g_b^+, g_a^+ - g_b^- \right] \tag{1b}$$

$$g_a^\pm \cdot g_b^\pm \in \left[\min \left\{ g_a^- \cdot g_b^-, g_a^+ \cdot g_b^+, g_a^- \cdot g_b^+, g_a^+ \cdot g_b^- \right\}, \right.$$
$$\left. \max \left\{ g_a^- \cdot g_b^-, g_a^+ \cdot g_b^+, g_a^- \cdot g_b^+, g_a^+ \cdot g_b^- \right\} \right] \tag{1c}$$

$$\frac{g_a^\pm}{g_b^\pm} \in \left[\min \left\{ \frac{g_a^-}{g_b^-}, \frac{g_a^+}{g_b^+}, \frac{g_a^-}{g_b^+}, \frac{g_a^+}{g_b^-} \right\}, \right.$$
$$\left. \max \left\{ \frac{g_a^-}{g_b^-}, \frac{g_a^+}{g_b^+}, \frac{g_a^-}{g_b^+}, \frac{g_a^+}{g_b^-} \right\} \right] \mid g_b^-, g_b^+ \neq 0. \tag{1d}$$

2.2   Fuzzy Grey Cognitive Maps

The mathematical formalism of grey numbers and the recurrent inference mechanism of traditional FCMs are the foundational underpinnings behind FGCM-based models. Roughly speaking, such models can be defined by means of the following 4-tuple:

$$\Theta = \langle C^\pm, A^\pm, W^\pm, f^\pm(\cdot), \ell(\psi) \rangle. \tag{2}$$

where $C^\pm = \{C_1^\pm, \ldots, C_i^\pm, \ldots, C_M^\pm\}$ is the set of M neurons with grey states $A^\pm = \{A_1^\pm, \ldots, A_i^\pm, \ldots, A_M^\pm\}$ such that $A_i^\pm \in [0,1]$, $f^\pm(\cdot)$ represents the grey transfer function and $\ell(\psi)$ is the range of the activation space. The grey weight connecting neurons $C_i^\pm$ and $C_j^\pm$ is denoted by $w_{ij}^\pm \in [-1,1]$ and gathered into the grey weight matrix $W^\pm$, which is defined as follows:

$$W^{\pm} = \begin{bmatrix} w_{11}^{\pm} & \cdots & w_{1i}^{\pm} & \cdots & w_{1M}^{\pm} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ w_{j1}^{\pm} & \ddots & w_{jj}^{\pm} & \ddots & w_{jM}^{\pm} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ w_{M1}^{\pm} & \cdots & w_{Mi}^{\pm} & \cdots & w_{MM}^{\pm} \end{bmatrix} .$$

The recurrent reasoning process of FGCMs is devoted to updating neurons' grey activation values given an initial stimulus [15]. Thus, in each iteration t the model produces a grey state vector $A^{\pm(t)} = [A_1^{\pm(t)}, \ldots, A_i^{\pm(t)}, \ldots, A_M^{\pm(t)}]$ containing the grey activation values of all neurons in the model. The first state vector $A^{\pm(0)}$ in this sequence is either provided by experts when performing WHAT-IF simulations or derived automatically from data. This iterative procedures is given below:

$$A_i^{\pm(t+1)} = f_i^{\pm} \left(\sum_{j=1}^{M} w_{ji}^{\pm} \cdot A_j^{\pm(t)} \right) \tag{3}$$

such that

$$f_i^{\pm}(\bar{A}_i^{\pm(t+1)}) = [f_i(\bar{A}_i^{-(t+1)}), f_i(\bar{A}_i^{+(t+1)})] \tag{4}$$

where $\bar{A}_i^{-(t+1)}$ and $\bar{A}_i^{+(t+1)}$ represents the neuron's lower and upper grey raw activation values, respectively, while $f_i(.)$ denotes the transfer function. Notice that self-feedback is allowed, given that the expert can avoid this by explicitly setting $w_{jj}^{\pm} = [0, 0]$.

The FGCMs' inference process finishes when the stability appears or a maximal number of iterations is reached. After the inference process, the FGCM either settles down to a fixed pattern of activation values (grey fixed-point attractor), keep cycling between several fixed grey states (limit grey cycle) or behaves chaotically (grey chaotic attractor) [16]. The last state occurs when, instead of stabilizing, the FGCM continues to produce different results for each iteration. Also, this state is only attainable with a continuous activation function.

The most widely used transfer functions are [3] the *sigmoid* function and the *hyperbolic tangent*. The *bivalent*, *trivalent* and *threshold* functions have also been employed. The former have continuous open intervals as their image set, while the latter are bounded into closed intervals instead (they also have discrete image set). Generally speaking, any bounded and monotonically increasing function over the set of real numbers is a candidate transfer function, since the image set of a bounded function belongs to an interval.

Let F be the set of all monotonically increasing functions bounded into non-negative intervals. Let $F^0 \subset F$ and $F' \subset F$ be the subsets bounded into open intervals and closed intervals respectively. Also, let $f_i \in F$ be the transfer function used in the activation process of neuron $C_i^{\pm}$ (i.e., every neuron has its own transfer function). This means that f_i is bounded into a non-negative interval (either open or closed). Observe that the hyperbolic tangent and the trivalent

functions do not belong to F. In this paper, we refer to an F-function as any function belonging to F. It should be highlighted that Eq. (4) holds because f_i is an F-function and thus monotonically increasing. Given that $A_i^\pm \in [0, 1]$, in this paper we assume that the image set of F-functions belongs to $[0, 1]$.

3   Raw Activation Values of Grey Neurons

Equations (1a)–(1d) provide tools for operating with grey values and they involve finding the extreme values of sets. When applied to the inference mechanism of FGCMs, the latter equations can be used in simpler ways given the intrinsic properties of these networks. The grey raw activation value for the neuron $C_i^\pm$ at the $(t + 1)$-th iteration is

$$\bar{A}_i^{\pm(t+1)} = \sum_{j=1}^{M} w_{ji}^\pm \cdot A_j^{\pm(t)}. \tag{5}$$

Equation (5) shows the summation of M terms, where each term is the product of two grey values. Equation (1c) is employed to solve such product, where we would have a set with four elements and we would select the minimum and the maximum among them. The above summation has M terms, so we would repeat the procedure in Eq. (1c) M times and then, we would add up the results using Eq. (1a). For a computer it is not a hard procedure, but using the restrictions of the grey numbers used in Eq. (5), we propose a formula which speeds up the computations while yielding a simplified representation for $\bar{A}_i^{\pm(t+1)}$. Also, this formula avoids explicitly finding minimum and maximum values of sets.

The raw activation value can also be represented by following grey dot product:

$$w_i^\pm . A^{\pm(t)} = \sum_{j=1}^{M} w_{ji}^\pm \cdot A_j^{\pm(t)}. \tag{6}$$

As derived from Eq. (3), the grey dot product between $w_i^\pm$ and $A^{\pm(t)}$ is considered for every neuron $C_i^\pm$ in order to compute its activation value. In this research, we assume that each neuron is influenced by, at least, another neural processing entity. In the case of input neurons, their activation values either remain unchanged or become inactive (depending on the FGCM implementation). Whichever the case, their values are easy to predict. Based on Eq. (1c), but using the fact that $A_i^\pm \subseteq [0, 1]$ and $w_{ji}^\pm \subseteq [-1, 1]$, we propose new formulas to calculate the bounds for the grey dot product between $w_i^\pm$ and $A^{\pm(t)}$. The latter product is the grey raw activation value used to calculate the next activation value for the neuron $C_i^\pm$.

It is known that $A_j^{\pm(t)} = [A_j^{-(t)}, A_j^{+(t)}]$ denotes the grey number associated to the j-th neuron at the t-th iteration. Then, the lower limit of $w_i^\pm . A^{\pm(t)}$, denoted

by $\bar{A}_i^{-(t+1)}$, is:

$$\sum_{j=1}^{M} \frac{w_{ji}^{-}\left(A_j^{+(t)}(1 - sgn(w_{ji}^{-})) + A_j^{-(t)}(1 + sgn(w_{ji}^{-}))\right)}{2} \tag{7}$$

and the upper limit, denoted by $\bar{A}_i^{+(t+1)}$, is

$$\sum_{j=1}^{M} \frac{w_{ji}^{+}\left(A_j^{-(t)}(1 - sgn(w_{ji}^{+})) + A_j^{+(t)}(1 + sgn(w_{ji}^{+}))\right)}{2} \tag{8}$$

where $sgn(.)$ is the sign function

Proof. Previous equations are a compact representation based on the following ideas. Analyzing Eq. (6) we realize that we need to focus on the grey product $w_{ji}^{\pm} \cdot A_j^{\pm(t)} \; \forall j$ to find the lower limit depicted in (7). As we know, the grey limits for $A_j^{\pm(t)}$ are non-negative while the grey limits for $w_{ji}^{\pm}$ could be either positive, negative or zero. According to the sign of $w_{ji}^{\pm}$, two cases emerge in order to find the lower limit of $w_{ji}^{\pm} \cdot A_j^{\pm(t)}$:

– IF $w_{ji}^{-} < 0$ THEN $(w_{ji}^{\pm} \cdot A_j^{\pm(t)})^{-} = w_{ji}^{-}A_j^{+(t)}$
– IF $w_{ji}^{-} \geq 0$ THEN $(w_{ji}^{\pm} \cdot A_j^{\pm(t)})^{-} = w_{ji}^{-}A_j^{-(t)}$

By applying the aforementioned reasoning and using the sign function ($sgn(.)$) we assemble both cases within a single formula without ramifications. This results in the lower limit shown at (7).

Analogously, to derive Eq. (8) we analyze the upper limit of the grey product $w_{ji}^{\pm} \cdot A_j^{\pm(t)} \; \forall j$. Two cases arise:

– IF $w_{ji}^{+} < 0$ THEN $(w_{ji}^{\pm} \cdot A_j^{\pm(t)})^{+} = w_{ji}^{+}A_j^{-(t)}$
– IF $w_{ji}^{+} \geq 0$ THEN $(w_{ji}^{\pm} \cdot A_j^{\pm(t)})^{+} = w_{ji}^{+}A_j^{+(t)}$

From this point, the upper limit shown at (8) is derived in the same way as the lower limit. □

4   Studying Convergence in FGCMs

In this section, we provide sufficient conditions for the existence and uniqueness of grey fixed-point attractors of FGCMs equipped with transfer F-functions. To present the first theorem about the FGCMs' behavior, we need to define when a grey value $a^{\pm}$ contains another $b^{\pm}$.

Definition 4.1. *The grey value $a^{\pm}$ contains the grey value $b^{\pm}$ if $a^{-} \leq b^{-}$ and $a^{+} \geq b^{+}$.*

In other words, the interval where the crisp value of $a^\pm$ lies, contains the interval where the crisp value of $b^\pm$ is located.

Definition 4.2. *The grey vector $A^\pm$ contains the grey vector $B^\pm$ if they have the same length and every grey value from $A^\pm$ contains the corresponding grey value in $B^\pm$.*

Now, we are in conditions to introduce the theorem. This theorem asserts that if some grey state vector contains the next one, then every successive state vector contains the following. Nevertheless, it is possible that $A^{\pm(t)} = A^{\pm(t+1)}$, which implies that $A^{\pm(t)} = A^{\pm(t+k)} \ \forall k \in \mathbb{N}$. So, the grey state vectors may not shrink forever.

Theorem 4.1 (Weak Shrinking Grey State Vector). *In an FGCM Θ, $A^{\pm(t)}$ contains $A^{\pm(t+1)} \ \forall t > t_0 : t_0 \in \mathbb{N}$, if $A^{\pm(t_0)}$ contains $A^{\pm(t_0+1)}$ with $f_i \in F \ \forall i \in \{1, 2, \dots, M\}$.*

Proof. Let $A^{\pm(t_0)} = \{A_1^{\pm(t_0)}, \dots, A_M^{\pm(t_0)}\}$, $A^{\pm(t_0+1)} = \{A_1^{\pm(t_0+1)}, \dots, A_M^{\pm(t_0+1)}\}$ and $A^{\pm(t_0+2)} = \{A_1^{\pm(t_0+2)}, \dots, A_M^{\pm(t_0+2)}\}$.

We must demonstrate that $A_i^{\pm(t_0+1)}$ contains $A_i^{\pm(t_0+2)}$ for every $i = 1, 2, \dots, M$, to prove that $A^{\pm(t_0+1)}$ contains $A^{\pm(t_0+2)}$. The fact that $A^{\pm(t_0)}$ contains $A^{\pm(t_0+1)}$ means that $A_i^{\pm(t_0)}$ contains $A_i^{\pm(t_0+1)}$ for every $i = 1, \dots, M$. Based on this knowledge, we will prove that $A_i^{\pm(t_0+1)}$ contains $A_i^{\pm(t_0+2)}$ for every $i = 1, \dots, M$.

According to Eqs. 3 and 4, we have that $A_i^{\pm(t_0+1)} = f_i^\pm(\bar{A}_i^{\pm(t_0+1)})$ and $A_i^{\pm(t_0+2)} = f_i^\pm(\bar{A}_i^{\pm(t_0+2)}) \ \forall i$. Given that f_i is monotonically increasing, it suffices to prove that $\bar{A}_i^{\pm(t_0+1)}$ contains $\bar{A}_i^{\pm(t_0+2)} \ \forall i$. Based on Definition 4.1 we need to prove two inequalities.

– **Inequality 1:** $\bar{A}_i^{-(t_0+1)} \leq \bar{A}_i^{-(t_0+2)}$ Formula (7) leads us to prove that:

$$\sum_{j=1}^{M} \frac{w_{ji}^- \left(A_j^{+(t_0)}(1 - sgn(w_{ji}^-)) + A_j^{-(t_0)}(1 + sgn(w_{ji}^-)) \right)}{2}$$

$$\leq \sum_{j=1}^{M} \frac{w_{ji}^- \left(A_j^{+(t_0+1)}(1 - sgn(w_{ji}^-)) + A_j^{-(t_0+1)}(1 + sgn(w_{ji}^-)) \right)}{2}$$

It is sufficient to prove that, for every j:

$$w_{ji}^- \left(A_j^{+(t_0)}(1 - sgn(w_{ji}^-)) + A_j^{-(t_0)}(1 + sgn(w_{ji}^-)) \right)$$

$$\leq w_{ji}^- \left(A_j^{+(t_0+1)}(1 - sgn(w_{ji}^-)) + A_j^{-(t_0+1)}(1 + sgn(w_{ji}^-)) \right)$$

There are three possible scenarios depending on the sign of w_{ji}^-:

- **Scenario 1.** If $sgn(w_{ji}^-) = -1$ then

$$A_j^{+(t_0+1)}(1 - sgn(w_{ji}^-)) \leq A_j^{+(t_0)}(1 - sgn(w_{ji}^-))$$

$$A_j^{+(t_0)} \leq A_j^{+(t_0+1)}$$

  which is true because the condition in Theorem 4.1 saying that $A^{\pm(t_0)}$ contains $A^{\pm(t_0+1)}$.

- **Scenario 2.** If $sgn(w_{ji}^-) = 1$ then

$$A_j^{-(t_0)}(1 + sgn(w_{ji}^-)) \leq A_j^{-(t_0+1)}(1 + sgn(w_{ji}^-))$$

$$A_j^{-(t_0)} \leq A_j^{-(t_0+1)}$$

  which is true because the aforementioned condition in Theorem 4.1.

- **Scenario 3.** If $sgn(w_{ji}^-) = 0$ then the inequality holds because both sides of the inequality are zero.

– **Inequality 2:** $\bar{A}_i^{+(t_0+1)} \geq \bar{A}_i^{+(t_0+2)}$
Formula (8) leads us to prove that:

$$\sum_{j=1}^{M} \frac{w_{ji}^+ \left(A_j^{-(t_0)}(1 - sgn(w_{ji}^+)) + A_j^{+(t_0)}(1 + sgn(w_{ji}^+)) \right)}{2}$$

$$\geq \sum_{j=1}^{M} \frac{w_{ji}^+ \left(A_j^{-(t_0+1)}(1 - sgn(w_{ji}^+)) + A_j^{+(t_0+1)}(1 + sgn(w_{ji}^+)) \right)}{2}$$

It is sufficient to prove that, for every j:

$$w_{ji}^+ \left(A_j^{-(t_0)}(1 - sgn(w_{ji}^+)) + A_j^{+(t_0)}(1 + sgn(w_{ji}^+)) \right)$$

$$\geq w_{ji}^+ \left(A_j^{-(t_0+1)}(1 - sgn(w_{ji}^+)) + A_j^{+(t_0+1)}(1 + sgn(w_{ji}^+)) \right)$$

Again, three possible scenarios arise, depending on the sign of w_{ji}^+:

- **Scenario 1.** If $sgn(w_{ji}^+) = -1$ then

$$A_j^{-(t_0+1)}(1 - sgn(w_{ji}^+)) \geq A_j^{-(t_0)}(1 - sgn(w_{ji}^+))$$

$$A_j^{-(t_0+1)} \geq A_j^{-(t_0)}$$

  which is true because the condition in Theorem 4.1 saying that $A^{\pm(t_0)}$ contains $A^{\pm(t_0+1)}$.

- **Scenario 2.** If $sgn(w_{ji}^+) = 1$ then

$$A_j^{+(t_0)}(1 + sgn(w_{ji}^+)) \geq A_j^{+(t_0+1)}(1 + sgn(w_{ji}^+))$$

$$A_j^{+(t_0)} \geq A_j^{+(t_0+1)}$$

  which is true because the aforementioned condition in Theorem 4.1.

- **Scenario 3.** If $sgn(w_{ji}^+) = 0$ then the inequality holds because both sides of the inequality are zero.

At this point, we have demonstrated that $A_i^{\pm(t_0+1)}$ contains $A_i^{\pm(t_0+2)}$, from $A_i^{\pm(t_0)}$ contains $A_i^{\pm(t_0+1)}$. Based on this and proceeding inductively, we can prove that $A_i^{\pm(t_0+2)}$ contains $A_i^{\pm(t_0+3)}$ and so on. Then we confirm that $A^{\pm(t)}$ contains $A^{\pm(t+1)}$, $\forall t > t_0, t_0 \in \mathbb{N}$ if $A^{\pm(t_0)}$ contains $A^{\pm(t_0+1)}$, having $f_i \in F \; \forall i \in \{1, 2, \dots, M\}$. □

As presented in Sect. 2, F-functions are bounded into either open or closed intervals. Therefore, we can associate a grey value with this kind of intervals.

Definition 4.3. *The induced grey value of neuron $C_i^{\pm}$ is such that its lower and upper limits are the same as the lower and upper limits of the interval that the f_i transfer function is bounded to.*

Note: It should be highlighted that function f_i is associated to neuron $C_i^{\pm}$.
A consequence of this definition is the following lemma:

Lemma 4.1. *The induced grey value of neuron $C_i^{\pm}$ contains $A_i^{\pm(t)}$ $\forall t \in \mathbb{N} : t > 0$, which is the grey activation value of neuron $C_i^{\pm}$.*

Proof. The f_i transfer function is bounded to an interval containing $A_i^{\pm t}$ $\forall t \in \mathbb{N} : t > 0$. This happens because the activation values are generated by f_i. When $t = 0$, the experts might decide to assign values out of this interval, but when the inference is triggered, the following activation values will meet the interval restriction. The lemma holds based on Definition 4.3, since the induced grey value of neuron $C_i^{\pm}$ has the same limits as the interval mentioned in such definition. □

Definition 4.4. *The induced grey vector for the FGCM Θ is such that the i-th component is the induced grey value for the neuron $C_i^{\pm}$.*

The following lemma extends Lemma 4.1 to the grey vector space. This lemma is the basis of a new corollary that will be defined next.

Lemma 4.2. *The induced grey vector for the FGCM Θ contains $A^{\pm(t)}$ $\forall t \in \mathbb{N} : t > 0$, which is the grey state vector of the FGCM at t-th iteration.*

Proof. The demonstration results by applying Lemma 4.1 for every activation value in $A^{\pm t}$ and also for every $t > 0$. □

The following corollary shows that the first state vector $A^{\pm(0)}$ plays a major role in the FGCM's behavior.

Corollary 4.1.1. *In an FGCM Θ, $A^{\pm(t)}$ contains $A^{\pm(t+1)}$ $\forall t \in \mathbb{N}$, if $A^{\pm(0)}$ is the induced grey vector of the FGCM, with $f_i \in F \; \forall i \in \{1, 2, \dots, M\}$.*

Proof. We only need to prove that $A^{\pm(0)}$ contains $A^{\pm(1)}$. According to Lemma 4.2, the induced grey vector contains $A^{\pm(1)}$. Having that $A^{\pm(0)}$ is the induced grey vector, the demonstration is completed. □

In order to affirm that grey state vectors will shrink forever and $A^{\pm(t)} = A^{\pm(t+k)}$ is not possible, for any $t, k \in \mathbb{N}$, we need to define when a grey value $a^{\pm}$ strictly contains $b^{\pm}$.

Definition 4.5. *The grey value $a^{\pm}$ strictly contains the grey value $b^{\pm}$ if $a^{-} < b^{-}$ and $a^{+} > b^{+}$.*

Definition 4.6. *The grey vector $A^{\pm}$ strictly contains the grey vector $B^{\pm}$ if they have the same length and every grey value from $A^{\pm}$ strictly contains the corresponding grey value in $B^{\pm}$.*

Similarly to Theorem 4.1, the next theorem declares that if some grey state vector strictly contains the next one, then every successive state vector strictly contains the following. Thus, it is not possible that $A^{\pm(t)} = A^{\pm(t+1)}$, which implies that $A^{\pm(t)} \neq A^{\pm(t+k)} \ \forall k \in \mathbb{N}$. Hence, grey state vectors will shrink forever.

Theorem 4.2 (Strong Shrinking Grey State Vector). *In an FGCM Θ, $A^{\pm(t)}$ strictly contains $A^{\pm(t+1)} \ \forall t > t_0 : t_0 \in \mathbb{N}$, if $A^{\pm(t_0)}$ strictly contains $A^{\pm(t_0+1)}$ with $f_i \in F \ \forall i \in \{1, 2, \ldots, M\}$.*

Proof. 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. Still, the scenario where $sgn(w_{ji}^{-}) = 0$ needs special attention. When this happens, both sides of inequality are equal to zero. If it were true for a fixed i and every j, then we would obtain that the grey activation value for neuron $C_i^{\pm}$ remains constant through iterations of the FGCM. This would imply that any grey state vector does not strictly contain the next one. It turns out that this situation is impossible given that, as we explained in Sect. 3, each neuron is influenced by at least another neural processing entity. Thus, for every neuron $C_i^{\pm}$, there must be at least a non-zero incoming connection. Therefore, the strong version of the theorem is true. □

To define a corollary for Theorem 4.2 we must go over F-functions and specially its subset F^0, which is a subset of F where functions are bounded into open intervals. Also, Definitions 4.4, 4.5 and 4.6 will be useful.

Lemma 4.3. *The induced grey value of neuron $C_i^{\pm}$ strictly contains $A_i^{\pm t} \ \forall t \in \mathbb{N} : t > 0$ if $f_i \in F^0 \ \forall i \in \{1, 2, \ldots, M\}$.*

Proof. Transfer functions are bounded into open intervals, which means that grey activation values will always be contained into a grey value with lower and upper limits matching the interval's limits. Such grey value is the induced grey

value and this concludes the proof. Being more explicit, let us take an example whit the sigmoid function $f(x) = 1/(1 + e^{-\lambda(x-h)})$, which belongs to F^0 because its image set is $(0, 1)$. In this case, the grey activation values of every neuron will take values between 0 and 1, but not inclusive. Also, according to Definition 4.3, the lower and upper limits of the induced grey value of any neuron are 0 and 1, respectively. Therefore, the lemma holds.   □

To define a corollary for Theorem 4.2, the following lemma extends Lemma 4.3 to the grey vector space.

Lemma 4.4. *The induced grey vector for the FGCM Θ strictly contains $A^{\pm(t)}$ $\forall t \in \mathbb{N} : t > 0$ if $f_i \in F^0$ $\forall i \in \{1, 2, \ldots, M\}$.*

Proof. The demonstration results by applying Lemma 4.3 for every activation value in $A^{\pm(t)}$ and also for every $t > 0$.   □

Again, the first state vector $A^{\pm(0)}$ plays a major role in the FGCM's behavior, together with the fact that the transfer function belongs to F^0.

Corollary 4.2.1. *In an FGCM Θ, $A^{\pm(t)}$ strictly contains $A^{\pm(t+1)}$ $\forall t \in \mathbb{N}$, if $A^{\pm(0)}$ is the induced grey vector of the FGCM, with $f_i \in F^0$ $\forall i \in \{1, 2, \ldots, M\}$.*

Proof. We only need to prove that $A^{\pm(0)}$ strictly contains $A^{\pm(1)}$. According to Lemma 4.4, the induced grey vector contains $A^{\pm(1)}$. Having that $A^{\pm(0)}$ is the induced grey vector, the demonstration is completed.   □

The above results lead to the question of whether the grey state vectors will shrink until they become a vector of white numbers or not. Becoming a white number would imply that, for every activation value, both grey limits have the same value. This would indicate that every FGCM converges to a white fixed-point attractor. This situation is false, as we know by other studies [4]. Such concerns serve as a motivation to define the limit grey state.

Let $\mathcal{G}$ be the set of all grey values with non-negative limits and let $\mathcal{S}^M$ be the set of all M-ary Cartesian products over the elements in $\mathcal{G}$. Formally, $\mathcal{S}^M = \{\mathcal{I}_1 \times \mathcal{I}_2 \times \ldots \times \mathcal{I}_M : \mathcal{I}_i \in \mathcal{G}, \forall i = 1, 2, \ldots, M\}$. Grey state vectors of an FGCM with M neurons, belong to $\mathcal{S}^M$.

Definition 4.7. *$A^{\pm(\infty)} \in \mathcal{S}^M$ is the **limit grey state** of Θ, such that $A^{\pm(\infty)} = \lim_{t \to \infty} A^{\pm(t)}$.*

Theorem 4.3. *The limit grey state $A^{\pm(\infty)}$ of FGCM Θ always exists if Θ fulfills the premises in Theorems 4.1 or 4.2.*

Proof. The shrinkage of the state vectors implies a contraction of the activation values. Also, from a mathematical point of view, iterative grey state vectors (e.g., $A^{\pm(0)}, A^{\pm(1)}, \ldots, A^{\pm(t)}, \ldots$) are a sequence of elements over $\mathcal{S}^M$ and iterative grey activation values of a neuron are a sequence of elements over $\mathcal{G}$. A sequence over $\mathcal{G}$ can be interpreted as two other sequences: the sequence of lower bounds and the sequence of upper bounds of the iterative grey activation values (both

sequences are defined over the set of real numbers). We say that a sequence over $\mathcal{G}$ is *convergent*, if the sequences of lower and upper bounds are also convergent.

Theorems 4.1 and 4.2 imply that grey activation values associated with neurons become smaller from one iteration to the following, meaning that the lower bound becomes bigger and the upper bound becomes smaller. This suggests that the lower bounds sequence increases and the upper bounds sequence decreases. Besides, both sequences are bounded from each other, so the lower bounds sequence is a lower bound for the higher bound sequence and vice versa. Thus, both sequences are convergent because the *monotone convergence theorem* [1] and have a limit (in the extreme case, the limit is a closed interval with identical lower and upper bounds). Now, we have that the sequence of grey activation values is convergent, which implies the convergence of the sequence of grey state vectors. At this point, there is no doubt about the existence and unicity of a limit for iterative feasible state spaces.   □

As long as the FGCM is equipped with transfer F-functions and any activation vector contains (or strictly contains) the next one, it will converge to a grey fixed-point attractor, the so-called limit grey state. Corollaries 4.1.1 and 4.2.1 are useful when the initial stimulus is unknown or the experts want to analyze the FGCM's behavior under full uncertainty to arrive at conclusions.

5   Comparison with Previous Convergence Results

In this section we discuss some results reported in the literature, specifically in [7]. Then we explain the reach and depth of our theoretical findings while we compare them with the latter results. Finally, we show the broad applicability of our results, particularly predicting the dynamic behavior of FCMs.

5.1   Literature Results

As mentioned earlier, in [7], conditions for the existence and uniqueness of fixed points of FGCMs are presented. The authors expanded and corrected the ideas presented by [2], who firstly addressed the convergence issues of FCMs from a mathematical perspective. Former conditions are expressed by matrix norms and applying the contraction mapping theorem [2] with suitable distance metrics. To do so, authors build on strict analytical bounds for the derivatives of the transfer functions. These functions may be the log-sigmoid threshold function, the hyperbolic tangent and, in general, the sigmoid-like (S-shaped) threshold function. Also, they assume that the human expert or the training process assigns the proper signs to the grey weights, so a weight is either non-positive or non-negative. This means that the type of relationship (direct or inverse) between the neurons is properly described by the FGCMs.

These sufficient conditions are summarized in Theorem 9 in [7]. This theorem is based upon some definitions that we are going to present briefly.

- W^* is the matrix where every element w_{ij}^* is the maximum between the absolute values of the lower and upper limits of the grey weight $w_{ij}^{\pm}$.
- The 1-norm of W^* is given by: $||W^*||_1 = \max_{1 \leq j \leq M} \left(\sum_{i=1}^{M} w_{ij}^* \right)$
- The ∞-norm of W^* is given by: $||W^*||_\infty = \max_{1 \leq i \leq M} \left(\sum_{j=1}^{M} w_{ij}^* \right)$
- The Frobenius norm of W^* is given by: $||W^*||_F = \left(\sum_{i=1}^{M} \sum_{j=1}^{M} (w_{ij}^*)^2 \right)^{\frac{1}{2}}$
- Let K be the maximal value of $f'(x)$, where f is an S-shaped transfer function. This real function is bounded, monotone increasing and continuously differentiable defined for all real values.

Essentially, Theorem 9 states that the FGCM has one and only one grey fixed point, regardless of the initial values of neurons, if at least one of the inequalities $||W^*||_1 < \frac{1}{K}$, $||W^*||_\infty < \frac{1}{K}$ or $||W^*||_F < \frac{1}{K}$ holds.

In the particular case of the log-sigmoid threshold function (the most widely used), taking into account its common set of parameters, the values of K are always higher than 1. So, it would be needed at least one norm (among the three) lower than 1. If we take a closer inspection on the norms' definitions, we notice that fulfilling this condition would need, we need FGCMs with a small number of neurons and with low values for the influences. In this case, if there exist only one grey weight whose upper limit is equal to 1, then no norm fulfills the inequality. In conclusion, the theorem conditions are quite restrictive to a particular set of the FGCMs' universe.

5.2   Our Findings

First, let us compare our definition of FGCM with the one reported in [7]. Two major differences appear:

- They assume that both limits of the grey weight connecting neurons are either non-positive or non-negative.
- Their transfer functions are S-shaped, bounded, monotone increasing and continuously differentiable.
- They use a single transfer function for the whole FGCM.

We do not restrict weights beyond the property $w_{ij}^{\pm} \in [-1, 1]$. A specific weight could normally be $w_{ij}^{\pm} = [-0.5, 0.5]$, which contradicts the assumption found in [7]. On the other hand, we allow transfer functions to be any bounded and monotonically increasing function, as long as its image set lies into $[0, 1]$ (intrinsic restriction in our FGCMs). We do not restrict to have a single transfer function in the model, since each neuron has one of these functions and the parameters may vary among them. Moreover, the estimation of K (the maximal value of $f'(x)$) may result hard working with more complex functions or loose bounds could be obtained.

By means of Theorems 4.1 and 4.2, we ensure the shrinkage of successive grey state vectors, where the only condition that must hold is that some state vector

contains (or strictly contains) the next one. No restrictions are required for the weights' matrix or the number of transfer functions or the functions' parameter set. Beyond the shrinkage, we also prove the convergence to a unique grey fixed point, the so-called grey limit state.

Corollaries 4.1.1 and 4.2.1 lead us to a particular application for the FGCMs: the prediction of the state space of FCMs. When the first grey state vector is the induced grey vector, our results match the findings in [4]. In that paper we showed that, approximating the state space of an FCMs is useful to predict fixed-point attractors and to find hidden behavior even in unstable FCMs, without running the inference mechanism. Hence, FGCMs serve as predictors for FCMs that can notify human experts about the limitations of the FCM-model with no computational burden caused by the inference process.

6   Concluding Remarks

In this paper, motivated by our findings regarding the dynamic behavior of FCMs [4], we have introduced a set of mathematical entities (i.e., definitions, lemmas, theorems, corollaries) to uncover the behavior of FGCMs equipped with transfer F-functions. The research conducted in [7] is quite similar to ours, in the sense that our goals are the existence and uniqueness of grey fixed-point attractors. However, we use a completely different approach producing less restrictive rules for the existence and unicity of FGCMs. Furthermore, our definition of such models is wider when we refer to the influences among neurons, the usage of a transfer function per neuron and the parameters of these functions. Our research presents more flexible models, while brings forward less restrictive sufficient conditions in order to fulfill the convergence goal.

Theorems 4.1 and 4.2 ensure the shrinkage of the grey state vectors through the inference process, for FGCMs equipped with transfer F-functions. In case the conditions are met, the inference process leads to the grey limit state. This knowledge could be used by experts or maybe in the learning procedure, to ensure the convergence for this neural networks. Finally, we prove that FGCMs are predictors for FCMs' state spaces, disclosing hidden behavior with no need to trigger the inference mechanism.

References

1. Binmore, K.G.: Mathematical Analysis: A Straightforward Approach. Cambridge University Press, Cambridge (1977)
2. Boutalis, Y., Kottas, T.L., Christodoulou, M.C.: Adaptive estimation of fuzzy cognitive maps with proven stability and parameter convergence. IEEE Trans. Fuzzy Syst. **17**, 874–889 (2009)
3. Bueno, S., Salmeron, J.L.: Benchmarking main activation functions in fuzzy cognitive maps. Expert Syst. Appl. **36**(3), 5221–5229 (2009)
4. Concepción, L., Nápoles, G., Falcon, R., Bello, R., Vanhoof, K.: Unveiling the dynamic behavior of fuzzy cognitive maps. IEEE Trans. Fuzzy Syst. (2020). https://doi.org/10.1109/TFUZZ.2020.2973853

5. Felix, G., Nápoles, G., Falcon, R., Froelich, W., Vanhoof, K., Bello, R.: A review on methods and software for fuzzy cognitive maps. Artif. Intell. Rev. **52**(3), 1707–1737 (2017)
6. Froelich, W., Salmeron, J.L.: Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series. Int. J. Approx. Reason. **55**(5), 1319–1335 (2014)
7. Harmati, I.Á., Kóczy, L.T.: On the convergence of sigmoidal fuzzy grey cognitive maps. Int. J. Appl. Math. Comput. Sci. **29**, 453–466 (2019)
8. Harmati, I.Á., Kóczy, L.T.: On the convergence of fuzzy grey cognitive maps. In: Kulczycki, P., Kacprzyk, J., Kóczy, L.T., Mesiar, R., Wisniewski, R. (eds.) ITSRCP 2018. AISC, vol. 945, pp. 74–84. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-18058-4_6
9. Nápoles, G., Salmeron, J.L., Vanhoof, K.: Construction and supervised learning of long-term grey cognitive networks. IEEE Trans. Cybern. 1–10 (2019)
10. Nápoles, G., Bello, R., Vanhoof, K.: How to improve the convergence on sigmoid fuzzy cognitive maps? Intell. Data Anal. **18**(6S), S77–S88 (2014)
11. Nápoles, G., Concepción, L., Falcon, R., Bello, R., Vanhoof, K.: On the accuracy-convergence tradeoff in sigmoid fuzzy cognitive maps. IEEE Trans. Fuzzy Syst. **26**(4), 2479–2484 (2018)
12. Nápoles, G., Papageorgiou, E., Bello, R., Vanhoof, K.: On the convergence of sigmoid fuzzy cognitive maps. Inf. Sci. **350**, 154–171 (2016)
13. Nápoles, G., Papageorgiou, E., Bello, R., Vanhoof, K.: Learning and convergence of fuzzy cognitive maps used in pattern recognition. Neural Process. Lett. **45**, 431–444 (2017)
14. Pedrycz, W., Jastrzebska, A., Homenda, W.: Design of fuzzy cognitive maps for modeling time series. IEEE Trans. Fuzzy Syst. **24**, 120–130 (2016)
15. Salmeron, J., Palos-Sanchez, P.: Uncertainty propagation in fuzzy grey cognitive maps with hebbian-like learning algorithms. IEEE Trans. Cybern. **49**, 211–220 (2019)
16. Salmeron, J.L.: Modelling grey uncertainty with fuzzy grey cognitive maps. Expert Syst. Appl. **37**(12), 7581–7588 (2010)
17. Salmeron, J.L.: An autonomous FGCM-based system for surveillance assets coordination. J. Grey Syst. **28**(1), 27–35 (2016)
18. Salmeron, J.L., Gutierrez, E.: Fuzzy grey cognitive maps in reliability engineering. Appl. Soft Comput. J. **12**(12), 3817–3823 (2012)
19. Salmeron, J.L., Papageorgiou, E.I.: A fuzzy grey cognitive maps-based decision support system for radiotherapy treatment planning. Knowl.-Based Syst. **30**(1), 151–160 (2012)
20. Yang, Y., John, R.: Grey sets and greyness. Inf. Sci. **185**(1), 249–264 (2012)