

# Verification of temporal-causal network models by mathematical analysis

Jan Treur<sup>1</sup>

Received: 28 February 2016 / Accepted: 21 March 2016 / Published online: 16 April 2016  
© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** Usually dynamic properties of models can be analysed by conducting simulation experiments. But sometimes, as a kind of prediction properties can also be found by calculations in a mathematical manner, without performing simulations. Examples of properties that can be explored in such a manner are:

- whether some values for the variables exist for which no change occurs (stationary points or equilibria), and how such values may depend on the values of the parameters of the model and/or the initial values for the variables
- whether certain variables in the model converge to some limit value (equilibria) and how this may depend on the values of the parameters of the model and/or the initial values for the variables
- whether or not certain variables will show monotonically increasing or decreasing values over time (monotonicity)
- how fast a convergence to a limit value takes place (convergence speed)
- whether situations occur in which no convergence takes place but in the end a specific sequence of values is repeated all the time (limit cycle)

Such properties found in an analytic mathematical manner can be used for verification of the model by checking them for the values observed in simulation experiments. If one of these properties is not fulfilled, then there will be some

error in the implementation of the model. In this paper some methods to analyse such properties of dynamical models will be described and illustrated for the Hebbian learning model, and for dynamic connection strengths in social networks. The properties analysed by the methods discussed cover equilibria, increasing or decreasing trends, recurring patterns (limit cycles), and speed of convergence to equilibria.

## 1 Introduction

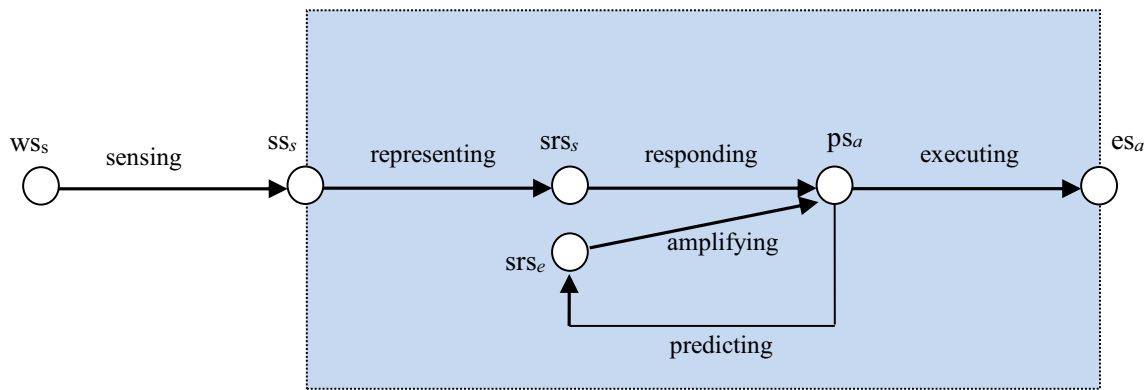
Usually dynamic properties of dynamic models can be analysed by conducting simulation experiments. But sometimes, as a kind of prediction properties can also be found by calculations in a mathematical manner, without performing simulations. Examples of properties that can be explored in such a manner are:

- whether some values for the variables exist for which no change occurs (stationary points or equilibria), and how such values may depend on the values of the parameters of the model and/or the initial values for the variables
- whether certain variables in the model converge to some limit value (equilibria) and how this may depend on the values of the parameters of the model and/or the initial values for the variables
- whether or not certain variables will show monotonically increasing or decreasing values over time (monotonicity)
- whether situations occur in which no convergence takes place but in the end a specific sequence of values is repeated all the time (limit cycle)

Mathematical techniques addressing such questions have been developed, starting with Poincaré [12, 13]; see also [3, 9, 11], and [7] for a historical perspective. Such types of prop-

✉ Jan Treur  
j.treur@vu.nl

<sup>1</sup> VU University Amsterdam, Behavioral Informatics Group, Amsterdam, The Netherlands



**Fig. 1** Conceptual representation of an example model

erties found in an analytic mathematical manner can be used for verification of the model by checking them for the values observed in simulation experiments. If one of these properties is not fulfilled, then there will be some error in the implementation of the model. This particular use of mathematical analysis is the focus of this paper. In this paper some methods to analyse such properties of temporal-causal network models will be described and illustrated for some example models, including a Hebbian learning model, and a model for dynamic connection strengths in social networks. The properties analysed by the methods discussed cover equilibria, increasing or decreasing trends, and recurring patterns: limit cycles.

To get the idea, first the general set up is discussed in Sect. 2. This is illustrated in Sect. 3 by an analysis of a simple example as discussed in [17], Section 2.4.1, using sum and identity combination functions. In simulations it is observed for this example model that when a constant stimulus level occurs in the world, for each state its activation value increases from 0 to some value that is then kept forever, until the stimulus disappears: an equilibrium state. In subsequent sections three more general examples of this type of analysis for which equilibrium states occur are addressed: for a scaled sum combination function (Sect. 4), for Hebbian learning (Sect. 5), and for dynamic networks based on the homophily principle (Sect. 6). In Sect. 7 the analysis is discussed for a case in which no equilibrium state occurs, but instead a limit cycle pattern emerges.

## 2 How to verify a temporal-causal network model by mathematical analysis

A stationary point of a state occurs at some point in time if for this time point no change occurs: the graph is horizontal at that point. Stationary points are usually maxima or minima (peaks or dips) but sometimes also other stationary points may occur. An equilibrium occurs when for all states

no change occurs. From the difference or differential equations describing the dynamics for a model it can be analysed when stationary points or equilibria occur. Moreover, it can be found when a certain state is increasing or decreasing when a state is not in a stationary point or equilibrium. First a definition for these notions is expressed; for example, see [3, 9, 11–13].

**Definition** (*increase, decrease, stationary point and equilibrium*) Let  $Y$  be a state

- $Y$  has a *stationary point* at  $t$  if  $\mathbf{d}Y(t)/\mathbf{d}t = 0$
- $Y$  is *increasing* at  $t$  if  $\mathbf{d}Y(t)/\mathbf{d}t > 0$
- $Y$  is *decreasing* at  $t$  if  $\mathbf{d}Y(t)/\mathbf{d}t < 0$

The model is in *equilibrium* at  $t$  if every state  $Y$  of the model has a stationary point at  $t$ .

To illustrate these notions, consider the example from [17], with conceptual representation depicted here in Fig. 1, and an example simulation shown in Fig. 2.

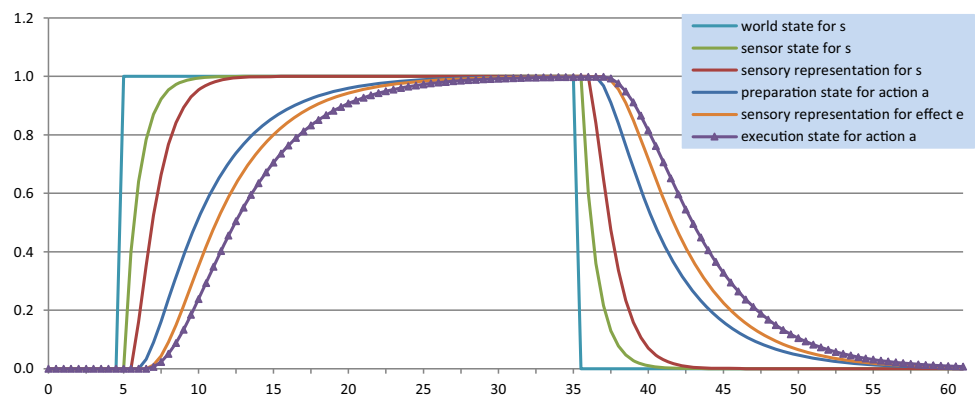
The systematic transformation from a conceptual representation of a temporal-causal model (as depicted in Fig. 1) into a numerical representation of this temporal-causal model works as follows [17]:

- At each time point  $t$  each state  $Y$  in the model has a real number value in the interval  $[0, 1]$ , denoted by  $Y(t)$
- At each time point  $t$  each state  $X$  connected to state  $Y$  has an *impact* on  $Y$  defined as  $\mathbf{impact}_{X,Y}(t) = \omega_{X,Y} X(t)$  where  $\omega_{X,Y}$  is the weight of the connection from  $X$  to  $Y$
- The *aggregated impact* of multiple states  $X_i$  on  $Y$  at  $t$  is determined using a *combination function*  $\mathbf{c}_Y(\cdot)$ :

$$\begin{aligned} \mathbf{aggimpact}_Y(t) &= \mathbf{c}_Y(\mathbf{impact}_{X_1,Y}(t), \dots, \mathbf{impact}_{X_k,Y}(t)) \\ &= \mathbf{c}_Y(\omega_{X_1,Y} X_1(t), \dots, \omega_{X_k,Y} X_k(t)) \end{aligned}$$

where  $X_i$  are the states with connections to state  $Y$

**Fig. 2** Simulation example for the model depicted in Fig. 1 using identity and sum combination functions for all states



- The effect of  $\mathbf{aggimpact}_Y(t)$  on  $Y$  is exerted over time gradually, depending on *speed factor*  $\eta_Y$ :

$$Y(t + \Delta t) = Y(t) + \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)]\Delta t$$

or

$$dY(t)/dt = \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)]$$

- Thus, the following *difference* and *differential equation* for  $Y$  are obtained:

$$Y(t + \Delta t) = Y(t) + \eta_Y[\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)]\Delta t$$

$$dY(t)/dt = \eta_Y[\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)]$$

For more details, see [17].

Combination functions used in this simple example are the scaled sum function and the identity function, and all connections have weight 1, except the connections to  $ps_a$ , which have weight 0.5.

In Fig. 2 it can be seen that as a result of the stimulus all states are increasing until time point 35, after which they start to decrease as the stimulus disappears. Just before time point 35 all states are almost stationary. If the stimulus is not taken away after this time point this trend is continued, and an equilibrium state is approximated. The question then is whether these observations based on one or more simulation experiments are in agreement with a mathematical analysis.

If it is found out that they are in agreement with the mathematical analysis, then this provides some extent of evidence that the implemented model is correct. If they turn out not to be in agreement with the mathematical analysis, then this indicates that probably there is something wrong, and further inspection and correction has to be initiated.

Considering the differential equation for a temporal-causal network model more specific criteria can be found:

$$dY(t)/dt = \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)]$$

with

$$\mathbf{aggimpact}_Y(t) = \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t))$$

and  $X_1, \dots, X_k$  the states connected toward  $Y$

For example, it can be concluded that

$$dY(t)/dt > 0 \Leftrightarrow \eta_Y[\mathbf{aggimpact}_Y(t) - Y(t)] > 0$$

$$\Leftrightarrow \mathbf{aggimpact}_Y(t) > Y(t)$$

$$\Leftrightarrow \mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t)$$

In this manner the following criteria can be found.

### 2.1 Criteria for a temporal-causal network model: increase, decrease, stationary point and equilibrium

Let  $Y$  be a state and  $X_1, \dots, X_k$  the states connected toward  $Y$ . Then the following hold

$Y$ has a stationary point at $t$	$\Leftrightarrow$	$\mathbf{aggimpact}_Y(t) = Y(t)$	$\Leftrightarrow$	$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$
$Y$ is increasing at $t$	$\Leftrightarrow$	$\mathbf{aggimpact}_Y(t) > Y(t)$	$\Leftrightarrow$	$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t)$
$Y$ is decreasing at $t < 0$	$\Leftrightarrow$	$\mathbf{aggimpact}_Y(t) < Y(t)$	$\Leftrightarrow$	$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) < Y(t)$
The model is in equilibrium at $t$	$\Leftrightarrow$	$\mathbf{aggimpact}_Y(t) = Y(t)$ for every state $Y$	$\Leftrightarrow$	$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$ for every state $Y$

These criteria can be used to verify (the implementation of) the model based on inspection of stationary points or equilibria in the following two different manners. Note that in a given simulation the stationary points that are identified are usually approximately stationary; how closely they are approximated depends on different aspects, for example on the step size, or on how long the simulation is done.

## 2.2 Verification by checking the criteria through substitution values from a simulation in the criteria

1. Generate a simulation
2. For a number of states  $Y$  identify stationary points with their time points  $t$  and state values  $Y(t)$
3. For each of these stationary points for a state  $Y$  at time  $t$  identify the values  $X_1(t), \dots, X_k(t)$  at that time of the states  $X_1, \dots, X_k$  connected toward  $Y$
4. Substitute all these values  $Y(t)$  and  $X_1(t), \dots, X_k(t)$  in the criterion  $\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$
5. If the equation holds (for example, with an accuracy  $< 10^{-2}$ ), then this test succeeds, otherwise it fails
6. If this test fails, then it has to be explored where the error can be found

This verification method can be illustrated for the example of Figs. 1 and 2 as follows. For example, consider state  $ps_a$  with numerical representation

$$\begin{aligned} ps_a(t + \Delta t) = ps_a(t) \\ + \eta_{ps_a} [\omega_{\text{respondingSRS}_s}(t) \\ + \omega_{\text{amplifyingSRS}_e}(t) - ps_a(t)] \Delta t \end{aligned}$$

The equation expressing that a state of  $ps_a$  is stationary at time  $t$  is

$$\omega_{\text{respondingSRS}_s}(t) + \omega_{\text{amplifyingSRS}_e}(t) = ps_a(t)$$

At time point  $t = 35$  (where all states are close to stationary) the following values occur:  $ps_a(35) = 0.99903$ ,  $srs_s(35) = 1.00000$  and  $srs_e(35) = 0.99863$ ; moreover  $\omega_{\text{responding}} = \omega_{\text{amplifying}} = 0.5$ . All these values can be substituted in the above equation:

$$\begin{aligned} 0.5 \times 1.00000 + 0.5 \times 0.99863 &= 0.99903 \\ 0.999315 &= 0.99903 \end{aligned}$$

It turns out that the equation is fulfilled with accuracy  $< 10^{-3}$ . This gives some evidence that the model as implemented indeed does what it was meant to do. If this is done for all other states, similar outcomes are found. This gives still more evidence. The step size  $\Delta t$  for the simulation here was 0.5, which is even not so small. For still more accurate results it is advisable to choose a smaller step size. So, having the equations for stationary points for all states provides a

means to verify the implemented model in comparison to the model description. The equations for stationary points themselves can easily be obtained from the model description in a systematic manner.

Note that this method works without having to solve the equations, only substitution takes place; therefore, it works for any choice of combination function. Moreover, note that the method also works when there is no equilibrium but the values of the states fluctuate all the time, according to a recurring pattern (a limit cycle). In such cases for each state there are maxima (peaks) and minima (dips) which also are stationary. The method can be applied to such a type of stationary points as well; here it is still more important to choose a small step size as each stationary point occurs at just one time point. In Sect. 7 it will be discussed how the approach can be applied to such limit cycles.

There is still another method possible that is sometimes proposed; this method is applied for the case of an equilibrium (where all states have a stationary point simultaneously), and is based on solving the equations for the equilibrium values first. This can provide explicit expressions for equilibrium values in terms of the parameters of the model. Such expressions can be used to predict equilibrium values for specific simulations, based on the choice of parameter values. This method provides more than the previous method, but a major drawback is that it cannot be applied in all situations. For example, when logistic combination functions are used it cannot be applied. However, in some cases it still can be useful. The method goes as follows.

## 2.3 Verification by solving the equilibrium equations and comparing predicted equilibrium values to equilibrium values in a simulation

1. Consider the equilibrium equations for all states  $Y$ :

$$\mathbf{c}_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$$

2. Leave the  $t$  out and denote the values as constants

$$\mathbf{c}_Y(\omega_{X_1,Y}\underline{\mathbf{X}}_1, \dots, \omega_{X_k,Y}\underline{\mathbf{X}}_k) = \underline{\mathbf{Y}}$$

An equilibrium is a solution  $\underline{\mathbf{X}}_1, \dots, \underline{\mathbf{X}}_k$  of the following set of  $n$  equilibrium equations in the  $n$  states  $X_1, \dots, X_n$  of the model:

$$\mathbf{c}_{X_1}(\omega_{X_1,X_1}\underline{\mathbf{X}}_1, \dots, \omega_{X_n,X_1}\underline{\mathbf{X}}_n) = \underline{\mathbf{X}}_1$$

...

$$\mathbf{c}_{X_n}(\omega_{X_1,X_n}\underline{\mathbf{X}}_1, \dots, \omega_{X_n,X_n}\underline{\mathbf{X}}_n) = \underline{\mathbf{X}}_n$$

3. Solve these equations mathematically in an explicit analytical form: for each state  $X_i$  a mathematical formula

$\underline{X}_i = \dots$  in terms of the parameters of the model (connection weights and parameters in the combination function  $\mathbf{c}_{X_i}(\cdot)$ , such as the steepness  $\sigma$  and threshold  $\tau$  in a logistic sum combination function); more than one solution is possible

4. Generate a simulation
5. Identify equilibrium values in this simulation
6. If for all states  $Y$  the predicted value  $\underline{Y}$  from a solution of the equilibrium equations equals the value for  $Y$  obtained from the simulation (for example, with an accuracy  $<10^{-2}$ ), then this test succeeds, otherwise it fails
7. If this test fails, then it has to be explored where the error can be found

In Sect. 2.3 it will be illustrated how this method works for the example depicted in Figs. 1 and 2. In general, whether or not the equilibrium equations can be solved in an explicit analytical manner strongly depends on the form of the combination functions  $\mathbf{c}_Y(\dots)$ . In a number of specific cases explicit analytical solutions can be found. Three examples of this are addressed in subsequent sections:

- for a (scaled) sum combination function (Sects. 3 and 4)
- for Hebbian learning (Sect. 5)
- for dynamic networks based on the homophily principle (Sect. 6)

However, there are also many cases in which an explicit analytical solution cannot be determined, for example, when logistic combination functions are used. In such cases equilibria can only be determined either by numerically solving the equations by some numerical approximation method, or by observing the behaviour of the model in simulation experiments. But in the latter case verification is not possible, as then only simulation results are available. An additional drawback is that in such cases specific values for the parameters of the model have to be chosen, whereas in the case of an explicit analytical solution a more generic expression can be obtained which depends, as a function, on the parameter values. For example, for the cases described in Sects. 3–6 expressions can be found for the equilibrium values in terms of the connection weights (for which no specific values are needed at forehand).

### 3 Mathematical analysis for equilibrium states: an example

Are there cases in which the types of behaviour considered above can be predicted without running a simulation? In particular, can equilibrium values be predicted, and how they depend on the specific values of the parameters of the model

(e.g. connection weights, speed factors)? Below, these questions will be answered for a relatively simple example. Indeed it will turn out that in this case it is possible to predict the equilibrium values from the connection weights (the equilibrium values turn out to be independent of the speed factors, as long as these are nonzero). As a first step, consider the sensor state  $ss_s$ .

#### LP<sub>ss<sub>s</sub></sub> Sensing a stimulus: determining values for state $ss_s$

$$dss_s(t)/dt = \eta_{ss_s} [\omega_{sensing} ws_s(t) - ss_s(t)]$$

Having an equilibrium value means that no change occurs at  $t$ :  $dss_s(t)/dt = 0$ . As it is assumed that  $\eta_{ss_s}$  is nonzero, this is equivalent to the following equilibrium equation for state  $ss_s$ , with  $\underline{ws}_s$  and  $\underline{ss}_s$  the equilibrium values for the two states  $ws_s$  and  $ss_s$ .

$$\omega_{sensing} \underline{ws}_s = \underline{ss}_s$$

In a similar manner this can be done for the other states, resulting in the following equations:

Equilibrium of state	Equilibrium criterion
$ss_s$	$\omega_{sensing} \underline{ws}_s = \underline{ss}_s$
$srs_s$	$\omega_{representing} \underline{ss}_s = \underline{srs}_s$
$ps_a$	$\omega_{responding} \underline{srs}_s + \omega_{amplifying} \underline{srs}_e = \underline{ps}_a$
$srs_e$	$\omega_{predicting} \underline{ps}_a = \underline{srs}_e$
$es_a$	$\omega_{executing} \underline{ps}_a = \underline{es}_a$

These are five equations with six unknowns  $\underline{ws}_s, \underline{ss}_s, \underline{srs}_s, \underline{ps}_a, \underline{srs}_e, \underline{es}_a$ ; however, the variable  $\underline{ws}_s$  can be considered given as it indicates the external stimulus. So the five equations can be used to find expressions for the equilibrium values for the five other states in terms of the connection weights and  $\underline{ws}_s$ . Note that for the sake of simplicity here it is assumed that  $\omega_{amplifying}$  and  $\omega_{predicting}$  are not both 1. Then this can be solved in an explicit analytical manner as follows. First two of them (the first two equations) are expressed in the externally given value  $\underline{ws}_s$ :

$$\underline{ss}_s = \omega_{sensing} \underline{ws}_s$$

$$\underline{srs}_s = \omega_{representing} \underline{ss}_s = \omega_{representing} \omega_{sensing} \underline{ws}_s$$

Moreover, the third and fourth equation can be solved as follows:

$$\omega_{responding} \underline{srs}_s + \omega_{amplifying} \underline{srs}_e = \underline{ps}_a$$

$$\omega_{predicting} \underline{ps}_a = \underline{srs}_e$$

Substitute  $\omega_{predicting} \underline{ps}_a$  for  $\underline{srs}_e$  in the third equation, resulting in the following equation in  $\underline{ps}_a$  and  $\underline{srs}_s$ :

$$\omega_{responding} \underline{srs}_s + \omega_{amplifying} \omega_{predicting} \underline{ps}_a = \underline{ps}_a$$



This can be used to express  $\underline{ps}_a$  in  $\underline{srs}_s$ , and subsequently in  $\underline{ws}_s$ :

$$\omega_{\text{responding}} \underline{srs}_s = (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \underline{ps}_a$$

$$\begin{aligned} \underline{ps}_a &= \omega_{\text{responding}} \underline{srs}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ &= \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \end{aligned}$$

Moreover, by the fourth equation it is found

$$\underline{srs}_e = \omega_{\text{predicting}} \underline{ps}_a = \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})$$

Based on these, the fifth equation can be used to get an expression for  $\underline{es}_a$ :

$$\underline{es}_a = \omega_{\text{executing}} \underline{ps}_a = \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}})$$

Summarizing, all equilibrium values have been expressed in terms of the external state  $\underline{ws}_s$  and the connection weights:

$$\begin{aligned} \underline{ss}_s &= \omega_{\text{sensing}} \underline{ws}_s \\ \underline{srs}_s &= \omega_{\text{representing}} \omega_{\text{sensing}} \underline{ws}_s \\ \underline{ps}_a &= \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{srs}_e &= \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{es}_a &= \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} \underline{ws}_s / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \end{aligned}$$

For example, if the external stimulus  $\underline{ws}_s$  has level 1 this becomes:

$$\begin{aligned} \underline{ss}_s &= \omega_{\text{sensing}} \\ \underline{srs}_s &= \omega_{\text{representing}} \omega_{\text{sensing}} \\ \underline{ps}_a &= \omega_{\text{responding}} \omega_{\text{representing}} \omega_{\text{sensing}} / \\ &\quad (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{srs}_e &= \omega_{\text{predicting}} \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \\ \underline{es}_a &= \omega_{\text{executing}} \omega_{\text{responding}} \omega_{\text{representing}} \\ &\quad \times \omega_{\text{sensing}} / (1 - \omega_{\text{amplifying}} \omega_{\text{predicting}}) \end{aligned}$$

Moreover, if all connection weights are 1, except that  $\omega_{\text{responding}} = 0.5$  and  $\omega_{\text{amplifying}} = 0.5$ , as in the example simulation shown in [17], Section 2.4.1, the values become:

$$\underline{ss}_s = 1$$

$$\begin{aligned} \underline{srs}_s &= 1 \\ \underline{ps}_a &= 0.5/0.5 = 1 \\ \underline{srs}_e &= 0.5/0.5 = 1 \\ \underline{es}_a &= 0.5/0.5 = 1 \end{aligned}$$

Indeed in the example simulation in [17], Section 2.4.1 Fig. 11 it can be seen that all values go to 1. The solution of the equilibrium equations in terms of the connection weights can be used to predict that when the connection weights have different values, also these equilibrium values will turn out different. Recall that the cases  $\omega_{\text{amplifying}} = 1$  and  $\omega_{\text{predicting}} = 1$  was excluded. In that case the combined third and fourth equation becomes trivial, as  $\underline{ps}_a$  is lost from the equation:

$$\begin{aligned} \omega_{\text{responding}} \underline{srs}_s + \omega_{\text{amplifying}} \omega_{\text{predicting}} \underline{ps}_a &= \underline{ps}_a \\ \omega_{\text{responding}} \underline{srs}_s + \underline{ps}_a &= \underline{ps}_a \\ \omega_{\text{responding}} \underline{srs}_s &= 0 \\ \underline{srs}_s &= 0 \end{aligned}$$

Here in the last step it is assumed that  $\omega_{\text{responding}} > 0$ . As a consequence by the first two equations also  $\underline{ss}_s$  and  $\underline{ws}_s$  are 0, and by the fourth and fifth equation also the values for the other states. It turns out that in this case there can only be an equilibrium if there is no stimulus at all. As soon as there is a nonzero stimulus in this case that  $\omega_{\text{amplifying}} = 1$  and  $\omega_{\text{predicting}} = 1$ , the values of  $\underline{ps}_a$ ,  $\underline{srs}_e$  and  $\underline{es}_a$  increase indefinitely to larger and larger values (and in particular do not stay within the interval  $[0, 1]$ ), as can be seen from simulations. Note that there was an additional assumption made that  $\omega_{\text{responding}} > 0$ . If, in contrast,  $\omega_{\text{responding}} = 0$ , then still more possibilities for equilibria are available. For example, in that case  $\underline{ps}_a$  and  $\underline{srs}_e$  can have any value, but they have to be equal due to the fourth equation, but this value is independent of the values of  $\underline{ws}_s$ ,  $\underline{ss}_s$  and  $\underline{srs}_s$ , as there is no nonzero connection between these parts of the graph. So, this would not be a very relevant case.

The analysis above can also be done to find out whether or not the activation level of a state is increasing. As a first step, again consider the sensor state  $\underline{ss}_s$ .

#### LP<sub>ss<sub>s</sub></sub> Sensing a stimulus: determining values for state $\underline{ss}_s$

$$\begin{aligned} d\underline{ss}_s(t)/dt &= \eta_{\underline{ss}_s} [\omega_{\text{sensing}} \underline{ws}_s(t) - \underline{ss}_s(t)] \\ \underline{ss}_s(t + \Delta t) &= \underline{ss}_s(t) + \eta_{\underline{ss}_s} [\omega_{\text{sensing}} \underline{ws}_s(t) - \underline{ss}_s(t)] \Delta t \end{aligned}$$

The activation value increases mean

$$d\underline{ss}_s(t)/dt > 0 \quad \text{or} \quad \underline{ss}_s(t + \Delta t) > \underline{ss}_s(t)$$

This is equivalent to:

$$\omega_{\text{sensing}} \text{WS}_S(t) - \text{SS}_S(t) > 0$$

This in turn is equivalent to the criterion that the impact on  $\text{ss}_S$  is higher than the current activation value:

$$\omega_{\text{sensing}} \text{WS}_S(t) > \text{SS}_S(t)$$

For example, when  $\text{ws}_S(t) = 1$  and  $\omega_{\text{sensing}} = 1$ , then the criterion  $\omega_{\text{sensing}} \text{WS}_S(t) > \text{SS}_S(t)$  indicates the activation of state  $\text{ss}_S$  will increase as long as it did not reach the value 1 yet. This gives as additional information that the equilibrium value 1 of sensor state  $\text{ss}_S$  is *attracting*: the value goes in that direction as long as it was not reached.

In a similar manner this can be done for the other states, thus obtaining the following criteria:

**State is increasing if and only if**

$\text{SS}_S$	$\omega_{\text{sensing}} \text{WS}_S(t) > \text{SS}_S(t)$
$\text{SRS}_S$	$\omega_{\text{representing}} \text{SS}_S(t) > \text{SRS}_S(t)$
$\text{PS}_a$	$\omega_{\text{responding}} \text{SRS}_S(t) + \omega_{\text{amplifying}} \text{SRS}_e(t) > \text{PS}_a(t)$
$\text{SRS}_e$	$\omega_{\text{predicting}} \text{PS}_a(t) > \text{SRS}_e(t)$
$\text{ES}_a$	$\omega_{\text{executing}} \text{PS}_a(t) > \text{ES}_a(t)$

**4 Mathematical analysis for equilibrium states: scaled sum combination function**

The approach described above can be applied easily for the case of a scaled sum combination function  $c_i(\dots)$  for each state  $X_i$ ; such a scaled sum function  $\text{ssum}_{\lambda_i}(\dots)$  with scaling factor  $\lambda_i$  is defined as

$$\text{ssum}_{\lambda_i}(V_1, \dots, V_k) = (V_1 + \dots + V_k)/\lambda_i$$

Suppose the differential equation for some state  $X_i$  connected to states  $X_j$  is given by

$$dX_i/dt = \eta_i[\text{aggimpact}_i(X_1, \dots, X_k) - X_i]$$

where

$$\begin{aligned} \text{aggimpact}_i(X_1, \dots, X_k) &= c_i(\omega_{1,i} X_1, \dots, \omega_{k,i} X_k) \\ &= \text{ssum}_{\lambda_i}(\omega_{1,i} X_1, \dots, \omega_{k,i} X_k) \\ &= (\omega_{1,i} X_1 + \dots + \omega_{k,i} X_k)/\lambda_i \end{aligned}$$

with  $\omega_{j,i}$  the specific weights for the connections from  $X_j$  to  $X_i$ . In this case the following holds:

$$\begin{aligned} \text{Increasing } X_i : X_i(t + \Delta t) &> X_i(t) \\ \Leftrightarrow (\omega_{1,i} X_1(t) + \dots + \omega_{k,i} X_k(t))/\lambda_i &> X_i(t) \end{aligned}$$

$$\begin{aligned} \text{Equilibrium of } X_i : X_i(t + \Delta t) &= X_i(t) \\ \Leftrightarrow (\omega_{1,i} X_1(t) + \dots + \omega_{k,i} X_k(t))/\lambda_i &= X_i(t) \end{aligned}$$

$$\begin{aligned} \text{Decreasing } X_i : X_i(t + \Delta t) &< X_i(t) \\ \Leftrightarrow (\omega_{1,i} X_1(t) + \dots + \omega_{k,i} X_k(t))/\lambda_i &< X_i(t) \end{aligned}$$

In particular, the equilibrium equations for the states  $X_i$  are

$$\begin{aligned} (\omega_{1,1} \underline{X}_1 + \dots + \omega_{k,1} \underline{X}_k)/\lambda_1 &= \underline{X}_1 \\ \dots & \\ (\omega_{1,k} \underline{X}_1 + \dots + \omega_{k,k} \underline{X}_k)/\lambda_k &= \underline{X}_k \end{aligned}$$

This means that in an equilibrium state the value  $\underline{X}_i$  for a state  $X_i$  may be a weighted average of the equilibrium values  $\underline{X}_j$  for the states  $X_j$ , in particular when

$$\lambda_i = \omega_{1,i} + \dots + \omega_{k,i}$$

Note that always at least one solution exists: when all are 0. But it is usually more interesting to know whether nonzero solutions exist.

The equilibrium equations are equivalent to

$$\begin{aligned} \omega_{1,1} \underline{X}_1 + \dots + \omega_{k,1} \underline{X}_k &= \lambda_1 \underline{X}_1 \\ \dots & \\ \omega_{1,k} \underline{X}_1 + \dots + \omega_{k,k} \underline{X}_k &= \lambda_k \underline{X}_k \end{aligned}$$

or

$$\begin{aligned} (\omega_{1,1} - \lambda_1) \underline{X}_1 + \omega_{2,1} \underline{X}_2 + \dots + \omega_{k,1} \underline{X}_k &= 0 \\ \dots & \\ \omega_{1,i} \underline{X}_1 + \dots + \omega_{i-1,i} \underline{X}_{i-1} + (\omega_{i,i} - \lambda_i) \underline{X}_i & \\ + \omega_{i+1,i} \underline{X}_{i+1} + \dots + \omega_{k,i} \underline{X}_k &= 0 \\ \dots & \\ \omega_{1,k} \underline{X}_1 + \dots + \omega_{k-1,k} \underline{X}_{k-1} + (\omega_{k,k} - \lambda_k) \underline{X}_k &= 0 \end{aligned}$$

In general these linear equilibrium equations can be solved analytically, which in principle can provide symbolic expressions for the equilibrium values of  $X_j$  in terms of the connection weights  $\omega_{j,i}$  and the scaling factor  $\lambda_i$ . However, for more than two states ( $k > 2$ ) such expressions may tend to become more and more complex, but this depends on the number of these  $\omega_{j,i}$  which are nonzero, i.e. how many connections between the states exist. For example, if all states have only one incoming and one outgoing connection (a cascade or loop), then these equations can easily be solved. In some cases no nonzero solution exists. This happens, for example, when the values of the parameters are such that two of the equations in a sense contradict each other, as in the equations  $X_1 - 2X_2 = 0$  and  $X_1 - 3X_2 = 0$ .

In some cases some properties of equilibrium values can be derived. For well-connected temporal-causal network models based on scaled sum functions with as scaling factor the sum of the weights of the incoming connections it can be derived that all states have the same equilibrium value.

**Definition 1** A network is called *strongly connected* if for every two nodes  $A$  and  $B$  there is a directed path from  $A$  to  $B$  and vice versa.

**Lemma 1** Let a temporal-causal network model be given based on scaled sum functions:

$$dY/dt = \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y]$$

Then the following hold.

- (a) If for some state  $Y$  at time  $t$  for all states  $X$  connected toward  $Y$  it holds  $X(t) \geq Y(t)$ , then  $Y(t)$  is increasing at  $t$  :  $dY(t)/dt \geq 0$ ; if for all states  $X$  connected toward  $Y$  it holds  $X(t) \leq Y(t)$ , then  $Y(t)$  is decreasing at  $t$  :  $dY(t)/dt \leq 0$ .
- (b) If for some state  $Y$  at time  $t$  for all states  $X$  connected toward  $Y$  it holds  $X(t) \geq Y(t)$ , and at least one state  $X$  connected toward  $Y$  exists with  $X(t) > Y(t)$  then  $Y(t)$  is strictly increasing at  $t$ :  $dY(t)/dt > 0$ . If for some state  $Y$  at time  $t$  for all states  $X$  connected toward  $Y$  it holds  $X(t) \leq Y(t)$ , and at least one state  $X$  connected toward  $Y$  exists with  $X(t) < Y(t)$  then  $Y(t)$  is strictly decreasing at  $t$  :  $dY(t)/dt < 0$ .

*Proof of Lemma 1* (a) From the differential equation for  $Y(t)$

$$\begin{aligned} dY/dt &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y] \\ &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X \\ &\quad - \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} Y] / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} \\ &= \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} (X - Y)] / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} \end{aligned}$$

it follows that  $dY(t)/dt \geq 0$ , so  $Y(t)$  is increasing at  $t$ . Similar for decreasing.

- (b) In this case it follows that  $dY(t)/dt > 0$ , so  $Y(t)$  is strictly increasing. Similar for decreasing.  $\square$

**Theorem 1** (convergence to one value) Let a strongly connected temporal-causal network model be given based on scaled sum functions:

$$dY/dt = \eta_Y [\sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} X / \sum_{X, \omega_{X,Y} > 0} \omega_{X,Y} - Y]$$

Then for all states  $X$  and  $Y$  the equilibrium values  $\underline{X}$  and  $\underline{Y}$  are equal:  $\underline{X} = \underline{Y}$ . Moreover, this equilibrium state is attracting.

*Proof of Theorem 1* Take a state  $Y$  with highest value  $\underline{Y}$ . Then for all states  $X$  it holds  $\underline{X} \leq \underline{Y}$ . Suppose for some state  $X$  connected toward  $Y$  it holds  $\underline{X} < \underline{Y}$ . Take a time point  $t$  and assume  $Z(t) = \underline{Z}$  for all states  $Z$ . Now apply Lemma 1b) to state  $Y$ . It follows that  $dY(t)/dt < 0$ , so  $Y(t)$  is not in equilibrium for this value  $\underline{Y}$ . This contradicts that this  $\underline{Y}$  is an equilibrium value for state  $Y$ . Therefore, the assumption that for some state  $X$  connected toward  $Y$  it holds  $\underline{X} < \underline{Y}$  cannot be true. This shows that  $\underline{X} = \underline{Y}$  for all states connected toward  $Y$ . Now this argument can be repeated for all states connected toward  $Y$  instead of  $X$ . By iteration every other state in the network is reached, due to the strong connectivity assumption; it follows that all other states  $X$  in the temporal-causal network model have the same equilibrium value  $\underline{X}$  as  $\underline{Y}$ . From Lemma 1b) it follows that such an equilibrium state is attracting: if for any state the value is deviating it will move to the equilibrium value.  $\square$

## 5 Mathematical analysis for equilibrium states: Hebbian learning

It can also be analysed from the difference or differential equation when a Hebbian adaptation process (e.g. [2, 4–6, 8, 15, 16]) has an equilibrium and when it increases or decreases. More specifically, assume the following dynamic model (also see [5]) for Hebbian learning for the strength  $\omega$  of a connection from a state  $X_1$  to a state  $X_2$  with maximal connection strength 1, learning rate  $\eta > 0$ , and extinction rate  $\zeta \geq 0$  (here  $X_1(t)$  and  $X_2(t)$  denote the activation levels of the states  $X_1$  and  $X_2$  at time  $t$ ; sometimes the  $t$  is left out of  $X_i(t)$  and simply  $X_i$  is written)

$$\begin{aligned} \omega(t + \Delta t) &= \omega(t) + [\eta X_1(t) X_2(t) (1 - \omega(t)) - \zeta \omega(t)] \Delta t \\ d\omega(t)/dt &= \eta X_1 X_2 (1 - \omega(t)) - \zeta \omega(t) \end{aligned}$$

Note that also for the states  $X_1$  and  $X_2$  equations may be given, but here the focus is on  $\omega$ .

From the expressions for  $\omega$  it can be analysed when each of the following cases occurs:

Increasing  $\omega$  :  $d\omega(t)/dt > 0$

$$\Leftrightarrow \eta X_1 X_2 (1 - \omega(t)) - \zeta \omega(t) > 0$$

Equilibrium of  $\omega$  :  $d\omega(t)/dt = 0$

$$\Leftrightarrow \eta X_1 X_2 (1 - \omega(t)) - \zeta \omega(t) = 0$$

Decreasing  $\omega$  :  $d\omega(t)/dt < 0$

$$\Leftrightarrow \eta X_1 X_2 (1 - \omega(t)) - \zeta \omega(t) < 0$$



### 5.1 Analysis of increase, decrease or equilibrium for Hebbian learning without extinction

To keep things a bit simple for a first analysis, for the special case that there is no extinction ( $\zeta = 0$ ), this easily leads to the following criteria

Increasing  $\omega$  :  $\eta X_1 X_2(1 - \omega(t)) > 0$   
 $\Leftrightarrow \omega(t) < 1$  and both  $X_1 > 0$  and  $X_2 > 0$

Equilibrium of  $\omega$  :  $\eta X_1 X_2(1 - \omega(t)) = 0$   
 $\Leftrightarrow \omega(t) = 1$  or  $X_1 = 0$  or  $X_2 = 0$

Decreasing  $\omega$  :  $\eta X_1 X_2(1 - \omega(t)) < 0$   
 this is never the case, as always  $X_i \geq 0$   
 and  $\omega(t) \leq 1$

So, in case that there is no extinction, the only equilibrium is when  $\omega = 1$ , and as long as this value was not reached yet and both  $X_1 > 0$  and  $X_2 > 0$ , the value of  $\omega$  increases: the equilibrium is attracting. Note that when  $X_1 = 0$  or  $X_2 = 0$ , also an equilibrium for  $\omega$  can be found: no (further) learning takes place; the value of  $\omega$  stays the same independent of which value it has, so in this case any value is an equilibrium value. In simulations this indeed can be observed: as long as both  $X_1 > 0$  and  $X_2 > 0$  the value of  $\omega$  keeps on increasing until it reaches 1, but if  $X_1 = 0$  or  $X_2 = 0$  then  $\omega$  always stays the same.

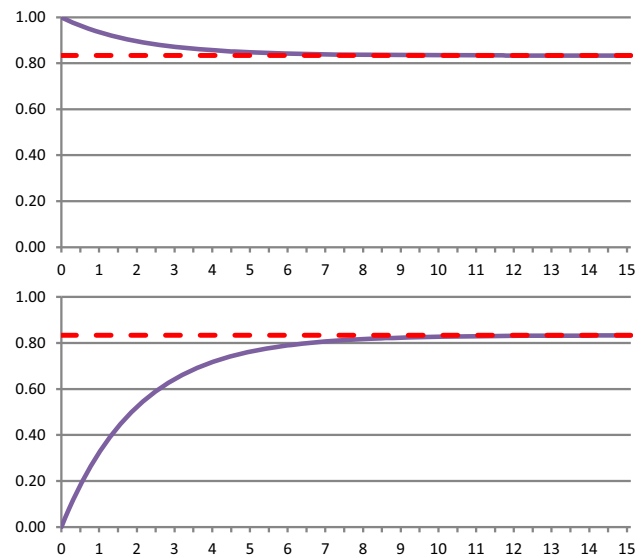
### 5.2 Analysis of increase, decrease or equilibrium for Hebbian learning with extinction

As a next step this analysis is extended to the case with extinction  $\zeta > 0$ . In this case the analysis requires slightly more work; here for convenience the  $t$  is left out of the expressions.

**Increasing  $\omega$**  :  $\eta X_1 X_2(1 - \omega) - \zeta \omega > 0$   
 $\Leftrightarrow \eta X_1 X_2 - \eta X_1 X_2 \omega - \zeta \omega > 0$   
 $\Leftrightarrow \eta X_1 X_2 - (\zeta + \eta X_1 X_2) \omega > 0$   
 $\Leftrightarrow (\zeta + \eta X_1 X_2) \omega < \eta X_1 X_2$   
 $\Leftrightarrow \omega < \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2}$   
 $\Leftrightarrow \omega < \frac{1}{1 + \zeta / (\eta X_1 X_2)}$   
 (when both  $X_1 > 0$  and  $X_2 > 0$ )

Note that when  $X_1 = 0$  or  $X_2 = 0$ , the value of  $\omega$  is never increasing. Similarly the following criteria can be found.

**Equilibrium of  $\omega$**  :  $\eta X_1 X_2(1 - \omega) - \zeta \omega = 0$   
 $\Leftrightarrow \omega = \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2}$



**Fig. 3** Hebbian learning for  $\eta = 0.4$ ,  $\zeta = 0.08$ ,  $\Delta t = 0.1$ , and activation levels  $X_1 = 1$  and  $X_2 = 1$  Equilibrium value 0.83 (dotted line)

$$\Leftrightarrow \omega = \frac{1}{1 + \zeta / (\eta X_1 X_2)}$$

(when both  $X_1 > 0$  and  $X_2 > 0$ )

$$\eta X_1 X_2(1 - \omega) - \zeta \omega = 0$$

$$\Leftrightarrow \omega = 0$$

(when  $X_1 = 0$  or  $X_2 = 0$ , and  $\zeta > 0$ )

**Decreasing  $\omega$**  :  $\eta X_1 X_2(1 - \omega) - \zeta \omega < 0$

$$\Leftrightarrow \omega > \frac{\eta X_1 X_2}{\zeta + \eta X_1 X_2}$$

$$\Leftrightarrow \omega > \frac{1}{1 + \zeta / (\eta X_1 X_2)}$$

(when both  $X_1 > 0$  and  $X_2 > 0$ )

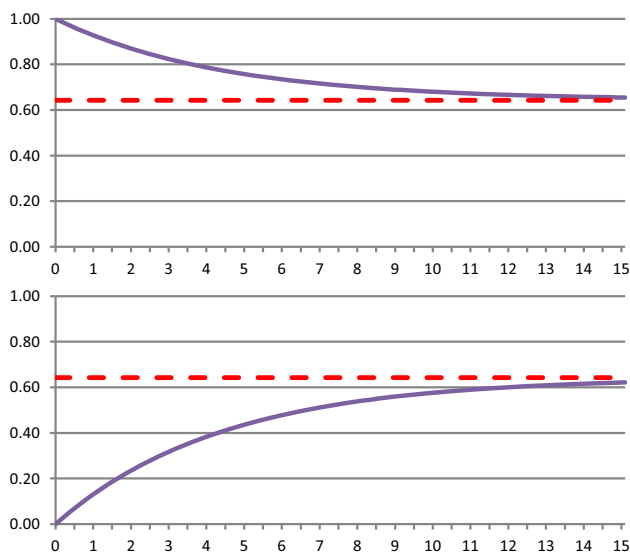
$$\eta X_1 X_2(1 - \omega) - \zeta \omega < 0$$

$$\Leftrightarrow \text{always}$$

(when  $X_1 = 0$  or  $X_2 = 0$ , and  $\zeta > 0, \omega > 0$ )

In this more general case with extinction, depending on the values of  $X_1$  and  $X_2$  there may be a positive equilibrium value (when both  $X_1 > 0$  and  $X_2 > 0$ ) but when  $\zeta > 0$  this value is  $< 1$ . Also 0 is an equilibrium value (when  $X_1 = 0$  or  $X_2 = 0$ ). This looks similar to the case without extinction. Moreover, as before, the value of  $\omega$  increases when it is under the positive equilibrium value and it decreases when it is above this value (it is an attracting equilibrium); for example patterns, see Figs. 3 and 4.

Note that this time this positive equilibrium value (indicated by the dotted line) is lower than 1. It may be close to



**Fig. 4** Hebbian learning for  $\eta = 0.4$ ,  $\zeta = 0.08$ ,  $\Delta t = 0.1$ , and activation levels  $X_1 = 0.6$  and  $X_2 = 0.6$  Equilibrium value 0.64 (dotted line)

1, but when  $\zeta > 0$  it never will be equal to 1. In fact the maximal value of this equilibrium is when both  $X_1 = 1$  and  $X_2 = 1$ , in which case the equilibrium value is

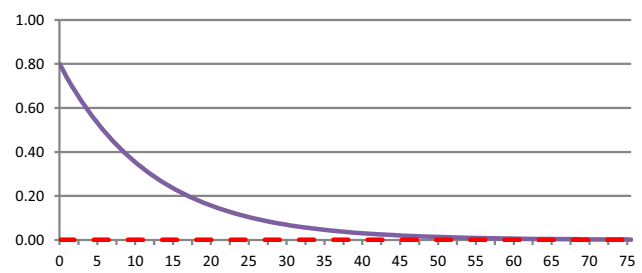
$$\frac{1}{1 + \zeta/\eta}$$

For example, for  $\eta = 0.4$ ,  $\zeta = 0.02$ , and  $X_1 = 1$  and  $X_2 = 1$ , the positive equilibrium value for  $\omega$  is about 0.95. Another example is  $\eta = 0.4$ ,  $\zeta = 0.08$ , and  $X_1 = 1$  and  $X_2 = 1$ , in which case the equilibrium value is 0.83. The graphs in Fig. 2 show what happens below this equilibrium and above it. If for the same settings for  $\eta$  and  $\zeta$ , the activation levels are lower ( $X_1 = 0.6$  and  $X_2 = 0.6$ ), then the equilibrium value is lower too (0.64), and the learning is much slower, as is shown in Fig. 3.

So, it is found that the positive equilibrium value occurs for  $X_1 > 0$  and  $X_2 > 0$ , and in that case this equilibrium is attracting. In contrast, the equilibrium value 0 does not occur for  $X_1 > 0$  and  $X_2 > 0$ , but it does occur for  $X_1 = 0$  or  $X_2 = 0$ , in which case no positive equilibrium value occurs. In this case pure extinction occurs:  $\omega$  is attracted by the equilibrium value 0; this pattern is different from the case without extinction. For an example of such a pure extinction process, see Fig. 5. Note that, given the lower value of the extinction rate  $\zeta$ , the extinction process takes a much longer time than the learning process.

### 5.3 How much activation of $X_1$ and $X_2$ is needed to let $\omega$ increase?

From a different angle, another question that can be addressed is for a given value of  $\omega$ , how high the value  $X_1 X_2$  should be



**Fig. 5** Pure extinction for  $\eta = 0.4$ ,  $\zeta = 0.08$ ,  $\Delta t = 0.1$ , and activation levels  $X_1 = X_2 = 0$ ; equilibrium value 0

to let  $\omega$  become higher. This can be determined in a similar manner as follows:

**Increasing  $\omega$  :**

$$\omega < \frac{1}{1 + \zeta/(\eta X_1 X_2)}$$

$$\Leftrightarrow (1 + \zeta/\eta X_1 X_2) \omega < 1$$

$$\Leftrightarrow 1 + \zeta/(\eta X_1 X_2) < 1/\omega$$

$$\Leftrightarrow \zeta/(\eta X_1 X_2) < 1/\omega - 1 = (1 - \omega)/\omega$$

$$\Leftrightarrow 1/(X_1 X_2) < \frac{\eta}{\zeta} (1 - \omega)/\omega$$

$$\Leftrightarrow X_1 X_2 > \frac{\zeta}{\eta} \omega/(1 - \omega)$$

So, for activation levels  $X_1$  and  $X_2$  with  $X_1 X_2 > \frac{\zeta}{\eta} \omega/(1 - \omega)$ , further learning takes place, and below this value extinction dominates and will decrease the level of  $\omega$ .

## 6 Mathematical analysis for equilibrium states: dynamic network connections

The connections between agents in a social network may change over time based on the *homophily principle*: the closer the states of the interacting agents, the stronger the connections of the agents will become. This principle may be formalized with as a general template

$$d\omega_{A,B}/dt = \eta_{A,B} [c_{A,B}(X_A, X_B, \omega_{A,B}) - \omega_{A,B}]$$

for some combination function  $c_{A,B}(V_1, V_2, W)$  for which it is assumed that  $c_{A,B}(V_1, V_2, 0) \geq 0$  and  $c_{A,B}(V_1, V_2, 1) \leq 1$ . The example used in this section is

$$c_{A,B}(V_1, V_2, W) = W + (\tau_{A,B}^2 - (V_1 - V_2)^2)W(1 - W)$$

In this case

$$d\omega_{A,B}/dt = \eta_{A,B} (\tau_{A,B}^2 - (X_A - X_B)^2) \omega_{A,B} (1 - \omega_{A,B})$$

In this section it is analysed which equilibrium values  $\underline{\omega}_{A,B}$  can occur for  $\omega_{A,B}(t)$  and when  $\omega_{A,B}(t)$  is increasing or decreasing.

The standard approach is to derive an inequality or equation from the differential equation by putting  $d\omega_{A,B}(t)/dt = 0$ ,  $d\omega_{A,B}(t)/dt \geq 0$  or  $d\omega_{A,B}(t)/dt \leq 0$ . For this case this provides

$$\begin{aligned} \text{Increasing } \omega_{A,B} \quad d\omega_{A,B}(t)/dt &\geq 0 \\ \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2) & \\ \times \omega_{A,B}(1 - \omega_{A,B}) &> 0 \end{aligned}$$

$$\begin{aligned} \text{Equilibrium of } \omega_{A,B} \quad d\omega_{A,B}(t)/dt &= 0 \\ \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2) & \\ \times \omega_{A,B}(1 - \omega_{A,B}) &= 0 \end{aligned}$$

$$\begin{aligned} \text{Decreasing } \omega_{A,B} \quad d\omega_{A,B}(t)/dt &\leq 0 \\ \Leftrightarrow \eta_{A,B}(\tau_{A,B}^2 - (X_A - X_B)^2) & \\ \times \omega_{A,B}(1 - \omega_{A,B}) &< 0 \end{aligned}$$

For  $\omega_{A,B} = 0$  or  $\omega_{A,B} = 1$  the middle condition is fulfilled. This means that  $\underline{\omega}_{A,B} = 0$  and  $\underline{\omega}_{A,B} = 1$  are equilibrium values. Now assume  $0 < \omega_{A,B} < 1$ . Then  $\omega_{A,B}(1 - \omega_{A,B}) > 0$ , and therefore this factor can be left out, and the same applies to  $\eta_{A,B} > 0$ ; this results in:

$$\begin{aligned} \text{Increasing } \omega_{A,B} \quad \tau_{A,B}^2 - (X_A - X_B)^2 &> 0 \\ \Leftrightarrow |X_A - X_B| &< \tau_{A,B} \end{aligned}$$

$$\begin{aligned} \text{Equilibrium of } \omega_{A,B} \quad \tau_{A,B}^2 - (X_A - X_B)^2 &= 0 \\ \Leftrightarrow |X_A - X_B| &= \tau_{A,B} \end{aligned}$$

$$\begin{aligned} \text{Decreasing } \omega_{A,B} \quad \tau_{A,B}^2 - (X_A - X_B)^2 &< 0 \\ \Leftrightarrow |X_A - X_B| &> \tau_{A,B} \end{aligned}$$

This shows that for cases that  $|X_A - X_B| < \tau_{A,B}$  the connection keeps on becoming stronger until  $\omega_{A,B}$  becomes in equilibrium at 1. Similarly for cases that  $|X_A - X_B| > \tau_{A,B}$  the connection keeps on becoming weaker until  $\omega_{A,B}$  comes in equilibrium at 0. This implies that the equilibria  $\underline{\omega}_{A,B} = 0$  and  $\underline{\omega}_{A,B} = 1$  can both become attracting, but under different circumstances concerning the values of  $X_A$  and  $X_B$ .

In exceptional situations it could be the case that  $|X_A - X_B| = \tau_{A,B}$  in which case  $\omega_{A,B}$  is also in equilibrium, with  $\omega_{A,B}$  having any value. So in principle the equilibrium equation has three solutions

$$\underline{\omega}_{A,B} = 0 \quad \text{or} \quad \underline{\omega}_{A,B} = 1 \quad \text{or} \\ |\underline{X}_A - \underline{X}_B| = \tau_{A,B} \quad \text{and} \quad \underline{\omega}_{A,B} \text{ has any value}$$

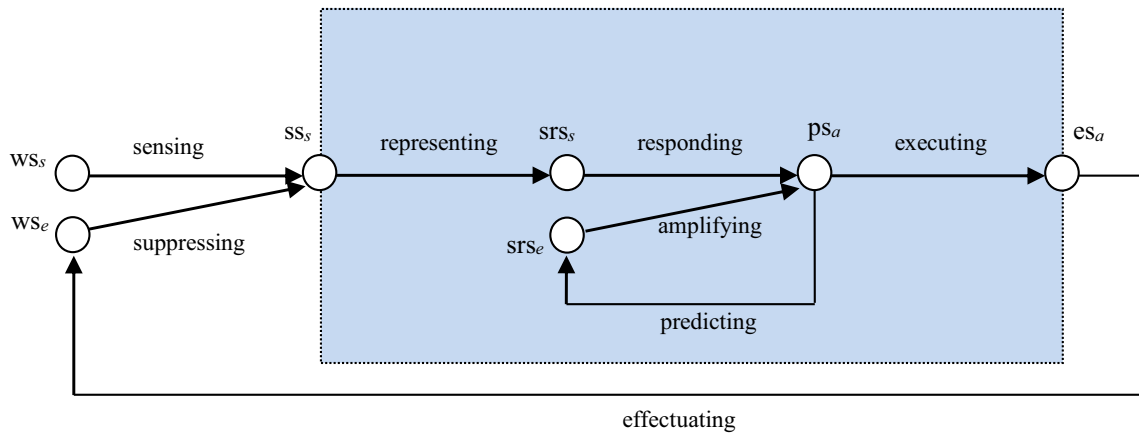
The analysis above can also be done for similar but slightly more complex variants of the model, of which the quadratic variant is described in [14]:

$$\begin{aligned} c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B} - |V_1 - V_2|))(1 - W) \\ &\quad - \text{Pos}(-\eta_{A,B}(\tau_{A,B} - |V_1 - V_2|))W \\ c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(\tau_{A,B}^2 - (V_1 - V_2)^2))(1 - W) \\ &\quad - \text{Pos}(-\eta_{A,B}(\tau_{A,B}^2 - (V_1 - V_2)^2))W \\ c_{A,B}(V_1, V_2, W) &= W + \text{Pos}(\eta_{A,B}(0.5 - 1/(1 \\ &\quad + e^{-\sigma_{A,B}(|V_1 - V_2| - \tau_{A,B})}))) (1 - W) \\ &\quad - \text{Pos}(-\eta_{A,B}(0.5 - 1/(1 + e^{-\sigma_{A,B}(|V_1 - V_2| - \tau_{A,B})})))W \end{aligned}$$

where  $\text{Pos}(x) = (|x| + x)/2$ , which returns  $x$  when  $x$  is positive and 0 when  $x$  is negative. These models make that the approaching of the boundaries 0 and 1 of the interval  $[0, 1]$  of  $\omega$  is slow, thus making  $\omega$  not cross these boundaries, but  $\omega$  departing from the neighbourhood of these boundaries is not slow. In [14] an analysis and example simulations can be found using the second, quadratic model. As part of the analysis, there it is also shown that different equilibrium values  $\underline{X}_A$  and  $\underline{X}_B$  have a distance of at least  $\tau_{A,B}$ , which implies that at most  $1/\tau_{A,B}$  clusters can emerge.

### 7 Mathematical analysis for behaviour ending up in a limit cycle pattern

Sometimes the values of the states of a model do not end up in an equilibrium value, but instead keep on fluctuating all the time, and after some time they do this according to a repeated pattern, called a *limit cycle*; for example, see [3,9,11–13]. The example model shown in Figs. 1 and 2 can be extended to show such behaviour; see Fig. 6. In this case it is assumed that action  $a$  directs the person (e.g. his or her gaze) away from the stimulus  $s$ , so that after (full) execution of  $a$  stimulus  $s$  is not sensed anymore. This type of behaviour can occur as a form of emotion regulation to down-regulate a stressful emotion triggered by  $s$ . The effect of this is as follows. The presence of stimulus  $s$  leads to high activation levels of sensor state and sensory representation for  $s$ , and subsequently for the preparation state and execution state of action  $a$ . But then the action leads to its effect in the world which is suppression of the sensor state for  $s$ . As a consequence the sensor state and sensory representation for  $s$ , and also the preparation state and execution state of action  $a$  get low activation levels. The effect is that there is no suppression of sensing the stimulus anymore and, therefore, all activation levels become high again. And so it goes on and on, forever (see also Fig. 8). At a longer timescale this type of pattern may also occur in so-called on-again-off-again relationships. This type of behaviour can be achieved by the following additions to the example model (see Fig. 6):



**Fig. 6** Simple example model incorporating suppression of sensing

- a connection from the execution state  $es_a$  of  $a$  to the world state  $ws_e$  for effect  $e$  of action  $a$
- a connection from this world state  $ws_e$  for  $e$  to the sensor state  $ss_s$  of  $s$
- a combination function for the sensor state  $ss_s$  of  $s$  that models that  $ws_e$  makes that  $s$  is not sensed

The aggregation used for  $ss_s$  is modelled by the following combination function  $c_{ss_s}(V_1, V_2)$ , where  $V_1$  refers to the impact  $\omega_{ws_s,ss_s} ws_s(t)$  from  $ws_s$  on  $ss_s$  and  $V_2$  to the impact  $\omega_{ws_e,ss_s} ws_e(t)$  from  $ws_e$  on  $ss_s$ :

$$c_{ss_s}(V_1, V_2) = V_1(1 + V_2)$$

Since the connection weight  $\omega_{ws_e,ss_s}$  is chosen negative (it is a suppressing link), for example -1, this function makes the sensing of stimulus  $s$  inversely proportional to the extent  $ws_e(t)$  of avoidance; e.g. sensing  $s$  becomes 0 when avoidance  $e$  is 1, and  $V_1$  when avoidance  $e$  is 0. According to this combination function the difference and differential equation for  $ss_s$  are as follows:

$$ss_s(t + \Delta t) = ss_s(t) + \eta_{ss_s} [\omega_{ws_s,ss_s} ws_s(t)(1 - \omega_{ws_e,ss_s} ws_e(t)) - ss_s(t)] \Delta t$$

$$dss_s/dt = \eta_{ss_s} [\omega_{ws_s,ss_s} ws_s(t)(1 - \omega_{ws_e,ss_s} ws_e(t)) - ss_s(t)]$$

The combination functions for all states with only one connection toward it are the identity function, except for  $es_a$  in which case the advanced logistic function  $\mathbf{alogistic}_{\sigma,\tau}(\dots)$  is used. The combination function for  $ps_a$  is also the advanced logistic function  $\mathbf{alogistic}_{\sigma,\tau}(\dots)$ .

In Fig. 8 an example simulation with the model depicted in Fig. 7 clearly shows how a limit cycle pattern emerges, with period 18.5.

Here all connection weights are 1, except the weight of the suppressing connection from  $ws_e$  to  $ss_s$ , which is -1. Moreover, the steepness  $\sigma$  and threshold  $\tau$  for  $ps_a$  are 4 and 0.9, respectively, and for  $es_a$  they are 40 and 0.7. The step size  $\Delta t$  was 0.1 and the speed factors  $\eta$  for  $es_a$  and  $ws_e$  were 0.4, and for the other (internal) states  $\eta$  was 1.

For this simulation an analysis of the stationary points has been performed for the maxima and minima in the final stage for all states. Recall from Sect. 2 that the equation expressing that a state  $Y$  is stationary at time  $t$  is

$$\mathbf{aggimpact}_Y(t) = Y(t)$$

which is equivalent to

$$c_Y(\omega_{X_1,Y} X_1(t), \dots, \omega_{X_k,Y} X_k(t)) = Y(t)$$

For example, for state  $ps_a$  the combination function is the sum function, so the aggregated impact is

$$\mathbf{aggimpact}_{ps_a}(t) = \omega_{\text{responding}SRS_s}(t) + \omega_{\text{amplifying}SRS_e}(t)$$

Then the stationary point equation expressing that state  $ps_a$  is stationary at time  $t$  is

$$\omega_{\text{responding}SRS_s}(t) + \omega_{\text{amplifying}SRS_e}(t) = ps_a(t)$$

It is this equation that has been checked for the minima and maxima for each of the states in the final stage of the simulation. The results are shown in Table 1. Here both for the maxima and for the minima the first rows show the time points at which the stationary point occurs. The next row (state value) shows the values of the right-hand side of the above equation, followed by rows (aggregated impact) showing the left-hand sides of this equation, and then a row with the absolute deviation between the values in the two rows above it.

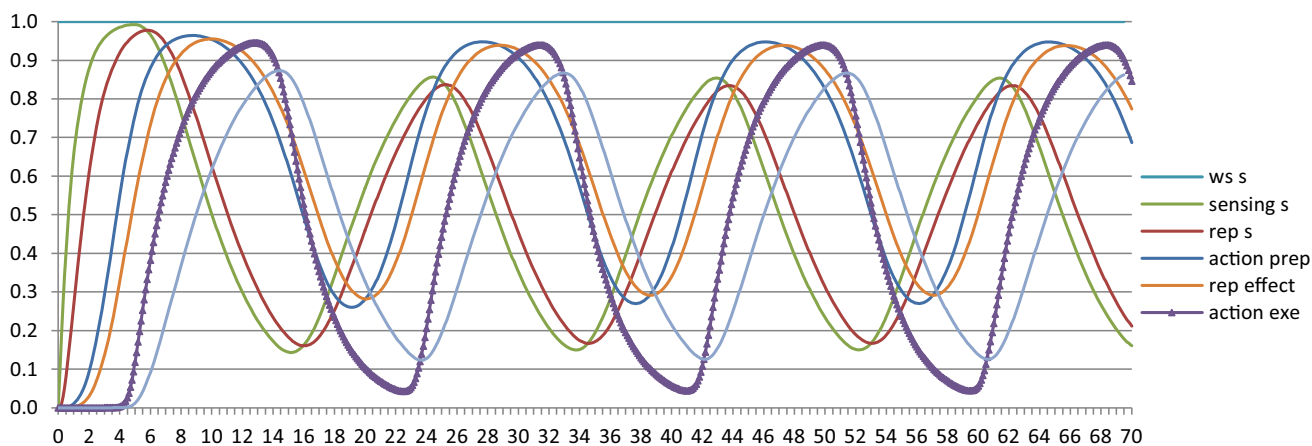


Fig. 7 Example simulation showing a limit cycle

Table 1 Overview of the outcomes of a mathematical analysis for stationary points in a limit cycle

	$ws_e$	$ss_s$	$srs_s$	$srs_e$	$ps_a$	$es_a$
<b>Maxima</b>						
Time point	69.9	61.4	62.2	65.8	64.6	68.3
State value	0.86700	0.85401	0.83455	0.93859	0.94754	0.93975
Aggregated impact	0.86105	0.85001	0.83325	0.93713	0.94703	0.94012
Absolute deviation	0.00595	0.00400	0.00131	0.00146	0.00051	0.00037
<b>Minima</b>						
Time point	60.6	52.2	53.0	57.1	56.1	59.4
State value	0.12553	0.14993	0.16699	0.29153	0.27012	0.04400
Aggregated impact	0.13033	0.15168	0.16689	0.29480	0.27159	0.04317
Absolute deviation	0.00480	0.00175	0.00009	0.00327	0.00147	0.00083

It turns out that the stationary point equations are fulfilled with an average accuracy over all states and stationary points of 0.002 and a maximal accuracy of 0.006, which both is  $< 10^{-2}$ . This provides evidence that the implemented model is correct in comparison to the model description. In Table 1 the more specific numbers are shown for the different states. For the maxima the average deviation is 0.00226, and the maximal absolute deviation is 0.00595 (which occurs for state  $ws_e$ ). For the minima the average absolute deviation is 0.00204, and the maximal absolute deviation is 0.00480 (which again is for state  $ws_e$ ). Taken minima and maxima together, the overall average absolute deviation is 0.00215, and the maximal absolute deviation is 0.00595 (for the maxima of state  $ws_e$ ).

As another type example of the emergence of limit cycle behaviour, consider that in a realistic context stimuli can be present for some time, but also may be absent for certain periods according to fixed periods, for example, day/night rhythms. As an example, for Hebbian learning, for activations based on stimuli that return from time to time an analysis can be made about when there is enough stimulation over time

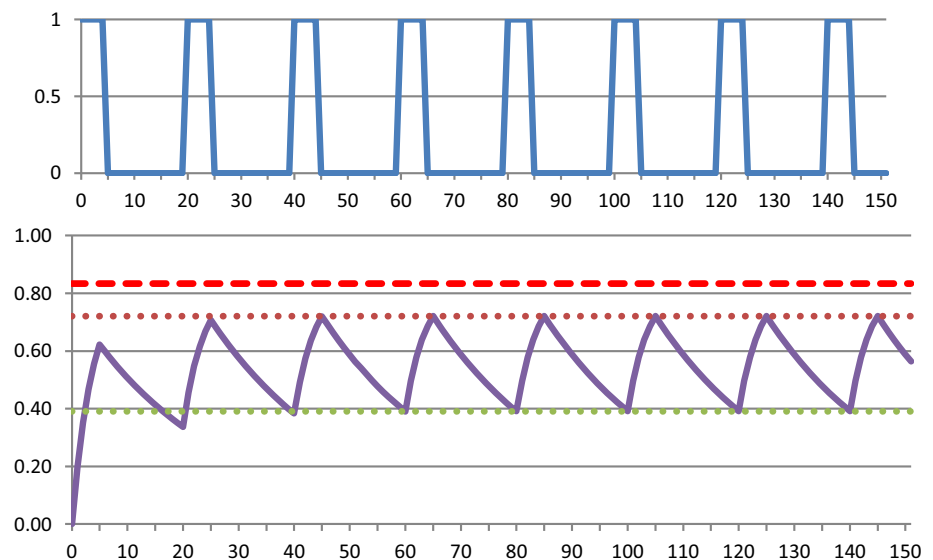
to achieve or maintain a value for the weight  $\omega$  of some connection. As an example, see the pattern in Fig. 7, where the upper graph shows the levels of both  $X_1$  and  $X_2$  (alternating between 0 and 1) and the lower graph shows how due to these activation periods, the periods of learning ( $d_1 = 5$  time units) and pure extinction ( $d_0 = 15$  time units) alternate. It turns out that there is a form of convergence not to one specific value of  $\omega$ , but to a recurring pattern that repeats itself; this is a specific case of a limit cycle, in this case induced by environmental fluctuations.

### 8 Discussion

In this paper it was discussed how mathematical analysis can be used to find out some properties of a model. An advantage is that this is done without performing simulations. This advantage makes that it can be used as an additional source of knowledge, independent of a specific implementation of the model. By comparing properties found by mathematical analysis and properties observed in simulation experiments



**Fig. 8** Limit cycle for  $d_1 = 5$  (learning),  $d_0 = 15$  (pure extinction), and  $\eta = 0.2$ ,  $\zeta = 0.04$  Equilibrium value 0.83,  $\omega_{\max} = 0.72$ ,  $\omega_{\min} = 0.39$  (dotted lines)



some form of verification can be done. If a discrepancy is found, for example, in the sense that the mathematical analysis predicts a certain property but some simulation does not satisfy this property, this can be a reason to inspect the implementation of the model carefully (and/or check whether the mathematical analysis is correct). Having such an option can be fruitful during a development process of a model, as to acquire empirical data for validation of a model may be more difficult or may take a longer time.

The techniques used for such mathematical analysis were adopted from [3, 9, 11–13]. In this literature many more techniques can be found than those covered in the current paper, for example, for the convergence speed (e.g. [10]) for attracting equilibria, but also for other types of properties. For example, there is underlying theory that proves the existence of certain patterns, for example, theorems from Poincaré (1881–1882) and Bendixson (1901) that state that under certain circumstances for two-dimensional systems (described by only two differential equations) limit cycles will occur. These are beyond the scope of this paper.

Mathematical analysis is not always easy or feasible. For example, linear equilibrium equations (for example, obtained when using scaled sum combination functions) in principle can be solved analytically in a generic form, thereby obtaining expressions for the equilibrium values in terms of the parameters of the model, but equilibrium equations involving logistic functions cannot be solved in such a manner. Nevertheless, for such cases often specific instances can be solved. Moreover, as discussed in Sect. 2, verification of a model does not depend on finding explicit analytical solutions of the equilibrium equations. For verification it is already sufficient if the equilibrium equations have been identified, which is always possible from the difference or differential equations. Then for each simulation trace observed equilibrium

values can be substituted in these equations and by this it is checked whether they satisfy the equations. Therefore, in general, mathematical analysis still can add some value, in addition to systematic simulation experiments. However, a limitation is that although verification is always possible, prediction is not. For prediction without having any simulation, it is needed to find explicit analytical solutions of the equilibrium equations, and in many realistic models this is not feasible.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Bendixson, I.: Sur les courbes définies par des équations différentielles. *Acta Math.* **24**, 1–88 (1901)
2. Bi, G., Poo, M.: Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* **24**, 139–166 (2001)
3. Brauer, F., Nohel, J.A.: *Qualitative Theory of Ordinary Differential Equations*. Benjamin, Amsterdam (1969)
4. Garagnani, M., Wennekers, T., Pulvermueller, F.: A neuroanatomically grounded Hebbian-learning model of attention-language interactions in the human brain. *Eur. J. Neurosci.* **27**(2008), 492–513 (2008)
5. Gerstner, W., Kistler, W.M.: *Mathematical formulations of Hebbian learning*. *Biol. Cybern.* **87**, 404–415 (2002)
6. Hebb, D.: (1949) *The Organisation of Behavior*. Wiley, New York (1949)
7. Hirsch, M.W.: The dynamical systems approach to differential equations. *Bull. (New Series) Am. Math. Soc.* **11**, 1–64 (1984)

8. Keysers, C., Perrett, D.I.: Demystifying social cognition: a Hebbian perspective. *Trends Cognit. Sci.* **8**(2004), 501–507 (2004)
9. Lotka, A.J.: *Elements of Physical Biology*. Williams and Wilkins Co. (1924), Dover Publications, 2nd ed. (1956)
10. Mathunjwa, J.S., Temple, J. (2006). Convergence behaviour in exogenous growth models. Discussion Paper No. 06/590, Department of Economics, University of Bristol
11. Picard, E.: *Traité d'Analyse*, vol. 1 (1891), vol. 2 (1893)
12. Poincaré, H.: *Mémoire sur les courbes définies par une équation différentielle (1881–1882)* (On curves defined by differential equations)
13. Poincaré, H.: *New methods of celestial mechanics*, vol. 3 (1892–1899) English translation (1967)
14. Sharpanskykh, A., Treur, J.: Modelling and analysis of social contagion in dynamic networks. *Neurocomputing J.* **146**(2014), 140–150 (2014)
15. Treur, J.: (2011a) A computational agent model for Hebbian learning of social interaction. In: Lu, B.-L., Zhang, L., Kwok, J. (eds.) *Proceedings of the 18th international conference on neural information processing, ICONIP'11, Part I. Lecture Notes in Artificial Intelligence*, vol. 7062, pp. 9–19. Springer, Berlin (2011)
16. Treur, J.: (2011b) Dreaming your fear away: a computational model for fear extinction learning during dreaming. In: Lu, B.-L., Zhang, L., Kwok, J. (eds.) *Proceedings of the 18th international conference on neural information processing, ICONIP'11, Part III. Lecture Notes in Artificial Intelligence*, vol. 7064, pp. 197–209. Springer, Berlin (2011)
17. Treur, J. (2016) Dynamic modeling based on a temporal-causal network modeling approach. *Biologically Inspired Cognitive Architectures*, 2016, ResearchGate [https://www.researchgate.net/publication/289193241\\_Dynamic\\_Modeling\\_Based\\_on\\_a\\_Temporal-Causal\\_Network\\_Modeling\\_Approach](https://www.researchgate.net/publication/289193241_Dynamic_Modeling_Based_on_a_Temporal-Causal_Network_Modeling_Approach) (to appear)