



# Global stability of fuzzy cognitive maps

István Á. Harmati<sup>1</sup> · Miklós F. Hatwágner<sup>2</sup> · László T. Kóczy<sup>2,3</sup>

Received: 2 March 2021 / Accepted: 10 November 2021 / Published online: 13 December 2021  
© The Author(s) 2021

## Abstract

Complex systems can be effectively modelled by fuzzy cognitive maps. Fuzzy cognitive maps (FCMs) are network-based models, where the connections in the network represent causal relations. The conclusion about the system is based on the limit of the iteratively applied updating process. This iteration may or may not reach an equilibrium state (fixed point). Moreover, if the model is globally asymptotically stable, then this fixed point is unique and the iteration converges to this point from every initial state. There are some FCM models, where global stability is the required property, but in many FCM applications, the preferred scenario is not global stability, but multiple fixed points. Global stability bounds are useful in both cases: they may give a hint about which parameter set should be preferred or avoided. In this article, we present novel conditions for the global asymptotical stability of FCMs, i.e. conditions under which the iteration leads to the same point from every initial vector. Furthermore, we show that the results presented here outperform the results known from the current literature.

**Keywords** Fuzzy cognitive map (FCM) · Stability · Fixed point · Convergence of fuzzy cognitive map · Bacterial evolutionary algorithm

## 1 Introduction

In [1], Axelrod used a directed graph to describe the connections between the political elites. This modelling technique was extended by Kosko [23, 24], who introduced Fuzzy Cognitive Maps (FCMs), by representing the strength of the causal connections using values from the  $[-1, 1]$  interval. The nodes of the graph represent the main subsystems, or system variables; while the weighted,

directed edges express the causal knowledge [39]. During the years, this modelling method proved to be very efficient in the representation of complex multicomponent systems, especially when the exact mathematical description was unknown, extremely complicated, thus difficult to deal with, or influenced by uncertain information. Successful applications of FCMs show a very diverse, colourful picture, including, but not limited to social sciences [5], economic problems [9, 26], educational applications [18], various decision-making problems and risk analysis [10, 34, 37], waste management [4, 12], medical problems [33], time series modelling and analysis [16]. The diversity of the fields where FCMs were applied with success, clearly demonstrates the flexibility and performance of this modelling paradigm. In the FCM terminology, the nodes of the weighted, directed graph are usually called ‘concepts’. These represent special characteristics or subsystems of the modelled system. Activation values are numbers from the unit interval (but sometimes the  $[-1, 1]$  interval is applied) assigned to the concepts to describe the state of the concepts. The initial activation vector usually changes rapidly in the simulation. A simulation always ends with one of the three possible outcomes: [41] (i) the value of the activation vector stabilizes, the iteration arrives to a *fixed point* (FP);

---

✉ István Á. Harmati  
harmati@sze.hu  
Miklós F. Hatwágner  
miklos.hatwagner@sze.hu  
László T. Kóczy  
koczy@sze.hu

<sup>1</sup> Department of Mathematics and Computational Sciences, Széchenyi István University, Egyetem tér 1, Győr 9026, Hungary

<sup>2</sup> Department of Information Technology, Széchenyi István University, Egyetem tér 1, Győr 9026, Hungary

<sup>3</sup> Department of Telecommunication and Media Informatics, Budapest University of Technology and Economics, Magyar tudósok körútja 2, Budapest 1117, Hungary

(ii) *limit cycle* means that activation vectors appear repeatedly in a specific order, and (iii) the system may show no stable or regular behaviour, which is usually called *chaotic* in the FCM literature.

In most decision-making applications ‘what-if’ questions are answered with the help of FCMs and simulations [20]. The simulation is started with a specific scenario (expressing the assumed, studied circumstances), and in the best case the simulation leads to a FP. In these cases the effect and usefulness of a decision can be easily analysed. Limit cycles express a continuously changing state of the system, but at least these states are known and can be examined. Chaotic behaviour, however, should be avoided in most application areas.

FCM models may have one single, or even multiple FPs, thus fixed points are not always stable. In some systems, it is important to know all possible stable states of the system, e.g. in the case of a safety-critical system, where a significant amount of investments may be damaged or even, people can be injured. The FPs of a FCM are usually explored empirically by a series of simulations [15], during which a lot of various scenarios are used to perform simulations with. Unfortunately, no one knows exactly the required minimum number of such scenarios. As Dickerson and Kosko pointed out in [8], the state space of an FCM contains attractor regions, and all simulations started from any point of the same region lead to the same FP. Thus, theoretically one scenario per FP region would be enough, but initially, the boundaries of those regions are unknown for the decision makers. In order to explore all FPs with high confidence, a tremendous number of scenarios have to be evaluated. Here, a practical problem arises: although the computational power required to perform a single simulation is not significant (it contains only a matrix multiplication and a relatively simple threshold function evaluation; moreover, FCMs usually converge to FPs quickly); a high number of repeated simulations may need very long execution times. The situation is absolutely not better, if someone wants to find scenarios leading to a specific stable state (e.g. to be aware of which initial system configurations should be avoided). If this goal-oriented decision support problem is solved with a population-based evolutionary algorithm like in [19], it also requires lots of executions of repeated simulation and long running times. Despite the tremendous efforts, the dynamic behaviour of FCMs can never be mapped with 100% reliability by using these empirical approaches, and therefore, there is an obvious need for a faster and more reliable analytical method.

Unfortunately, it is not easy to provide an analytical method to determine the FPs themselves, or at least their number, or to analyse the stability of FCMs in general, even if they are similar to neural networks [6]. Conditions

expressed by the weights of connections between FCM concepts that guarantee the existence and uniqueness of the FPs, were first introduced by Boutalis et al. [2, 3] for a special case when the steepness parameter value of the so-called sigmoid threshold function is  $\lambda = 1$ . In [13], the authors generalized the results of [3] for arbitrary steepness parameters.

Some aspects of the problem of fixed points were also discussed by Knight et al. [21]. Stability issues of FCMs were also investigated by Lee and Kwon [28], where the authors presented an analytical condition for global exponential stability of FCMs based on the Lyapunov method. Moreover, they applied the theoretical results in clinical decision making in [27]. Luo et al. [29] studied the algebraic dynamics of  $k$ -valued fuzzy cognitive maps.

This paper also addresses the same problem, the issue of global stability of fuzzy cognitive maps is discussed. Naturally, the question of the significance and applicability of global asymptotic stability arises. If the FCM is globally asymptotically stable then, any arbitrary initial stimulus leads to the same, unique fixed point. There are some applications, where this property is useful. For example, in Section 4 of [21], the authors investigated a large and diverse industrial area called the Humber region (UK). Sixteen key concepts (bio-based energy production, by-products, competitiveness, etc.) and 27 weighted, directed connections were considered in the model, based on the stakeholders’ opinion. The ranking between the importance of the factors was based on the activation values at the fixed point of the corresponding FCM. For the ranking to be unique, the authors required the FCM has a unique, stable fixed point. To ensure the global stability of the fixed point, they used the mathematical results presented in the same article (in Sect. 5, we compare their mathematical findings and the results presented in this paper).

Nevertheless, in the majority of applications, the unique fixed point is not the desired scenario. Although global stability is not the required, preferred feature of the FCM in these cases, it is important to know what parameter sets lead to this disadvantageous property. It is somewhat similar to diabetes or high blood pressure: we want to avoid them, so we have to know everything (or at least a lot) about their causes. To summarize it, we definitely do not state that global stability is always an advantageous feature. Although it may be rather a curse than a blessing in certain cases, this feature potentially comes with FCM models, so we have to explore its causes, even if to be able to avoid it.

In this paper, we present some novel results on the global stability of FCMs. Moreover, we show that these results are better than the previous stability bounds known from the literature. Additionally, we also show that the weight-independent condition for global stability (reported

earlier in the literature can be improved using our results. This new weight-independent condition is not only better than the previous one, but is also extremely simple. Finally, as a side result, we point out that a recent result regarding the existence and uniqueness of fixed points of FCMs is not valid.

The paper is organized as follows. In Sect. 2, we review the basic notions of fuzzy cognitive maps. Section 3 summarizes the basic mathematical concepts applied for the investigation of the problem of fixed points. Section 4 presents new theoretical achievements for globally asymptotically stable fixed points of FCMs, providing a better upper bound for the parameter of the threshold function. Moreover, conditions related to the structure of the FCM are also presented. In Sect. 5, the comparison of the result of the current paper and other authors’ findings is presented, pointing out that the approach of Sect. 4 gives a better upper bound for the parameter of the threshold function. The results of the paper are summarized in Sect. 6. In Appendix 1, we present the proof of Theorem 6, in Appendix 2 we discuss the validity of a recent result.

## 2 Basic notions of fuzzy cognitive maps

From the mathematical point of view, an FCM contains the following components: a weighted, directed graph expressing the causal relations between the concepts; and the updating rule, including the transformation function, which squashes the weighted sum of activation values into the allowed range (it is usually  $[0, 1]$ , but sometimes  $[-1, 1]$ ) [41]. In graph theory, the adjacency matrix contains the whole information about the connections in the graph. If we deal with a weighted, directed graph, then matrix  $W$  containing the weights ( $w_{ij} \in W$ ) of the connections (and zeros, if there are no causal connections) stores the causalities of the model. The nonnegative number  $|w_{ij}|$  describes the strength of influence of concept  $C_j$  on concept  $C_i$ ; moreover, if  $w_{ij} > 0$ , then a positive change in the activation value of  $C_j$  causes a positive change in the activation value of  $C_i$ ; if  $w_{ij} < 0$ , then positive change causes negative change. The weighted sum of the incoming activation values is transformed into the required range. The transformation is computed by a threshold function. Well-known discrete threshold functions are the bivalent and trivalent functions, while in continuous case we find various sigmoid-like functions. In most of the cases, FCM users choose the sigmoid function, see Eq. (1).

$$f(x) = \frac{1}{1 + e^{-\lambda x}} \tag{1}$$

The steepness parameter  $\lambda > 0$  controls the speed of transition from low values (close to zero) to high values (close

to one). If  $\lambda$  is small, then the function is close to a linear function; if  $\lambda$  is large, then the function is similar to the Heaviside function.

FCM simulation starts with a vector of initial activation values  $A(0) = [A_1(0), \dots, A_n(0)]^T$ . In each simulation step, the activation vector is re-calculated according to the updating rule. The simulation ends when (i) the activation vector is stabilized; (ii) the number of iteration steps reaches the prescribed maximum. In some applications, the updating rule contains self-feedback, but in some other cases self-feedback is not preferred. The general form of the updating rule is

$$A_i(k) = f_i \left( \sum_{j=1, j \neq i}^n w_{ij} A_j(k-1) + d_i A_i(k-1) \right). \tag{2}$$

Here,  $A_i(k)$  is the activation value of concept  $C_i$  at simulation step  $k$ ,  $f_i$  is the threshold function applied at concept  $C_i$ ,  $w_{ij}$  is the weight of causal edge from  $C_j$  to  $C_i$ ,  $d_i$  is the strength of the self-feedback. If  $d_i = 0$ , then there is no self-feedback, as it was used in the first FCM models. Although self-feedbacks were not allowed in Kosko’s original FCMs and are also avoided in some applications, they may be useful in specific cases. Without self-feedbacks, the activation value of a concept is defined by other concepts only. It is not realistic in some cases, however. It is easy to imagine a car, where the speed of the car not only depends on, e.g. the current position of the gas pedal, but on the speed of the car at the previous moment as well. In this example, the current velocity is not independent of the speed measured at the previous time step and the driver can also influence the speed by pushing the gas pedal. Many other, similar examples can be given, where a concept has some kind of ‘memory’. The intensity of that memory can be expressed by the weight of the self-feedback ( $d_i$ ). The theoretical background of self-feedbacks is already laid [11, 40], and several real-life examples can be found for their application [7, 22, 38] as well.

Self-feedback can be built into the weight matrix ( $d_i$ s into the diagonal, i.e.  $w_{ii} = d_i$ ), then the updating rule turns into the following:

$$A_i(k) = f_i \left( \sum_{j=1}^n w_{ij} A_j(k-1) \right). \tag{3}$$

In the present paper, we use this type of  $W$ , so in our terminology the weight matrix already contains the possible self-feedback, i.e. if self-feedback is applied then the diagonal of  $W$  contains the weights of the feedback ( $d_i$ s), if not, then the diagonal of  $W$  contains zeros.

### 3 Mathematical tools

In this section, we shortly summarize the mathematical notions and tools applied in Sects. 4 and 5. For more detailed and precise information about fixed points and fixed point theorems we refer to [35], while for linear algebra and matrix analysis, see [17]. A fixed point of a function  $G$  is a point of the state space such that  $G$  maps this point to itself:  $G(x^*) = x^*$ . Fixed point  $x^*$  is *locally asymptotically stable* if starting the iteration at an arbitrary point close enough to  $x^*$ , the iteration converges to  $x^*$ . If the iteration converges to  $x^*$  for every initial value, then  $x^*$  is a *globally asymptotically stable* fixed point.

According to Brouwer's theorem (see [35], pp. 296–299), every continuous function, which maps a convex, bounded and closed set  $K \subset \mathbb{R}^n$  to itself has (at least one) fixed point. Consequently, this theorem ensures the *existence* of at least one fixed point for any fuzzy cognitive map with continuous threshold function. Since the FCM reasoning is based on the limit of the iteration, we may wonder under what conditions this limit does exist. If by the application of the iteration rule the activation vectors get closer and closer to each other and their difference goes to the zero vector, then the iteration will converge to a certain point. In this case, points of the state space get closer to each other by applying a certain function. This property can be formalized as follows (see [36], page 220):

'Let  $(X, d)$  be a metric space, with metric  $d$ . If  $\varphi$  maps  $X$  into  $X$  and if there is a number  $c < 1$  such that

$$d(\varphi(x), \varphi(y)) \leq cd(x, y) \quad (4)$$

for all  $x, y \in X$ , then  $\varphi$  is said to be a *contraction* of  $X$  into  $X$ .'

The famous contraction mapping theorem (a.k.a. Banach's fixed point theorem, see [36], pp. 220–221 or [35], pp. 236–237) states that if a mapping is a contraction over a nonempty complete metric space, then it has exactly one fixed point. The proof of this statement tells more than the theorem: this fixed point can be found as a limit of the iteration  $x_{n+1} = G(x_n)$ , starting from an arbitrary point in the state space. Since the iteration converges to this unique equilibrium point from any initial values, this fixed point is asymptotically stable in the global sense.

In the results presented in Sect. 4, we prove the globally asymptotic stability of the unique fixed point using the contraction mapping theorem. The theorems and proofs require the basic knowledge of some notions from linear algebra, such as matrix norms, spectral radius and relations and inequalities between them. For these facts, we refer to [17]. However, there are two theorems which will be applied in Sect. 4, and which should be mentioned here. (Below,  $\rho(M)$  denotes the spectral radius of matrix  $M$ ):

**Theorem 1** (see [17], page 349) Let  $M \in \mathbb{R}^{n \times n}$  and  $\varepsilon > 0$  be given. There is matrix norm  $\|\cdot\|$ , such that  $\rho(M) \leq \|M\| \leq \rho(M) + \varepsilon$ .

**Theorem 2** (see [17], page 373) If  $\|\cdot\|_m$  is a matrix norm, then there is a vector norm  $\|\cdot\|_v$  that is compatible with it (i.e.  $\|Mx\|_v \leq \|M\|_m \cdot \|x\|_v$ ).

### 4 Conditions for the global stability of fuzzy cognitive maps

In this section, we prove two theorems for the global asymptotical stability of fixed points of FCMs. Consider again the updating rule of an FCM:

$$A_i(k) = f_i \left( \sum_{j=1, j \neq i}^n w_{ij} A_j(k-1) + d_i A_i(k-1) \right) \quad (5)$$

Let us introduce mapping  $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$  generating the next concept vector from the preceding one. Then the mapping with coordinates:

$$A(k+1) = \begin{bmatrix} A_1(k+1) \\ \vdots \\ A_n(k+1) \end{bmatrix} = \begin{bmatrix} f_1(w_1 A(k)) \\ \vdots \\ f_n(w_n A(k)) \end{bmatrix} = G(A(k)), \quad (6)$$

where  $f_i$  is the transformation function assigned to the  $i$ th concept and  $w_i = (w_{i1}, \dots, w_{in})$ ,  $w_{ij} \in W$ . We know from Banach's theorem that a contraction has a unique fixed point and it can be determined by an iteration starting from any point of the space. Since the FCM reasoning is based on an iteration, the application of this theorem is straightforward. Although we cannot compute the unique fixed point analytically from the given parameters ( $W$  and  $f_i$ s), we are able to state some conditions, for which the examined mapping  $G$  is a contraction. As we have seen from its definition, the notion of contraction requires a distance metric. This distance metric can be generated by a vector norm  $\|\cdot\|_v$  (we do not specify this norm, it can be an arbitrary vector norm), the distance of two concept vectors is defined as the norm of their difference. Using this vector norm, we can define a matrix norm (a.k.a. induced matrix norm or natural matrix norm) as  $\|M\|_* = \sup \left\{ \frac{\|Mx\|_v}{\|x\|_v} : x \in \mathbb{R}^n, x \neq \mathbf{0} \right\}$ . Besides matrix and vector norms, we are going to use the maximum value of the derivative of the threshold function. In general, a sigmoid-like threshold function is a monotone increasing, differentiable function, with finite limits at negative and positive infinity. Let  $f(x)$  be a threshold function of this type, then the maximal value of its derivative is finite, let

us denote it by  $K$ . Then this  $K$  can be considered as a Lipschitz constant:  $|f(x) - f(y)| \leq K \cdot |x - y|$  (Fig. 1).

**Theorem 3** Consider a fuzzy cognitive map (FCM) with weight matrix  $W$ . Moreover, let  $K_i$  be the maximum of the derivative of the threshold function  $f_i$  applied at the  $i$ th concept. If the inequality

$$\|\text{diag}(K_i) \cdot W\|_* < 1 \tag{7}$$

holds with an induced matrix norm  $\|\cdot\|_*$ , then the fuzzy cognitive map has the same fixed point for every initial activation vector.

**Proof** Let

$$G(A) = [f_1(w_1A), f_2(w_2A), \dots, f_n(w_nA)]^T \tag{8}$$

where  $w_i = (w_{i1}, \dots, w_{in})$ . Moreover, let  $\|\cdot\|_v$  be the vector norm generating matrix norm  $\|\cdot\|_*$ . In the following, we give an upper estimation of the value of  $\|G(A) - G(A')\|_v$ :

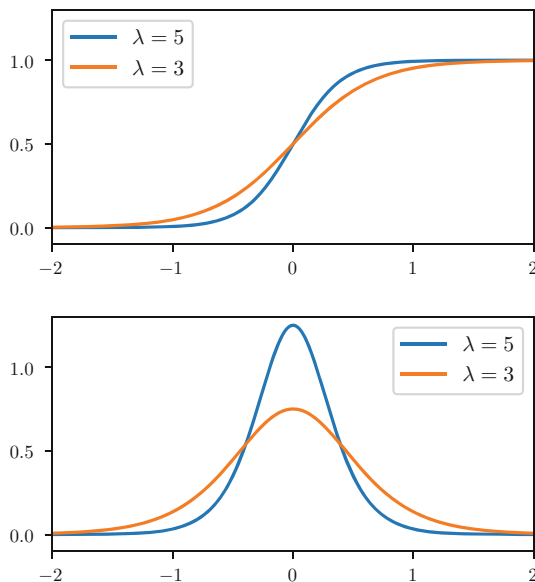
$$\begin{aligned} \|G(A) - G(A')\|_v &= \|f_1(w_1A) - f_1(w_1A'), \dots, f_n(w_nA) - f_n(w_nA')\|_v \end{aligned} \tag{9}$$

From the Mean Value Theorem, we have

$$|f_i(w_iA) - f_i(w_iA')| \leq K_i |w_iA - w_iA'|. \tag{10}$$

So we have

$$\|G(A) - G(A')\|_v \tag{11}$$



**Fig. 1** The sigmoid threshold function with parameter  $\lambda = 5$  and  $\lambda = 3$  (top) and their bell-shaped derivatives (bottom). The maximum value of the derivative is  $\lambda/4$ , i.e. 1.25 and 0.75, respectively

$$\leq \left\| [K_1(w_1A - w_1A'), \dots, K_n(w_nA - w_nA')]^T \right\|_v \tag{12}$$

$$= \|\text{diag}(K_i)W(A - A')\|_v \tag{13}$$

$$= \frac{\|\text{diag}(K_i)W(A - A')\|_v}{\|A - A'\|_v} \cdot \|A - A'\|_v \tag{14}$$

$$\leq \|\text{diag}(K_i)W\|_* \cdot \|A - A'\|_v, \tag{15}$$

where the last row is the consequence of the definition of matrix norm  $\|\cdot\|_*$ . If  $\|\text{diag}(K_i)W\|_* < 1$ , then  $G$  is a contraction mapping. It means that starting from any arbitrary initial values, the repetitive application of the FCM updating rule leads to the same equilibrium point.  $\square$

In Theorem 3, we did not specify the matrix norm (it can be, for example  $\|\cdot\|_2$ ), neither the threshold function  $f_i$  or the maximum value  $K_i$  of its derivative. If we choose them appropriately, we get some additional interesting results:

- If the threshold function is the most widely used one (i.e.  $f_i(x) = \frac{1}{1+e^{-\lambda x}}$ ), then  $K_i = \frac{\lambda}{4}$  and the condition turns to  $\|\text{diag}(\lambda_i)W\|_* < 4$ .
- If the threshold function is the same sigmoid function for all the concepts, i.e.  $f_i(x) = f(x) = \frac{1}{1+e^{-\lambda x}}$ , then  $K_i = \frac{\lambda}{4}$ , so the condition reduces to  $\|W\|_* < \frac{4}{\lambda}$ . In other words, if parameter  $\lambda < 4/\|W\|_*$ , then every initial stimulus leads to the unique fixed point. It is clear that if  $\|W\|_*$  is smaller, then global stability is ensured for a larger set of possible values of  $\lambda$ .
- Weighted in-degree and weighted out-degree are widely used descriptive measures of networks. If the matrix norm is the 1-norm or the  $\infty$ -norm, then condition for global stability can be expressed by the weighted in-degree and weighted out-degree, similarly as it was done for input-output FCMs in [14]. Nevertheless, although weighted in-degree and out-degree are very useful for descriptive analysis of networks, the convergence conditions expressed by them can be easily overperformed by other matrix norms.

Based on the relation between spectral radius and induced matrix norms, we show a better condition for global asymptotical stability of FCMs ( $\rho(\cdot)$  denotes the spectral radius of the matrix).

**Theorem 4** Consider a fuzzy cognitive map (FCM) with weight matrix  $W$ . Moreover, let  $K_i$  be the maximum of the derivative of the threshold function  $f_i$  applied at the  $i$ th concept. If

$$\rho(\text{diag}(K_i) \cdot W) < 1, \tag{16}$$

then the fuzzy cognitive map has the same fixed point for every initial activation vector.

**Proof** We have seen already that if  $\|\text{diag}(K_i) \cdot W\|_* < 1$  with a matrix norm, then the mapping generating the iteration is a contraction. Moreover, if

$$\rho(\text{diag}(K_i) \cdot W) < 1 \quad (17)$$

then Theorem 1 ensures the existence of a matrix norm (let us denote it by  $\|\cdot\|_M$ ) such that

$$\|\text{diag}(K_i)W\|_M < 1 \quad (18)$$

Given this matrix norm, Theorem 2 ensures the existence of a compatible vector norm to this matrix norm (let us denote it by  $\|\cdot\|_v$ ). If we measure the distance of the concept vectors with this norm, then we have

$$\|G(A) - G(A')\|_v \leq \|\text{diag}(K_i)W(A - A')\|_v \quad (19)$$

$$\leq \|\text{diag}(K_i)W\|_M \cdot \|A - A'\|_v \quad (20)$$

According to Eq. (18), the coefficient of  $\|A - A'\|_v$  is less than one. It means that using distance metric  $d(x, y) = \|x - y\|_v$ , mapping  $G$  is a contraction. Similar to the previous theorem, it means that starting the iteration from anywhere in the state space, the iterative FCM updating leads to the same equilibrium point.  $\square$

Since the spectral radius is the infimum of the induced matrix norms, the condition using  $\rho(W)$  is better than the conditions provided by matrix norms.

In a special case, when the threshold function is the sigmoid function and the slope parameter is the same for every concept (i.e.  $\lambda_1 = \lambda_2 = \dots = \lambda_n = \lambda$ ), the condition simplifies to

$$\rho(W) < \frac{4}{\lambda} \quad (21)$$

**Remark 1** In Theorems 3 and 4, we assumed the differentiability of the continuous threshold function. Differentiability is an advantageous property in learning the weights of the FCM; nevertheless, it is not a necessity. Theoretically, one may choose other continuous threshold function, for example a continuous piecewise linear function. On the other hand, one may recognize that in the proof of Theorem 3, we used the maximal value of the derivative as a Lipschitz constant. Consequently, Theorems 3 and 4 are valid for every Lipschitz continuous threshold function, with the modification that  $K_i$  is the Lipschitz constant belonging to threshold function  $f_i$ .

## 5 Comparison of the results with previous results

In this section, we shortly summarize previous theoretical research carried out on the problem of fixed points of FCMs, and compare these results to the results presented in previous section.

### 5.1 Comparison with the results of Boutalis et al.

According to our best knowledge, the first theoretical study discussing the existence and uniqueness of fixed points of FCMs was given by to Boutalis et al. [3] and [2]. They investigated the case, when the transformation function is  $f(x) = 1/(1 + e^{-x})$ , i.e. parameter  $\lambda$  equals one. They arrived to the conclusion that if the following inequality

$$\left( \sum_{i=1}^n \|w_i\|^2 \right)^{1/2} < 4 \quad (22)$$

holds, then the FCM has a unique fixed point and the iteration starting from an arbitrary initial activation vector eventually converges to this point (on the left,  $\|w_i\| = \sqrt{w_{i1}^2 + w_{i2}^2 + \dots + w_{in}^2}$ ). Note that the expression in Eq. (22) is just the Frobenius norm of weight matrix  $W$ . Their findings were generalized for sigmoid FCMs equipped with arbitrary positive parameter  $\lambda$  in [13]. Namely, it was proved that if  $\|W\|_F < 4/\lambda$ , then the iteration process of the FCM leads to the unique equilibrium point, independently from the initial activation values. The well-known inequalities between different matrix norms (and spectral radius) provide an easy way for comparison of results above and the results presented in Sect. 4. We can find a matrix norm  $\|\cdot\|_*$ , such that  $\|W\|_* \leq \|W\|_F$ , thus the global convergence to a unique fixed point is proved for a larger set of possible values of parameter  $\lambda$ . For example, we may choose the operator norm ( $\|\cdot\|_2$ ), or in some cases the infinity norm ( $\|W\|_\infty$ ) or the taxicab norm ( $\|W\|_1$ ). Furthermore, we may take the spectral radius, since  $\rho(W) \leq \|W\|_F$ , thus  $4/\rho(W) \geq 4/\|W\|_F$ . In Kottas et al. [25], the authors attempted to extend the results of [3]. We discuss this issue in Appendix 2.

### 5.2 Comparison with the findings of Knight et al.

In [21] Knight, Lloyd and Penn stated two theorems (Theorem 3.1 and 3.2) regarding the possible number of fixed points of fuzzy cognitive maps. Theorem 3.1 of [21] states that if parameter  $\lambda \geq 0$  of the sigmoid function is *small enough* then there is a unique fixed point, that is linearly stable. Conversely, the theorem also states that if  $\lambda \geq 0$  is large enough there can be multiple fixed points. In

Theorem 3.2 of [21], they clarify the notion of *small enough*. We cite this theorem literally:

Theorem 3.2 of [21]: ‘For  $W \in \mathbb{R}^{n \times n}$  given, the sigmoid FCM has a unique fixed point for all  $\lambda$  such that  $0 \leq \lambda \leq \bar{\lambda}(n)$ , this fix point is stable.  $\bar{\lambda}(n)$  satisfies

$$\left(1 - \frac{\bar{\lambda}(n)}{4}\right)^n - \sum_{i=1}^n b_i C_i^n \left(\frac{\bar{\lambda}(n)}{4}\right)^i = 0 \tag{23}$$

where  $C_i^n$  are the binomial coefficients, and  $b_i$  is given by the recursion relation  $b_i = ib_{i-1} + (-1)^i$ ,  $b_0 = 1$ .

Both of the theorems of [21] deal with FCMs equipped with the same parameter of  $\lambda$  for all of the concepts. Moreover, the weight matrix  $W$  was not taken into consideration in the theorems (it is not a fault, but a possible loss of information). From our results it follows that Theorem 3.2 of [21] can be improved. First we prove that for  $\lambda < 4/n$  ( $n$  is the number of concepts of the FCM), the FCM has exactly one fixed point. Then we show that this extremely simple bound ( $4/n$ ) is better than the bound provided by Eq. (23).

**Theorem 5** Consider a sigmoid FCM with weight matrix  $W \in \mathbb{R}^{n \times n}$  and sigmoid parameter  $\lambda$ . If  $0 \leq \lambda < \frac{4}{n}$  then the FCM has a unique, globally asymptotically stable fixed point.

**Proof** The statement below is an immediate consequence of Theorem 3. If

$$\|\text{diag}(\lambda_i) \cdot W\|_2 < 4 \tag{24}$$

then there is exactly one fixed point. Since in the current case  $\lambda_1 = \lambda_2 = \dots = \lambda_n = \lambda$ , this becomes

$$\lambda \cdot \|W\|_2 < 4 \tag{25}$$

which implies

$$\lambda < \frac{4}{\|W\|_2} \tag{26}$$

Moreover, we do know that  $\|\cdot\|_2 \leq \|\cdot\|_F$ . Thus, if  $\lambda < \frac{4}{\|W\|_F}$  holds, then  $\lambda < \frac{4}{\|W\|_2}$  holds, too. In weight matrix  $W$ , all of the entries are between  $-1$  and  $1$ . Furthermore, by its definition, the Frobenius norm is  $\|W\|_F = \sqrt{\sum_{i,j} w_{ij}^2}$ , so an upper estimation on the Frobenius norm of the weight matrix  $W$  is:

$$\|W\|_F \leq n \tag{27}$$

Since  $\frac{4}{n} \leq \frac{4}{\|W\|_F}$ , this completes the proof.  $\square$

If self-feedbacks are not allowed, then we have  $\|W\|_F \leq \sqrt{n(n-1)}$ , resulting in the bound  $\lambda < \frac{4}{\sqrt{n(n-1)}}$ ,

which is slightly higher than  $\frac{4}{n}$ .

One may think that better upper estimation can be given if we use other norm or the spectral radius of  $W$ , but this is not the case. Since the entries of  $W$  are between  $-1$  and  $1$  (and in this case we have no further information about the weights), the inequality  $\rho(W) \leq n$  holds (equality holds in the extreme case, when  $w_{ij} = 1$  for every  $i, j$ ), so the spectral radius leads to the same upper bound, consequently, we cannot get better bound with any matrix norm.

If we have more information about the weights, then this upper bound  $\frac{4}{n}$  can be further increased. For example, if  $W \in \mathbb{R}^{n \times n}$  has exactly  $k$  nonzero elements (i.e. the FCM has exactly  $k$  connections with nonzero weights), then  $\|W\|_F \leq \sqrt{k} \leq n$ , and  $\frac{4}{\sqrt{k}} \geq \frac{4}{n}$ .

Theorem 5 improves Theorem 3.2 of [21], since it ensures the uniqueness of fixed points for a larger set of values of  $\lambda$ . We can observe it in Fig. 2 and the next theorem shows that the inequality  $\bar{\lambda}(n) \leq \frac{4}{n}$  holds for every FCM (equality holds for  $n = 1, 2$ ).

**Theorem 6** Let  $\bar{\lambda}(n)$  be defined as in Eq. 23. Then the inequality

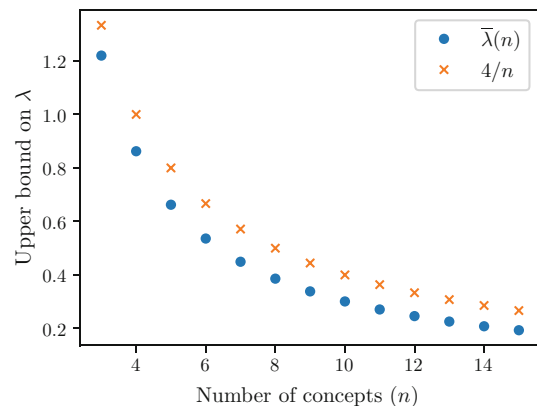
$$\bar{\lambda}(n) \leq \frac{4}{n} \tag{28}$$

holds for every  $n \geq 1$ .

**Proof** See Appendix 1.  $\square$

### 5.3 Comparison with the result of Lee and Kwon

Lee and Kwon examined the problem of equilibrium points via Lyapunov stability analysis [28]. In their approach, the inference rule is described by the following equation (we changed the notations for convenience):



**Fig. 2** Proven upper bounds on parameter  $\lambda$  for global stability,  $\bar{\lambda}(n)$  (blue bullet) and  $4/n$  (orange cross) versus number of concepts ( $n$ ). The values of  $4/n$  are slightly higher than the values of  $\bar{\lambda}(n)$  (color figure online)

$$A_i(k) = f\left(r_1 A_i(k-1) + r_2 \sum_{j=1, j \neq i}^n w_{ij} A_j(k-1)\right) \quad (29)$$

Here,  $f$  is the sigmoid function and its parameter  $\lambda$  is the same for every concept. They proved a criteria for global exponential stability (see [28]):

‘If the inequality

$$0 < \lambda < \frac{4}{r_1 + r_2 \|W^*\|_2} \quad (30)$$

holds, then the equilibrium point of the corresponding FCM is globally exponentially stable’.

We note here that in their approach the weight matrix does not contain the self-feedback, but the self-feedback is expressed by the term  $r_1$ . Because of this reason this weight matrix is denoted here by  $W^*$ .

To compare their result with the findings of Sect. 4, we start with a lower bound of the denominator of Eq. (30):

$$r_1 + r_2 \|W^*\|_2 = \|\text{diag}(r_1)\|_2 + r_2 \|W^*\|_2 \quad (31)$$

$$= r_2 \|\text{diag}(r_1/r_2)\|_2 + r_2 \|W^*\|_2 \quad (32)$$

$$\geq r_2 \|\text{diag}(r_1/r_2) + W^*\|_2 \quad (33)$$

In our approach,  $r_2 = 1$  and  $r_1 = d_i$ , so we get

$$\begin{aligned} r_2 \|W^* + \text{diag}(r_1/r_2)\|_2 &= \|W^* + \text{diag}(d_i)\|_2 = \|W\|_2 \\ &\leq \|W^*\|_2 + \|\text{diag}(d_i)\|_2 \end{aligned} \quad (34)$$

where  $W$  stands for the weight matrix including self-feedback ( $d_i$ s in the diagonal). Based on the inequality above, we get that

$$\frac{4}{r_1 + r_2 \|W^*\|_2} \leq \frac{4}{\|W\|_2} \quad (35)$$

The last inequality ensures that the bound for parameter  $\lambda$  provided in Sect. 4 is better than the bound given in [28].

#### 5.4 Example and ordering of the bounds

As we have seen, various approaches provide different bounds with the property that if the steepness parameter ( $\lambda$ ) of the sigmoid function is less than a number computed from some parameters of the model, then the FCM iteration rule produces the same equilibrium point for every initial activation vector. Based on the properties of matrix norms and the spectral radius, a simple ordering of the proven bounds can be given ( $\bar{\lambda}$  refers to the bound by Knight et al.):

$$\bar{\lambda} \leq \frac{4}{n} \leq \frac{4}{\|W\|_F} \leq \frac{4}{\|W\|_2} \leq \frac{4}{\rho(W)} \quad (36)$$

The following toy example illustrates the different

performance of these bounds. Consider a fuzzy cognitive map with weight matrix  $W$ :

$$W = \begin{pmatrix} 0 & -1 & 0.5 & 0 & 0 \\ 0 & 0 & -0.5 & 0.5 & -0.5 \\ -1 & 1 & 0 & -0.5 & 0 \\ 0 & 1 & -1 & 0 & -0.5 \\ -1 & 0 & 1 & -0.5 & 0 \end{pmatrix} \quad (37)$$

The threshold function is the same for all the concepts,  $f(x) = \frac{1}{1+e^{-\lambda x}}$ . Table 1 shows the different upper bounds on parameter  $\lambda$ , with proven global stability, provided by different methods (i.e. if the value of  $\lambda$  is less than the given number, then the FCM is globally asymptotically stable). We can observe that the method applying the spectral radius gives the best result. Numerical experiments show that the unique fixed point loses its global stability at about  $\lambda \approx 3.6708$ .

## 6 Summary

Fuzzy cognitive map-based reasoning relies on the behaviour of repeated application of the updating rule, i.e. it depends on the behaviour of an iteration. This iteration may or may not converge to an equilibrium point (fixed point). If the iteration converges to a fixed point, then this fixed point (and the whole FCM model) may or may not be globally asymptotically stable. Moreover, if the model is globally asymptotically stable, then this fixed point is unique and the iteration arrives at this point from every initial state. In other words, if the model is globally asymptotically stable, then the system reaches the same equilibrium point, regardless of the initial stimulus.

It has been previously known from the literature that, in case of sigmoidal threshold functions, this property is somewhat related to the value of the steepness parameter  $\lambda$ . Namely, if the value of  $\lambda$  is small (i.e. the transition from close to zero to close to one is not so drastic), then the FCM has the global stability property. Moreover, it has been also

**Table 1** Comparison of different bounds on  $\lambda$ , using weight matrix  $W$  (Eq. 37)

Method	Bound on $\lambda$
$\bar{\lambda}(n)$	0.6624
$4/n$	0.8
$4/\ W\ _\infty$	1.3333
$4/\ W\ _1$	1.6
$4/\ W\ _F$	1.3522
$4/\ W\ _2$	1.8920
$4/\rho(W)$	3.3237



clear, that this property is influenced by the structure of the weighted connections of the network (i.e. weight matrix  $W$ ).

In this paper, several novel analytical conditions have been presented for the global stability of fuzzy cognitive maps. These conditions involve the usual parameters of the model, namely the weight matrix and the parameter of the threshold function. Comparing to the existing results in the literature, these conditions are simpler and give a more efficient upper bound on the parameter of the threshold function. Moreover, the corresponding matrix norms and spectral radius can be easily determined by free mathematical software.

The results presented in this paper can be used in at least two different ways: in some applications, a unique fixed point is a required property of the model. It means that different initial stimuli should lead to the same equilibrium state. On the other hand, there are applications (for example pattern recognition) where the FCM should have more than one equilibrium point. In other words, in the first case, global stability is a required property, while in the second case we should avoid globally stable models. The simple analytical results help FCM users to decide about some model parameters before evaluation of the full model, decreasing the number of trial-and-error simulations.

### Appendix 1

Proof of Theorem 6:

First of all we simplify the expression given in Theorem 3.2 of [21] where we use  $x = \frac{\bar{\lambda}(n)}{4}$  for simplicity.

$$(1 - x)^n - \sum_{i=1}^n b_i C_i^n x^i \tag{38}$$

$$= \sum_{k=0}^n (-1)^k \cdot x^k \cdot \binom{n}{k} - \sum_{i=1}^n b_i \binom{n}{i} x^i \tag{39}$$

$$= 1 + \sum_{k=1}^n (-1)^k \cdot x^k \cdot \binom{n}{k} - \sum_{i=1}^n b_i \binom{n}{i} x^i \tag{40}$$

$$= 1 + \sum_{i=1}^n x^i \left( (-1)^i \binom{n}{i} - b_i \binom{n}{i} \right) \tag{41}$$

using that  $b_i = i b_{i-1} + (-1)^i$  we get that 
$$\tag{42}$$

$$= 1 + \sum_{i=1}^n x^i \left( -i \cdot b_{i-1} \binom{n}{i} \right) \tag{43}$$

which means that  $\frac{\bar{\lambda}(n)}{4}$  is root of the polynomial

$$p_n(x) = \sum_{i=1}^n i \cdot b_{i-1} \binom{n}{i} x^i - 1. \tag{44}$$

The first few values of the recursive sequence  $b_i$ :

$$b_0 = 1, \quad b_1 = 0, \quad b_2 = 1, \quad b_3 = 2, \quad b_4 = 9, \quad b_5 = 44 \tag{45}$$

The polynomials  $p_n(x)$ , the values of  $\bar{\lambda}(n)$  and  $4/n$  for  $n = 1, \dots, 5$  are the following:

$n$	$p_n(x)$	$\bar{\lambda}(n)$	$4/n$
1	$x - 1$	4	4
2	$2x - 1$	2	2
3	$3x^3 + 3x - 1$	1.2199	1.3333
4	$8x^4 + 12x^3 + 4x - 1$	0.8624	1
5	$45x^5 + 40x^4 + 30x^3 + 5x - 1$	0.6624	0.8

As it can be recognized, the upper bound provided by Theorem 5 is slightly higher than that given by the recursion in [21]. The following proof shows that it is true for every value of  $n$ . The main ideas of the proof are listed below:

- Since the polynomial  $p_n(x)$  has positive coefficient (except for the constant term), it is strictly increasing if  $x > 0$ .
- $\frac{\bar{\lambda}(n)}{4}$  is a positive root of polynomial  $p_n(x)$ , so  $p_n\left(\frac{\bar{\lambda}(n)}{4}\right) = 0$ .
- This implies that if  $p_n\left(\frac{4/n}{4}\right) > 0$  then  $\frac{4/n}{4} > \frac{\bar{\lambda}(n)}{4}$ .
- So we have to show that  $p_n(1/n) > 0$ .

We have seen (and simple computations show) that for  $n = 1$  and  $n = 2$   $\bar{\lambda}(n) = 4/n$ . Let us consider the case when  $n \geq 3$ , so we have to prove that

$$\sum_{i=1}^n i \cdot b_{i-1} \binom{n}{i} \left(\frac{1}{n}\right)^i - 1 > 0 \quad \text{for every } n \geq 3. \tag{46}$$

The sequence  $b_i$  is monotone increasing if  $i \geq 1$ , which can be easily proved by induction. Moreover, the value of  $i \cdot b_{i-1}$  is greater than or equal to 3 for every  $i \geq 3$ :

$$i \cdot b_{i-1} \geq i \cdot b_2 \geq 3 \cdot b_2 = 3. \tag{47}$$

We give a lower estimation of  $\sum_{i=1}^n i \cdot b_{i-1} \binom{n}{i} \left(\frac{1}{n}\right)^i$  by replacing the value of  $i \cdot b_{i-1}$  with 3 for  $i \geq 3$  and taking the exact value for  $i = 1, 2$ :

$$\sum_{i=1}^n i \cdot b_{i-1} \binom{n}{i} \left(\frac{1}{n}\right)^i \tag{48}$$

$$\geq 3 \cdot \sum_{i=1}^n \binom{n}{i} \left(\frac{1}{n}\right)^i - 2 \cdot \binom{n}{1} \left(\frac{1}{n}\right)^1 - 3 \cdot \binom{n}{2} \left(\frac{1}{n}\right)^2 \tag{49}$$

$$= 3 \cdot \sum_{i=1}^n \binom{n}{i} \left(\frac{1}{n}\right)^i - 2 - 3 \cdot \frac{n-1}{2n} \tag{50}$$

$$= 3 \cdot \sum_{i=0}^n \binom{n}{i} \left(\frac{1}{n}\right)^i - 5 - 3 \cdot \frac{n-1}{2n} \tag{51}$$

Now we will prove that the last expression is greater than one:

$$3 \cdot \sum_{i=0}^n \binom{n}{i} \left(\frac{1}{n}\right)^i - 5 - 3 \cdot \frac{n-1}{2n} > 1 \tag{52}$$

$$3 \cdot \sum_{i=0}^n \binom{n}{i} \left(\frac{1}{n}\right)^i > 6 + 3 \cdot \frac{n-1}{2n} \tag{53}$$

$$\left(1 + \frac{1}{n}\right)^n > 2.5 - \frac{1}{2n} \tag{54}$$

which is true for every  $n \geq 3$ .

Actually, we have just proved that the lower estimation of the sum is greater than one, which implies that  $p_n(1/n) > 0$  for every  $n \geq 3$ . As a consequence, we can conclude that  $\frac{4}{n} \geq \bar{\lambda}(n)$  (equality holds if and only if  $n = 1$  or  $n = 2$ ).

### Appendix 2

As a kind of generalization of the findings of [3], in [25] the authors stated the theorem below about the number of equilibrium points (limits of the iterations) of a FCM (Theorem 2 of [25]), literally:

‘ There is one and only one solution for any concept value  $A_i$  of any FCM where the sigmoid function  $f = 1/(1 + e^{-cx})$  is used, if:

$$\left(\sum_{i=1}^n (c_i \ell_i \|w_i\|)^2\right)^{1/2} < 1 \tag{55}$$

where  $w_i$  is the  $i$ th row of matrix  $W$ ,  $\|w_i\|$  is the  $L_2$  norm of  $w_i$ ,  $\ell_i$  is the inclination of function  $f$  equal to  $\ell_i = \frac{c_i}{e^{c_i w_i A}} f^2(c_i w_i \cdot A)$ , and  $c_i$  is the  $c_l$  factor of function  $f$  corresponding to  $A_i$  concept’.

In the following, we will show that this theorem is not valid in its current form. First we give an upper estimation of the formula given on the left-hand side of Eq. (55), then we provide numerical counterexamples by pointing out that

the requirements stated in the theorem are fulfilled, but the FCM does not converge to a unique fixed point.

The so-called inclination parameter (or the slope)  $\ell_i = \frac{c_i}{e^{c_i w_i A}} f^2(c_i w_i \cdot A)$  equals the derivative of the sigmoid function at point  $w_i \cdot A$ :

$$\left(\frac{1}{1 + e^{-c_i x}}\right)' = \frac{c_i e^{-c_i x}}{(1 + e^{-c_i x})^2} \tag{56}$$

$$= \frac{c_i}{e^{c_i x}} \left(\frac{1}{1 + e^{-c_i x}}\right)^2 = \frac{c_i}{e^{c_i x}} f^2(x) \tag{57}$$

On the other hand, it is easy to check that

$$\left(\frac{1}{1 + e^{-c_i x}}\right)' = \frac{c_i e^{-c_i x}}{(1 + e^{-c_i x})^2} = c_i \cdot f(x)(1 - f(x)) \tag{58}$$

The maximum value of the derivative of the sigmoid function is  $c_i/4$  (attained at  $f(x) = 0.5$ , i.e. at  $x = 0$ ). Using this value, we can give an upper estimation of the left-hand side of the inequality in the theorem (here  $\|w_i\|$  is the Euclidean norm of the  $i$ th row of  $W$ , i.e.

$$\|w_i\| = \left(\sum_{j=1}^n w_{ij}^2\right)^{1/2}:$$

$$\left(\sum_{i=1}^n (c_i \ell_i \|w_i\|)^2\right)^{1/2} \leq \left(\sum_{i=1}^n \left(\frac{c_i^2}{4} \|w_i\|^2\right)\right)^{1/2} \tag{59}$$

$$= \frac{1}{4} \left(\sum_{i=1}^n (c_i^2 \|w_i\|)^2\right)^{1/2} \tag{60}$$

If  $c_i = \lambda$  for every  $i$ , it becomes:

$$\frac{1}{4} \left(\sum_{i=1}^n (c_i^2 \|w_i\|)^2\right)^{1/2} = \frac{1}{4} \left(\sum_{i=1}^n (\lambda^2 \|w_i\|)^2\right)^{1/2} \tag{61}$$

$$= \frac{\lambda^2}{4} \left(\sum_{i=1}^n \|w_i\|^2\right)^{1/2} \tag{62}$$

In the following, we show that there exist FCMs, such that the conditions of Theorem 2 of [25] are fulfilled, but the iteration process does not lead to a unique fixed point (equilibrium point). First we show a counterexample with limit cycle, then an other counterexample with multiple fixed points. Although these are artificial counterexamples, not real-life scenarios, they show that the statement of [25] is not valid.

We applied the Bacterial Evolutionary Algorithm (BEA) [30] to find an FCM model (more specifically, a connection matrix and parameter  $\lambda$ ) that behaves in the assumed way in order to confirm our hypothesis. BEA is a member of the well-known family of evolutionary algorithms. It is a global optimizer, which can be used if an approximate solution is acceptable. The algorithm can solve any multi-

modal, non-continuous, nonlinear or high-dimensional problem, but the original goal of Nawa and Furuhashi, the researchers who suggested this algorithm, was to optimize the parameters of fuzzy systems [31, 32]. BEA mimics the process of the evolution of bacteria, which explains its name. Like other evolutionary algorithms, BEA works with a collection of possible solutions. These are called the population of ‘bacteria’. The algorithm develops the consecutive generations of populations until the termination criterion is fulfilled. In the simplest case, it can be a limit on the number of generations or on the objective value of the best bacterium. The last population, or at least some of the best bacteria of it can be considered as result. The current population is based on the previous population, and is created by two main operators: bacterial mutation and gene transfer. The former operator explores the search space with random modifications of bacteria, while the latter tries to combine the genetic information of better bacteria with worse bacteria in the hope that it may increase the objective/fitness value of worse bacteria. With other words, gene transfer does the exploitation of genetic data.

Bacterial mutation optimizes the bacteria one by one. First, it creates  $K$  clones of every bacteria of the current population, then it iterates over the genes of clones in a random order to modify them. The modification is made in a random way, too, and the clones are evaluated after every single gene modification. If the best clone outperforms the original one, it will be replaced by the new one. At the end, all clones are dropped.

The gene transfer operator divides the population into two, equally sized parts. One of these contains the bacteria with better objective values, which is called the ‘superior half’. The name of the other part is ‘inferior part’. The operator chooses  $T$  times a bacterium from the inferior half and overwrites some of its randomly selected genes with the genes of an other bacterium of the superior half. The modified bacterium must be evaluated and the elements of population halves must be determined again. After a successful modification the bacterium has a chance to move to the superior half.

The authors created a computer program based on the above expounded BEA to find a connection matrix with given properties according to their hypothesis. The bacteria represent different connection matrices. The first population is generated randomly, the following generations are created by the operators of BEA using the objective value of bacteria. The calculation of this value has several steps. First of all, simulations have to be performed in order to explore the behaviour of the model, differentiate fixed point attractors (FP), limit cycles and chaotic behaviours. A hundred-element set of initial state vectors, also called ‘scenarios’ is generated randomly when the program starts and this set is used in all simulations.

Finally, numerous counterexamples were generated, this one below was created for simple demonstration of our statement. One may observe that the non-diagonal elements of weight matrix  $W$  have the same absolute value. It is not an essential feature of the counterexamples, the main reason of this is that we have chosen simple and clear-cut example for demonstration.

Consider weight matrix  $W \in \mathbb{R}^{7 \times 7}$ , which has only zero elements in its diagonal. It means that the iteration rule of the corresponding FCM does not contain self-feedback.

$$W_1 = \begin{pmatrix} 0 & -0.9 & 0.9 & 0.9 & -0.9 & 0.9 & -0.9 \\ -0.9 & 0 & 0.9 & 0.9 & -0.9 & 0.9 & -0.9 \\ 0.9 & 0.9 & 0 & -0.9 & 0.9 & -0.9 & 0.9 \\ 0.9 & 0.9 & -0.9 & 0 & 0.9 & -0.9 & 0.9 \\ -0.9 & -0.9 & 0.9 & 0.9 & 0 & 0.9 & -0.9 \\ 0.9 & 0.9 & -0.9 & -0.9 & 0.9 & 0 & 0.9 \\ -0.9 & -0.9 & 0.9 & 0.9 & -0.9 & 0.9 & 0 \end{pmatrix} \tag{63}$$

If  $\lambda = 0.8$ , then  $\frac{\lambda^2}{4} \left( \sum_{i=1}^n \|w_i\|^2 \right)^{1/2} = 0.9332$ , which means that this scenario meets condition Eq. (55). Contrary to the statement of Theorem 2 of [25], the iteration produces a limit cycle, not an equilibrium point (the initial activation vector was  $A(0) = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]^T$ ). This limit cycle has two elements:

$$LC_1 = \begin{bmatrix} 0.39897 \\ 0.39897 \\ 0.78157 \\ 0.78157 \\ 0.39897 \\ 0.78157 \\ 0.39897 \end{bmatrix} \quad LC_2 = \begin{bmatrix} 0.69559 \\ 0.69559 \\ 0.50590 \\ 0.50590 \\ 0.69559 \\ 0.50590 \\ 0.69559 \end{bmatrix} \tag{64}$$

Now let us see an example with multiple fixed points. The weight matrix is

$$W_2 = \begin{pmatrix} 0 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 0 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 0 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 0 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 0 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 0 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 0 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 0 \end{pmatrix} \tag{65}$$

If  $\lambda = 0.7$ , then  $\frac{\lambda^2}{4} \left( \sum_{i=1}^n \|w_i\|^2 \right)^{1/2} = 0.9167$ , which means that this scenario also meets condition Eq. (55). On the other hand, if the initial activation vector is

$A^{(0)} = [0, 0, 0, 0, 0, 0, 0, 0]^T$ , then the iteration converges to  $FP_1$ , but if  $A^{(0)} = [1, 0, 0, 0, 0, 0, 0, 0]^T$ , then the iteration converges to  $FP_2$  and for  $A^{(0)} = [0, 0, 0, 0, 0, 0, 0, 1]^T$  the limit is  $FP_3$  (results are rounded to four decimals):

$$FP_1 = \begin{bmatrix} 0.4260 \\ 0.4260 \\ 0.4260 \\ 0.4260 \\ 0.4260 \\ 0.4260 \\ 0.4260 \\ 0.4260 \end{bmatrix} \quad FP_2 = \begin{bmatrix} 0.7847 \\ 0.1266 \\ 0.7847 \\ 0.1266 \\ 0.7847 \\ 0.1266 \\ 0.7847 \\ 0.1266 \end{bmatrix} \quad FP_3 = \begin{bmatrix} 0.1266 \\ 0.7847 \\ 0.1266 \\ 0.7847 \\ 0.1266 \\ 0.7847 \\ 0.1266 \\ 0.7847 \end{bmatrix} \quad (66)$$

A rigorous mathematical analysis leads to the conclusion that the weakness of the statement of [25] comes from the mistake that the contraction property was applied using local values of the derivative, while a global upper bound should be applied. Moreover, an algebraic mistake has also occurred in the derivation of the result, which caused a duplicated presence of the parameter of the sigmoid function ( $c_i$ ).

**Acknowledgements** This research was supported by National Research, Development and Innovation Office (NKFIH) K124055.

**Funding** Open access funding provided by Széchenyi István University (SZE).

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Axelrod R (1976) Structure of decision: the cognitive maps of political elites. Princeton University Press, Princeton
2. Boutalis Y, Kottas T, Christodoulou M (2008) On the existence and uniqueness of solutions for the concept values in fuzzy cognitive maps. In: 2008 47th IEEE conference on decision and control, pp 98–104. <https://doi.org/10.1109/CDC.2008.4738897>
3. Boutalis Y, Kottas TL, Christodoulou M (2009) Adaptive estimation of fuzzy cognitive maps with proven stability and parameter convergence. IEEE Trans Fuzzy Syst 17(4):874–889
4. Buruzs A, Hatwagner MF, Kóczy LT (2014) Fuzzy cognitive maps and bacterial evolutionary algorithm approach to integrated waste management systems. J Adv Comput Intell Inform 18(4):538–548
5. Carvalho JP (2010) On the semantics and the use of fuzzy cognitive maps in social sciences. In: WCCI 2010 IEEE World Congress on Computational Intelligence, pp 2456–2461
6. Carvalho JP, Tomé J (2002) Issues on the stability of fuzzy cognitive maps and rule-based fuzzy cognitive maps. In: Fuzzy Information Processing Society, 2002. Proceedings. NAFIPS. 2002 annual meeting of the North American. IEEE, pp 105–110
7. Chen Y, Mazlack LJ, Minai AA, Lu LJ (2015) Inferring causal networks using fuzzy cognitive maps and evolutionary algorithms with application to gene regulatory network reconstruction. Appl Soft Comput 37:667–679
8. Dickerson JA, Kosko B (1994) Virtual worlds as fuzzy cognitive maps. Presence Teleoperators Virtual Environ 3(2):173–189
9. Ferreira FAF, Jalali MS, Ferreira JJM, Stankevičienė J, Marques CSE (2016) Understanding the dynamics behind bank branch service quality in Portugal: pursuing a holistic view using fuzzy cognitive mapping. Serv Bus 10(3):469–487. <https://doi.org/10.1007/s11628-015-0278-x>
10. Fons S, Achari G, Ross T (2004) A fuzzy cognitive mapping analysis of the impacts of an eco-industrial park. J Intell Fuzzy Syst 15(2):75–88
11. Groumpos PP (2010) Fuzzy cognitive maps: basic theories and their application to complex systems. In: Fuzzy cognitive maps. Springer, pp 1–22
12. Hanan D, Burnley S, Cooke D (2013) A multi-criteria decision analysis assessment of waste paper management options. Waste Manag 33:566–573
13. Harmati IÁ, Hatwagner MF, Kóczy LT (2018) On the existence and uniqueness of fixed points of fuzzy cognitive maps. In: International conference on information processing and management of uncertainty in knowledge-based systems. Springer, pp 490–500 (2018)
14. Harmati IÁ, Kóczy LT (2020) On the convergence of input-output fuzzy cognitive maps. In: International joint conference on rough sets. Springer, pp 449–461
15. Hatwagner MF, Vastag G, Niskanen VA, Kóczy LT (2018) Improved behavioral analysis of fuzzy cognitive map models. In: International conference on artificial intelligence and soft computing. Springer, pp 630–641 (2018)
16. Homenda W, Jastrzebska A, Pedrycz W (2014) Computer information systems and industrial management: 13th IFIP TC8 International Conference, CISIM 2014, Ho Chi Minh City, Vietnam, November 5–7, 2014. Proceedings, chap. Time series modeling with fuzzy cognitive maps: simplification strategies. Springer, Berlin, pp 409–420
17. Horn RA, Johnson CR (2013) Matrix analysis, 2nd edn. Cambridge University Press, New York
18. Hossain S, Brooks L (2008) Fuzzy cognitive map modelling educational software adoption. Comput Educ 51(4):1569–1588
19. Khan MS, Khor S, Chong A (2004) Fuzzy cognitive maps with genetic algorithm for goal-oriented decision support. Int J Uncertain Fuzziness Knowl Based Syst 12(supp02):31–42
20. Khan MS, Quaddus M (2004) Group decision support using fuzzy cognitive maps for causal reasoning. Group Decis Negot 13(5):463–480. <https://doi.org/10.1023/B:GRUP.0000045748.89201.f3>
21. Knight CJ, Lloyd DJ, Penn AS (2014) Linear and sigmoidal fuzzy cognitive maps: an analysis of fixed points. Appl Soft Comput 15:193–202

22. Kok K (2009) The potential of fuzzy cognitive maps for semi-quantitative scenario development, with an example from Brazil. *Glob Environ Change* 19(1):122–133
23. Kosko B (1986) Fuzzy cognitive maps. *Int J Man Mach Stud* 24:65–75
24. Kosko B (1992) *Neural networks and fuzzy systems*. Prentice-Hall, Upper Saddle River
25. Kottas TL, Boutalis Y, Christodoulou M (2012) Bi-linear adaptive estimation of fuzzy cognitive networks. *Appl Soft Comput* 12(12):3736–3756
26. Koulouriotis DE, Diakoulakis IE, Emiris DM (2001) A fuzzy cognitive map-based stock market model: synthesis, analysis and experimental results. In: 10th IEEE international conference on fuzzy systems (Cat. No. 01CH37297), vol 1, pp 465–468. <https://doi.org/10.1109/FUZZ.2001.1007349>
27. Lee IK, Kim HS, Cho H (2012) Design of activation functions for inference of fuzzy cognitive maps: application to clinical decision making in diagnosis of pulmonary infection. *Healthc Inform Res* 18(2):105–114
28. Lee IK, Kwon SH (2010) Design of sigmoid activation functions for fuzzy cognitive maps via Lyapunov stability analysis. *IEICE Trans Inf Syst* 93(10):2883–2886
29. Luo C, Song X, Zheng Y (2020) Algebraic dynamics of k-valued fuzzy cognitive maps and its stabilization. *Knowl Based Syst* 209:106424
30. Nawa NE, Furuhashi T (1998) A study on the effect of transfer of genes for the bacterial evolutionary algorithm. In: 1998 second international conference on knowledge-based intelligent electronic systems, 1998. *Proceedings KES'98*, vol 3. IEEE, pp 585–590
31. Nawa NE, Furuhashi T (1999) Fuzzy system parameters discovery by bacterial evolutionary algorithm. *IEEE Trans Fuzzy Syst* 7(5):608–616
32. Nawa NE, Hashiyama T, Furuhashi T, Uchikawa Y (1997) A study on fuzzy rules discovery using pseudo-bacterial genetic algorithm with adaptive operator. In: *IEEE International Conference on Evolutionary Computation*. IEEE, pp 589–593 (1997)
33. Nápoles G, Grau I, Bello R, Grau R (2014) Two-steps learning of fuzzy cognitive maps for prediction and knowledge discovery on the HIV-1 drug resistance. *Expert Syst Appl* 41(3):821–830. <https://doi.org/10.1016/j.eswa.2013.08.012> (**Methods and applications of artificial and computational intelligence**)
34. Papageorgiou E, Kontogianni A (2012) Using fuzzy cognitive mapping in environmental decision making and management: a methodological primer and an application. INTECH Open Access Publisher
35. Pathak HK (2018) *An introduction to nonlinear analysis and fixed point theory*. Springer
36. Rudin W (1976) *Principles of mathematical analysis*, 3rd edn. McGraw-Hill, New York
37. Salmeron JL (2010) Fuzzy cognitive maps-based it projects risks scenarios. In: *Fuzzy cognitive maps*. Springer, pp 201–215
38. Salmeron JL (2012) Fuzzy cognitive maps for artificial emotions forecasting. *Appl Soft Comput* 12(12):3704–3710
39. Stylios CD, Groumpos PP (2004) Modeling complex systems using fuzzy cognitive maps. *IEEE Trans Syst Man Cybern Part A Syst Hum* 34(1):155–162
40. Stylios CD, Groumpos PP et al (1999) Mathematical formulation of fuzzy cognitive maps. In: *Proceedings of the 7th Mediterranean Conference on Control and Automation*, pp 2251–2261
41. Tsadiras AK (2008) Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. *Inf Sci* 178(20):3880–3894

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.