

RESEARCH

Open Access



Mathematical analysis of the emergence of communities based on coevolution of social contagion and bonding by homophily

Jan Treur 

Correspondence: j.treur@vu.nl;
https://www.researchgate.net/profile/Jan_Treur
Behavioural Informatics Group, Vrije
Universiteit Amsterdam,
Amsterdam, Netherlands

Abstract

In this paper it is analysed how community formation in an adaptive network for bonding based on similarity (homophily) can be related to characteristics of the adaptive network's structure, which includes the structure of the adaptation principles incorporated. In particular, this is addressed for adaptive social networks for bonding based on homophily. To this end, relevant properties of the network and the adaptation principle have been identified, such as a tipping point for similarity. As one of the results, it has been found how the emergence of communities strongly depends on the value of this tipping point. Moreover, it is shown that some properties of the structure of the network and the adaptation principle entail that the connection weights all converge to 0 (for persons in different communities) or 1 (for persons within one community).

Introduction

In this paper, it is analysed how emerging communities based on the coevolution of social contagion (Levy and Nail 1993) and bonding by homophily (McPherson et al. 2001; Pearson et al. 2006) can be related to characteristics of the adaptive network's structure. For this adaptive case, this network structure includes the structure of the homophily adaptation principle that is incorporated. The homophily adaptation principle expresses how 'being alike' strengthens the connection between two persons (McPherson et al. 2001; Pearson et al. 2006). Social contagion implies that the stronger two persons are connected, the more they will become alike (Levy and Nail 1993). Thus a reciprocal causal relation between the two processes occurs. It is known from simulations (Axelrod 1997; Holme and Newman 2006; Sharpanskykh and Treur 2014; Vazquez 2013; Vazquez et al. 2007) that the emerging behaviour of adaptive network models combining these two processes as a form of coevolution often shows community formation. In the resulting network structure within a community persons have high mutual connection weights and a high extent of 'being alike', and persons from different communities have low mutual connection weights and a low extent of 'being alike'.

Relevant properties of the network and the homophily adaptation principle have been identified, such as a tipping point for homophily. As one of the results, it has been found how the emergence of communities strongly depends on the value of this tipping point. Moreover, it is shown that some properties of the structure of the network and the homophily adaptation principle entail that the connection weights all converge to 0 (for states in different emerging communities) or 1 (for states within one emerging community).

In general, it is a challenging issue for dynamic models to predict what patterns of behaviour will emerge, and how their emergence depends on the structure of the model. This structure includes chosen parameter values used as characteristics of the model's structure (and settings). This also applies to network models, where behaviour depends in some way on the network structure, defined by network characteristics such as connections and their weights; e.g., (Turnbull et al. 2018). In this context, the issue is how emerging network behaviour relates to network structure. For example, see (Treur, 2018b) it can be an even more challenging issue when coevolution that occurs in adaptive networks is considered, in which case by a mutual causal interaction both the states and the network characteristics change over time. In this case the connections in the network change according to certain adaptation principles which themselves depend on certain adaptation characteristics represented by their own parameters.

In the current paper, the emergence of communities based on the coevolution of social contagion (Levy and Nail 1993) and bonding by homophily (McPherson et al. 2001; Pearson et al. 2006) will be analysed in some depth. By mathematical analysis, it is found out how emerging communities relate to the characteristics of the network and of the specific homophily adaptation principle used in combination with a social contagion principle.

The issue was addressed using the Network-Oriented Modeling approach based on temporal-causal networks (Treur 2016b, Treur 2019) as a vehicle. For temporal-causal networks, parameters characterising the network structure are connection weights, combination functions, and speed factors. For the type of adaptive networks considered, the connection weights are dynamic based on the homophily principle. So they are considered as states themselves and not part of the static characteristics of the network structure anymore. In the research reported here characteristics of the homophily adaptation principle have been identified that play an important role in community formation, among which the tipping point for the difference between two persons. This is the point where 'being alike' turns into 'not being alike', or conversely. In this paper, results are discussed that have been proven mathematically for this relationship between network structure and network behavior for the coevolution process. In particular, for emerging communities what the connections become between persons from one community or persons from different communities, and how different persons from one community or from different communities become. Note that these results have not been proven for one specific model or function, but for whole classes of models with a variety of functions that fulfill certain properties. The technical content has been adopted mostly from the conference paper (Treur 2018c), but extended by more than 50%. One of the additions is the analysis of the relation of the obtained results to the strongly connected components of a network, for example, as defined in (Harary et al. 1965, Ch. 3; Kuich 1970, Section 6).

In this paper, in Section 2 the Network-Oriented Modeling approach used is briefly outlined. In Section 3 adaptive networks based on homophily are described and a number of functions that can be used to model them. Section 4 shows some example simulations. In Section 5 relevant properties are defined that are used in Section 6 to prove results on the relation between network structure and behaviour. In Section 7 it is shown how the obtained results relate to the strongly connected components of a network. Finally, Section 8 is a discussion.

Network-oriented modeling by temporal-causal networks

In order to undertake any mathematical analysis of networks, in the first place, a solid definition of the concept of network is needed, based on well-defined semantics. In the current paper, the interpretation of connections based on causality and dynamics forms a basis of the structure and semantics of the considered networks. It is a deterministic dynamic modeling approach, for example in the line of (Ashby 1960), in contrast to, stochastic modeling approaches as, for example, used in (Axelrod 1997). It can be positioned as a branch in the causal modelling area which has a long tradition in AI; e.g., see (Kuipers and Kassirer 1983; Kuipers 1984; Pearl 2000). It distinguishes itself by a dynamic perspective on causal relations, according to which causal relations exert causal effects over time, and these causal relations themselves can also change over time.

More specifically, the nodes in a network are interpreted here as states (or state variables) that vary over time, and the connections are interpreted as causal relations that define how each state can affect other states over time. This type of network has been called a *temporal-causal network* (Treur 2016b); note that the word temporal here refers to the causality, not to the network. Temporal-causal networks that themselves change over time as well are called adaptive temporal-causal networks; e.g., (Gross and Sayama 2009). So, in cases of adaptive temporal-causal networks, in addition to the node states also the connections are assumed to change over time and are therefore treated like states as well.

A conceptual representation of a temporal-causal network model by a *labeled graph* provides a fundamental basis. Such a conceptual representation includes representing in a declarative manner states (also called nodes) and connections between them that represent (causal) impacts of states on each other. This part of a conceptual representation is often depicted in a *conceptual picture* by a graph with nodes and directed connections. However, a *complete conceptual representation* of a temporal-causal network model also includes a number of labels for such a graph. A notion of *strength of a connection* is used as a label for connections, some way to *aggregate multiple causal impacts* on a state is used, and a notion of *speed of change* of a state is used for timing of the processes. Note that states have one value; they can relate one by one to states of persons or agents, for example, the strength of their opinion states. It is also possible to model each person by more than one state, for example, an opinion and an emotion state per person. In such a case a person does not relate to a single state but to a sub-network consisting of multiple states.

The three described notions, called connection weight $\omega_{X,Y}$, combination function $c_Y(\cdot)$, and speed factor η_Y , make the graph of states and connections a labeled graph (e.g., see Fig. 1), and form the defining structure of a temporal-causal network model in

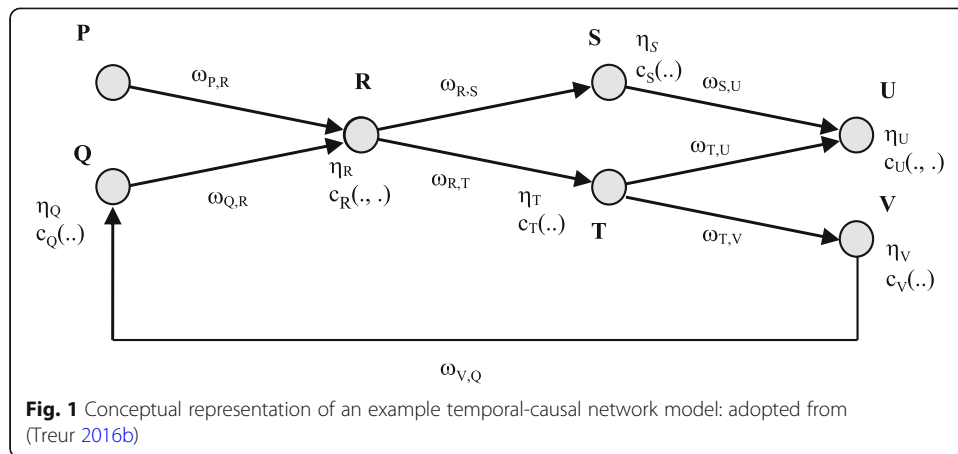


Fig. 1 Conceptual representation of an example temporal-causal network model: adopted from (Treur 2016b)

the form of a conceptual representation; for a summary, see also Table 1, top half: rows 1 to 5. Note that for the current paper all connection weights are assumed nonnegative: $\omega_{X,Y} \in [0, 1]$ for all X and Y .

The interpretation of a network based on causality and dynamics can be expressed in a formal-numerical way, thus associating semantics to any conceptual temporal-causal network representation in a detailed numerical-mathematically defined manner. For a summary, see Table 1, bottom half: rows 6 to 10. This shows how a conceptual representation based on states and connections enriched with labels for connection weights, combination functions, and speed factors, can be transformed into a numerical

Table 1 Conceptual and numerical representations of a temporal-causal network model

Concept	Conceptual representation	Explanation
States and connections	$X, Y, X \rightarrow Y$	Describes the nodes and links of a network structure (e.g., in graphical or matrix format)
Connection weight	$\omega_{X,Y}$	The <i>connection weight</i> $\omega_{X,Y} \in [-1, 1]$ represents the strength of the causal impact of state X on state Y through connection $X \rightarrow Y$
Aggregating multiple impacts on a state	$c_Y(\cdot)$	For each state Y (a reference to) a <i>combination function</i> $c_Y(\cdot)$ is chosen to combine the causal impacts of other states on state Y
Timing of the effect of causal impact	η_Y	For each state Y a <i>speed factor</i> $\eta_Y \geq 0$ is used to represent how fast a state is changing upon causal impact
Concept	Numerical representation	Explanation
State values over time t	$Y(t)$	At each time point t each state Y in the model has a real number value in $[0, 1]$
Single causal impact	$\text{impact}_{X,Y}(t) = \omega_{X,Y}X(t)$	At t state X with a connection to state Y has an impact on Y , using connection weight $\omega_{X,Y}$
Aggregating multiple causal impacts	$\text{aggimpact}_Y(t) = c_Y(\text{impact}_{X_1,Y}(t), \dots, \text{impact}_{X_k,Y}(t)) = c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t))$	The aggregated causal impact of multiple states X_i on Y at t , is determined using combination function $c_Y(\cdot)$
Timing of the causal effect	$Y(t + \Delta t) = Y(t) + \eta_Y [\text{aggimpact}_Y(t) - Y(t)] \Delta t = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) - Y(t)] \Delta t$	The causal impact on Y is exerted over time gradually, using speed factor η_Y ; here the X_i are all states with outgoing connections to state Y

representation (Treur 2016b), Ch. 2. A more detailed explanation of the difference equation format, taken from (Treur 2016b), Ch 2, p. 60–61, is as follows; see also Fig. 2. The aggregated impact value $\mathbf{aggimpact}_Y(t)$ at time t pushes the value of Y up or down, depending on how it compares to the current value of Y . So, $\mathbf{aggimpact}_Y(t)$ is compared to the current value $Y(t)$ of Y at t by taking the difference between them (also see Fig. 2): $\mathbf{aggimpact}_Y(t) - Y(t)$. If this difference is positive, which means that $\mathbf{aggimpact}_Y(t)$ at time t is higher than the current value of Y at t , in the time step from t to $t + \Delta t$ (for some small Δt) the value $Y(t)$ will increase in the direction of the higher value $\mathbf{aggimpact}_Y(t)$. This increase is done proportionally to the difference, with proportion factor $\eta_Y \Delta t$: the increase is $\eta_Y [\mathbf{aggimpact}_Y(t) - Y(t)] \Delta t$; see Fig. 2. By this format, the parameter η_Y indeed acts as a speed factor by which it can be specified how fast state Y should change upon causal impact.

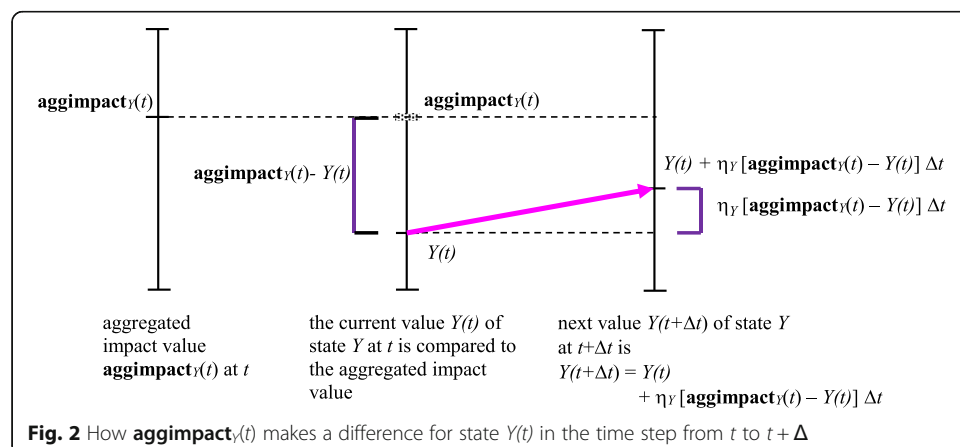
There are many different approaches possible to address the issue of combining multiple causal impacts. To provide sufficient flexibility, the Network-Oriented Modeling approach based on temporal-causal networks incorporates for each state a way to specify how multiple causal impacts on this state are aggregated by a combination function. For this aggregation, a library with a number of standard combination functions are available as options (currently 35), but also own-defined functions can be added. The difference equations in Fig. 2 and in the last row in Table 1 constitute the overall numerical representation of the temporal-causal network model and can be used for simulation and mathematical analysis; it can also be written in differential equation format:

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

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

For adaptive networks, connection weights are dynamic and therefore handled in the same way as states, with their own combination functions and speed factors; connections of them have weight value 1 by default. This will be shown in more detail in Sections 3 to 5.

Often used examples of combination functions are the *identity* function $\mathbf{id}(\cdot)$ for states with impact from only one other state, the *scaled sum* function $\mathbf{ssum}_\lambda(\cdot)$ with scaling factor λ , and the *advanced logistic sum* combination function $\mathbf{alogistic}_{\sigma,\tau}(\cdot)$ with steepness σ and threshold τ :



$$\begin{aligned} \mathbf{id}(V) &= V \\ \mathbf{ssum}_\lambda(V_1, \dots, V_k) &= \frac{V_1 + \dots + V_k}{\lambda} \\ \mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k) &= \left[\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right] (1 + e^{-\sigma\tau}) \end{aligned}$$

Note that for $\lambda = 1$, the scaled sum function is just the sum function $\mathbf{sum}(\cdot)$, and this sum function can also be used as the identity function in case of one incoming connection. In addition to the above functions, a *Euclidean combination function* is defined as

$$\mathbf{c}(V_1, \dots, V_k) = \mathbf{eucl}_{n, \lambda}(V_1, \dots, V_k) = \sqrt[n]{\frac{V_1^n + \dots + V_k^n}{\lambda}}$$

where n is the *order* (which can be any natural number but also any positive real number), and λ is a scaling factor. This can be used when all connection weights are non-negative, but in the specific case that n is an odd natural number, also negative connection weights can be allowed. A Euclidean combination function is called *normalised* if the scaling factor λ is chosen in such a way that $\mathbf{c}(\omega_{X_1, Y}, \dots, \omega_{X_k, Y}) = 1$; this is achieved for $\lambda = \omega_{X_1, Y}^n + \dots + \omega_{X_k, Y}^n$. Note that for $n = 1$ (first order) the scaled sum function is obtained:

$$\mathbf{eucl}_{1, \lambda}(V_1, \dots, V_k) = \mathbf{ssum}_\lambda(V_1, \dots, V_k)$$

Then $\lambda = \omega_{X_1, Y} + \dots + \omega_{X_k, Y}$ makes it normalised. For $n = 2$ it is the second-order Euclidean combination function $\mathbf{eucl}_{2, \lambda}(\cdot)$ defined by:

$$\mathbf{eucl}_{2, \lambda}(V_1, \dots, V_k) = \sqrt{\frac{V_1^2 + \dots + V_k^2}{\lambda}}$$

Such a second-order Euclidean combination function is also often applied in aggregating the error value in optimisation and in parameter tuning using the root-mean-square deviation (RMSD), based on the Sum of Squared Residuals (SSR).

Adaptive networks based on homophily

The homophily principle addresses bonding between persons. It describes how connections between two persons are strengthened or weakened depending on the extent of similarity between them: more ‘being alike’ will make the persons more ‘like each other’ (McPherson et al. 2001; Pearson et al. 2006). This is modelled by the states $X(t)$ and $Y(t)$ the persons X and Y have at time t , for example, indicating whether at time t they enjoy being physically active (high value) or not (low value), or indicating the extent to which they agree with a certain opinion. According to the homophily principle the weight $\omega_{X, Y}$ of the connection from X to Y is changing over time dynamically, depending on how the state levels $X(t)$ and $Y(t)$ differ. As this connection weight $\omega_{X, Y}$ is dynamic, in a general setting it is handled as a state with its own combination function $\mathbf{c}_{\omega_{X, Y}}(V_1, V_2, W)$, called *homophily combination function*. Then the standard difference and differential equation format as shown in Section 2 is applied; here V_1 stands for $X(t)$, V_2 stands for $Y(t)$, and W stands for connection weight $\omega_{X, Y}(t)$:

$$\begin{aligned} \omega_{X,Y}(t + \Delta t) &= \omega_{X,Y}(t) + \eta_{\omega_{X,Y}} [\mathbf{c}_{\omega_{X,Y}}(X(t), Y(t), \omega_{X,Y}(t)) - \omega_{X,Y}(t)] \Delta t \\ \mathbf{d}\omega_{X,Y}/\mathbf{d}t &= \eta_{\omega_{X,Y}} [\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y}) - \omega_{X,Y}] \end{aligned}$$

The parameter $\eta_{\omega_{X,Y}}$ is the speed parameter of connection weight $\omega_{X,Y}$. Note that $\omega_{X,Y}$ increases if and only if $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y}) > \omega_{X,Y}$, $\omega_{X,Y}$ decreases if and only if $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y}) < \omega_{X,Y}$, and $\omega_{X,Y}$ is stationary if and only if $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y}) = \omega_{X,Y}$. Specific adaptation principles formalising bonding by homophily can be obtained by picking specific functions for the homophily combination function $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y})$ for a given application domain, or even different combination functions $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y})$ for different persons within a given domain. There are many options for such choices; see, for example, Table 2 and Fig. 3. Therefore it may be more useful to define classes of such homophily combination functions characterised by certain properties of them. Then as shown in Section 6, results can be proven for such a class instead of for each homophily combination function separately. Formal definitions of such properties of the function $\mathbf{c}_{\omega_{X,Y}}(X, Y, \omega_{X,Y})$ will be given in Section 5, but here some examples of homophily combination functions are shown.

A very simple example of a homophily combination function $\mathbf{c}_{\omega_{X,Y}}(V_1, V_2, W)$ is a linear function in $D = |V_1 - V_2|$ defined as follows:

$$\mathbf{c}_{\omega_{X,Y}}(V_1, V_2, W) = W + \beta (\tau - D)$$

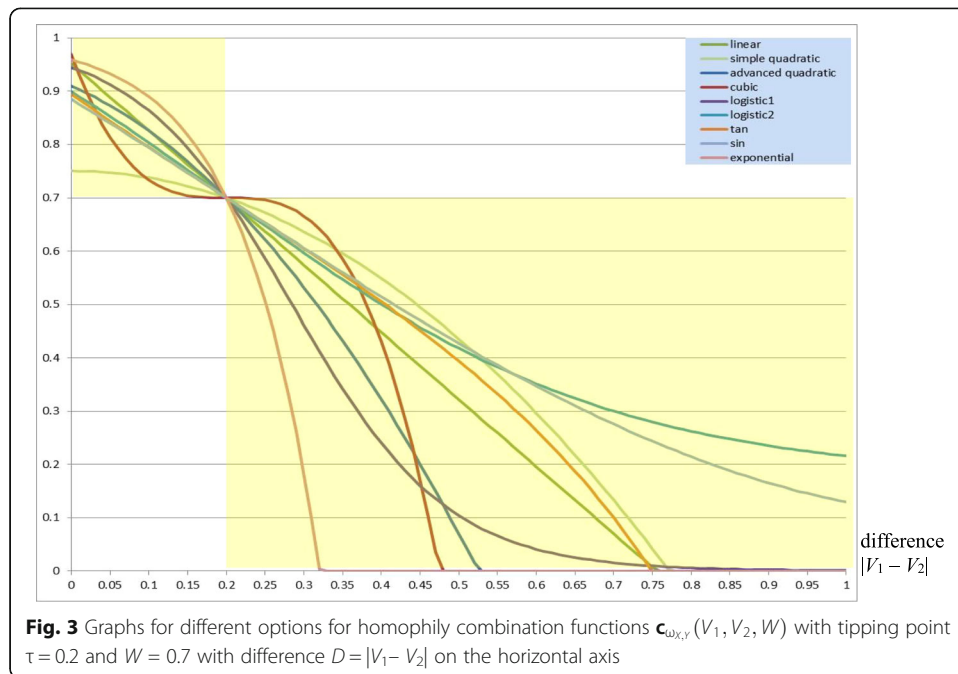
Here β is a modulation factor that still can be chosen, and τ is a tipping point (or threshold) parameter. This function may have the disadvantage that when $W = 0$ it may be negative (when $D > \tau$) or when $W = 1$ it may be higher than 1 (when $D < \tau$), so that it has to be cut off to avoid that the weight value $\omega_{X,Y}$ goes outside the interval $[0, 1]$. This can also be remedied in a smooth manner by choosing β as a function $\beta(W) = \alpha W(1-W)$ of W which can suppress the term $\tau - D$ when W comes closer to 0 or 1. This function makes that $\omega_{X,Y}$ will not cross the boundaries 0 and 1:

$$\mathbf{c}_{\omega_{X,Y}}(V_1, V_2, W) = W + \alpha W (1-W) (\tau - |V_1 - V_2|)$$

Here α is a modulation or amplification parameter; the higher its value, the stronger the effect (either positive or negative) of homophily on the connection. This combination function for bonding by homophily was considered in (Treur 2017). Using this homophily combination function, the dynamic relations for $\omega_{X,Y}$ are:

Table 2 Different options for combination functions $\mathbf{c}_{\omega_{X,Y}}(V_1, V_2, W)$ for the homophily principle based on a tipping point τ ; for the graphs depending on $D = |V_1 - V_2|$, see Fig. 3

Function type	Function name	Numerical representation	Parameters in Fig. 3
Simple Linear	slhomo $_{\tau, \alpha}(V_1, V_2, W)$	$W + \alpha W (1-W) (\tau - D)$	$\alpha = 6$
Simple quadratic 1	sq1homo $_{\tau, \alpha}(V_1, V_2, W)$	$W + \alpha W (1 - W) (\tau^2 - D^2)$	$\alpha = 6$
Simple quadratic 2	sq2homo $_{\tau, \alpha, \delta}(V_1, V_2, W)$	$W + \alpha ((\tau + \delta)^2 - (D + \delta)^2)$	$\delta = 0.15, \alpha = 10$
Cubic	cubehomo $_{\tau, \alpha}(V_1, V_2, W)$	$W + \alpha (1-W) (1-D/\tau)^3$	$\alpha = 0.9$
Logistic 1	log1homo $_{\tau, \alpha}(V_1, V_2, W)$	$\frac{W}{W + (1-W) e^{\alpha(D-\tau)}}$	$\sigma = 10$
Logistic 2	slog2homo $_{\tau, \alpha}(V_1, V_2, W)$	$W + \alpha \frac{W(1-W)}{1 + e^{-\alpha(D-\tau)}}$	$\sigma = 4, \alpha = 5$
Sine-based	sinhomo $_{\tau, \alpha}(V_1, V_2, W)$	$W - \alpha (1-W) \sin(\pi (D-\tau)/2)$	$\alpha = 2$
Tangent-based	tanhomo $_{\tau, \alpha}(V_1, V_2, W)$	$W - \alpha (1-W) \tan(\pi (D-\tau)/2)$	$\alpha = 2$
Exponential	exphomo $_{\tau, \alpha}(V_1, V_2, W)$	$1 - (1-W) e^{\alpha(D-\tau)}$	$\sigma = 10$



$$\begin{aligned} \frac{d\omega_{X,Y}}{dt} &= \eta_{\omega_{X,Y}} \alpha \omega_{X,Y} (1 - \omega_{X,Y}) (\tau - |X - Y|) \\ \omega_{X,Y}(t + \Delta t) &= \omega_{X,Y}(t) + \eta_{\omega_{X,Y}} \alpha \omega_{X,Y}(t) (1 - \omega_{X,Y}(t)) (\tau - |X(t) - Y(t)|) \Delta t \end{aligned}$$

As a variant of this homophily combination function $c_{\omega_{X,Y}}(V_1, V_2, W)$ that is linear in D , the following function can be obtained as a quadratic function of $D = |V_1 - V_2|$:

$$c_{\omega_{X,Y}}(V_1, V_2, W) = W + \alpha W(1 - W) (\tau^2 - (V_1 - V_2)^2)$$

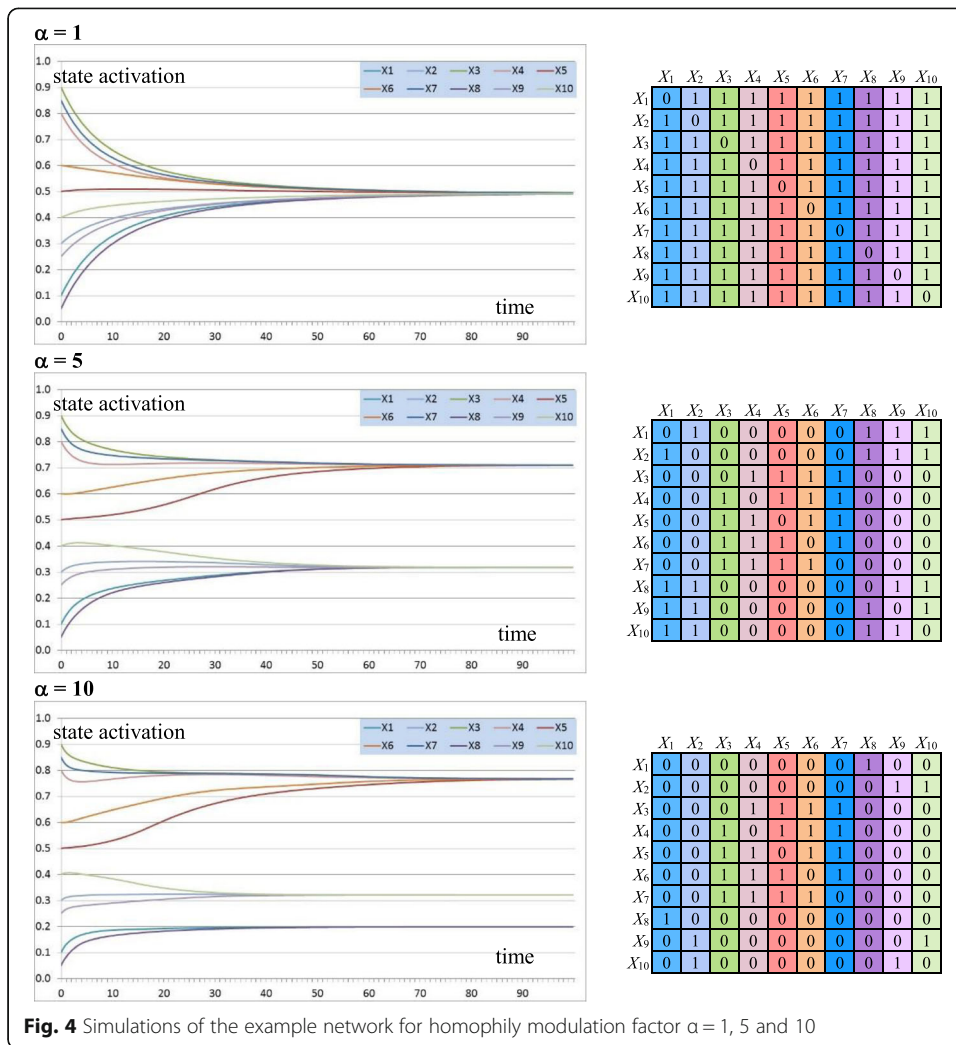
Using this combination function, the dynamic relations for $\omega_{X,Y}$ are:

$$\begin{aligned} \frac{d\omega_{X,Y}}{dt} &= \eta_{\omega_{X,Y}} \alpha \omega_{X,Y} (1 - \omega_{X,Y}) (\tau - (X - Y)^2) \\ \omega_{X,Y}(t + \Delta t) &= \omega_{X,Y}(t) + \eta_{\omega_{X,Y}} \alpha \omega_{X,Y}(t) (1 - \omega_{X,Y}(t)) (\tau - (X(t) - Y(t))^2) \Delta t \end{aligned}$$

In Table 2 and Fig. 3 these linear and quadratic combination function and a number of other examples of homophily combination functions are depicted, with $D = |V_1 - V_2|$ on the horizontal axis and $c_{\omega_{X,Y}}(V_1, V_2, W)$ on the vertical axis for $\tau = 0.2$ and $W = 0.7$. In Section 6 the results that have been proven are general in the sense that they apply to all of such functions, not just one specific function. To this end, in Section 5 relevant properties shared by these functions are defined. The simple linear combination function $\mathbf{slhomo}_{\tau,\alpha}(V_1, V_2, W)$ was used in (Blankendaal et al. 2016) and (Treur 2017), and the simple quadratic combination function $\mathbf{sqhom}_{\tau,\alpha}(V_1, V_2, W)$ in (van Beukel et al. 2017). A more advanced logistic combination function based on $\mathbf{log2homo}_{\tau,\alpha,\alpha}(V_1, V_2, W)$ was explored in (Boomgaard et al. 2018).

Example simulations for the coevolution of social contagion and bonding by homophily

In this section, a few example simulations for the coevolution of social contagion and bonding based on homophily are described. It will be shown how communities emerge



just above the tipping point $\tau = 0.1$. This may suggest that the model keeps the clusters at a distance of at least τ ; this is one of the issues that will be analysed further in Section 6. For higher values of α , such as $\alpha = 15$ no more than these four clusters emerge.

Relevant properties of combination functions for homophily and social contagion

This section addresses definitions for the properties for the network structure and in particular also for the homophily adaptation principle that have been identified as relevant for the adaptive network’s behavior. As adaptation principles are specified by their own combination functions, these properties are described as properties of homophily combination functions. The following are considered plausible assumptions for a homophily combination function; here $D = |V_1 - V_2|$:

- $c_{\omega_{X,Y}}(V_1, V_2, W)$ is a function: $[0, 1] \times [0, 1] \times [0, 1] \rightarrow [0, 1]$
- $c_{\omega_{X,Y}}(V_1, V_2, W)$ is a monotonically decreasing function of D
- For D close to 0 and $W < 1$ it holds $c_{\omega_{X,Y}}(V_1, V_2, W) > W$ (i.e., $\omega_{X,Y}$ will increase)

be shown in Section 6 how these properties relate to the equilibrium behaviour of the network.

Definition 1 (Tipping point and strict tipping point)

- a) The function $c(V_1, V_2, W): [0, 1] \times [0, 1] \times [0, 1] \rightarrow [0, 1]$ has *tipping point* τ for V_1 and V_2 if for all W with $0 < W < 1$ and all V_1, V_2 it holds
 - (i) $c(V_1, V_2, W) > W \quad |V_1 - V_2| < \tau$
 - (ii) $c(V_1, V_2, W) = W \quad |V_1 - V_2| = \tau$
 - (iii) $c(V_1, V_2, W) < W \quad |V_1 - V_2| > \tau$
- b) The function $c(V_1, V_2, W)$ has a *strict tipping point* τ if it has tipping point τ and in addition it holds:
 - (i) If $|V_1 - V_2| < \tau$ then $c(V_1, V_2, 0) > 0$
 - (ii) If $|V_1 - V_2| > \tau$ then $c(V_1, V_2, 1) < 1$

Note Definition 1a) can be reformulated in the sense that any function $c(V_1, V_2, W)$ that is monotonically decreasing in $D = |V_1 - V_2|$ and goes for any W with $0 < W < 1$ through the point with $D = |V_1 - V_2| = \tau$ and $c(V_1, V_2, W) = W$ has tipping point τ . This is illustrated in Fig. 3 by the yellow areas. In particular, this applies to all example functions shown in Table 2 as Fig. 3 shows that they all are monotonically decreasing and go through the point $(0.2, 0.7)$ for which $D = 0.2$ and $c(V_1, V_2, W) = 0.7$.

Note that condition b) of this Definition 1 is not fulfilled when $|V_1 - V_2| < \tau$ and $c(V_1, V_2, 0) = 0$ or when $|V_1 - V_2| > \tau$ and $c(V_1, V_2, 1) = 1$. For example, for **slhomo** $_{\tau,\alpha}(V_1, V_2, W)$ and **sqhomo** $_{\tau,\alpha}(V_1, V_2, W)$ these cases do occur as always **slhomo** $_{\tau,\alpha}(V_1, V_2, 0) = 0$ and **slhomo** $_{\tau,\alpha}(V_1, V_2, 1) = 1$ due to the factor $W(1-W)$, and the same for **sqhomo** $_{\tau,\alpha}(V_1, V_2, W)$. Therefore they do have a tipping point but they do not have a strict tipping point. In the third paragraph in Section 8 it is discussed how in a practical application having a tipping point but not having a strict tipping point is not a big problem, but some care is needed in not assigning initial values 0 or 1 to connection weights, as these values will never change in such cases. Note that when a function with strict tipping point τ is used, any two nodes X, Y with $\omega_{X,Y}(t) = 0$ (which usually is interpreted as not being connected) can become connected over time: when $|X(t) - Y(t)| < \tau$, then by Definition 1b)(i) it holds $\omega_{X,Y}(t') > 0$ for $t' > t$.

Some examples of combination functions having a strict tipping point τ have been explored in more detail in (Sharpanskykh and Treur 2014) and (Boomgaard et al. 2018). In the former the advanced quadratic combination function **aqhomo** $_{\tau,\sigma}(V_1, V_2, W)$ and in the latter the advanced logistic function **alog2homo** $_{\tau,\sigma,\alpha}(V_1, V_2, W)$ has been explored; see also (Treur 2016a, b), Ch 11, p. 309:

$$\mathbf{aqhomo}_{\tau,\sigma}(V_1, V_2, W) = W + \text{Pos}(\alpha(\tau^2 - (V_1 - V_2)^2))(1-W) - \text{Pos}(-\alpha(\tau^2 - (V_1 - V_2)^2))W$$

$$\mathbf{alog2homo}_{\tau,\sigma,\alpha}(V_1, V_2, W) = W +$$

$$\text{Pos}\left(\alpha\left(0.5 - \frac{1-W}{1 + e^{-\alpha(|V_1 - V_2| - \tau)}}\right)\right) - \text{Pos}\left(-\alpha\left(0.5 - \frac{W}{1 + e^{-\alpha(|V_1 - V_2| - \tau)}}\right)\right)$$

Here $\text{Pos}(x) = (|x| + x)/2$. The following lemma shows some properties of this operator.

Lemma 1

- a) $\text{Pos}(x) = x$ when x is positive, else $\text{Pos}(x) = 0$.

So always $\text{Pos}(x) \geq 0$, and when $\text{Pos}(x) > 0$, then $\text{Pos}(-x) = 0$.

- b) For any numbers α and β the following are equivalent:
 - (i) $\alpha \text{Pos}(x) + \beta \text{Pos}(-x) = 0$
 - (ii) $\alpha \text{Pos}(x) = 0$ and $\beta \text{Pos}(-x) = 0$
 - (iii) $x = 0$ or $x > 0$ and $\alpha = 0$ or $x < 0$ and $\beta = 0$

This format using the double $\text{Pos}(\cdot)$ function can often be used to create a combination function satisfying the strict tipping point condition as a variation on a combination function satisfying only the tipping point condition. This is shown in the following proposition that assumes any close distance function denoted by $d(\tau, D)$. More specifically, the following proposition makes it quite easy to obtain functions with tipping point τ or with strict tipping point τ .

Proposition 1

Suppose for any function $d(\tau, D)$ it holds

$$\begin{aligned} d(\tau, D) > 0 & \text{ iff } D < \tau \\ d(\tau, D) < 0 & \text{ iff } D > \tau \end{aligned}$$

Then the following hold:

- a) For any $\alpha > 0$ the function.

$$c(V_1, V_2, W) = W + \alpha W(1-W) d(\tau, |V_1 - V_2|)$$

satisfies the tipping point condition, but not strict

- b) For any $\alpha > 0$ the function.

$$c'(V_1, V_2, W) = W + \alpha \text{Pos}(d(\tau, |V_1 - V_2|))(1-W) - \alpha \text{Pos}(-d(\tau, |V_1 - V_2|))W$$

satisfies the strict tipping point condition.

Proof

- a) The proof is mainly based on some algebraic rewriting.

Here as a first step it has to be proven that for any W with $0 < W < 1$ and all V_1, V_2 it holds.

$$|V_1 - V_2| < \tau \Leftrightarrow c(V_1, V_2, W) > W$$

This follows from

$$|V_1 - V_2| < \tau \Leftrightarrow d(\tau, |V_1 - V_2|) > 0 \Leftrightarrow \alpha W(1-W) d(\tau, |V_1 - V_2|) > 0 \Leftrightarrow c(V_1, V_2, W) > W$$

Similarly the other two cases for $|V_1 - V_2| > \tau$ and $|V_1 - V_2| = \tau$ can be verified.

From $c(V_1, V_2, 0) = 0$ for all V_1, V_2 it follows that the strict tipping point requirement is not fulfilled.

b) Also this proof is mainly based on some algebraic rewriting, thereby using Lemma 1.

First, $c'(V_1, V_2, W)$ satisfies the tipping point condition; for any W with $0 < W < 1$ and all V_1, V_2 by Lemma 1 it holds:

$$|V_1 - V_2| < \tau \Leftrightarrow d(\tau, |V_1 - V_2|) > 0 \Leftrightarrow \text{Pos}(d(\tau, |V_1 - V_2|)) > 0 \text{ and } \text{Pos}(-d(\tau, |V_1 - V_2|)) = 0 \\ \Leftrightarrow c'(V_1, V_2, W) > W$$

Similarly the other two conditions.

It is strict because

$$|V_1 - V_2| < \tau \Rightarrow d(\tau, |V_1 - V_2|) > 0 \Rightarrow c'(V_1, V_2, 0) = \alpha \text{Pos}(d(\tau, |V_1 - V_2|)) > 0$$

and

$$|V_1 - V_2| > \tau \Rightarrow d(\tau, |V_1 - V_2|) < 0 \Rightarrow c'(V_1, V_2, 1) = 1 - \alpha \text{Pos}(-d(\tau, |V_1 - V_2|)) < 1$$

Proposition 1 can easily be applied to simple linear or quadratic functions $d(\tau, D)$ such as:

$$d(\tau, D) = \tau - D \\ d(\tau, D) = \tau^2 - D^2$$

but also to functions such as

$$d(\tau, D) = 0.5 - \frac{1}{1 + e^{-\sigma(D-\tau)}}$$

as is shown in Proposition 2 b) and c).

The following proposition shows for four cases how it can be proven that some homophily combination function satisfies the tipping point or strict tipping point conditions.

Proposition 2

- a) **log1hom** $_{\tau,\alpha}(V_1, V_2, W)$ has tipping point τ , and is not strict
- b) **slog2hom** $_{\tau,\alpha}(V_1, V_2, W)$ has tipping point τ , and is not strict
- c) **alog2hom** $_{\tau,\alpha}(V_1, V_2, W)$ has a strict tipping point τ
- d) **exphomo** $_{\tau,\sigma}(V_1, V_2, W)$ has a tipping point τ and is not strict

Proof

a) This is based on some algebraic rewriting.

For **log1hom** $_{\tau,\alpha}(V_1, V_2, W)$ suppose $0 < W < 1$, then for $D = |V_1 - V_2|$ it holds

$$\begin{aligned}
 D < \tau &\Leftrightarrow e^{\sigma(D-\tau)} < 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} < 1-W \\
 &\Leftrightarrow W + (1-W) \frac{e^{\sigma(D-\tau)}}{W} < W + 1-W = 1 \\
 &\Leftrightarrow \frac{W + (1-W) e^{\sigma(D-\tau)}}{W} > W
 \end{aligned}$$

Similarly the other conditions can be verified:

$$\begin{aligned}
 D = \tau &\Leftrightarrow e^{\sigma(D-\tau)} = 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} = 1-W \\
 &\Leftrightarrow W + (1-W) \frac{e^{\sigma(D-\tau)}}{W} = W + 1-W = 1 \\
 &\Leftrightarrow \frac{W + (1-W) e^{\sigma(D-\tau)}}{W} = W
 \end{aligned}$$

$$\begin{aligned}
 D > \tau &\Leftrightarrow e^{\sigma(D-\tau)} > 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} > 1-W \\
 &\Leftrightarrow W + (1-W) \frac{e^{\sigma(D-\tau)}}{W} > W + 1-W = 1 \\
 &\Leftrightarrow \frac{W + (1-W) e^{\sigma(D-\tau)}}{W} < W
 \end{aligned}$$

To verify that it has no strict tipping point τ : for any V_1, V_2 it holds

$$\mathbf{log1\ hom}_{\tau,\alpha}(V_1, V_2, 0) = 0$$

- b) This proof is based on Proposition 1. Applying Proposition 1a) to $\mathbf{slog2hom}_{\tau,\alpha}(V_1, V_2, W)$ consider

$$d(\tau, D) = 0.5 - \frac{1}{1 + e^{-\sigma(D-\tau)}}$$

This function indeed satisfies the conditions of Proposition 1; therefore from Proposition 1a) it follows that it has tipping point τ but not strict tipping point τ .

- c) Also this proof is based on Proposition 1. For $\mathbf{alog2hom}_{\tau,\alpha}(V_1, V_2, W)$ it follows from Proposition 1b) with the same $d(\tau, D)$ as above that it has strict tipping point τ .
- d) This is based on some algebraic rewriting.

For W with $0 < W < 1$ it holds

$$\begin{aligned}
 D < \tau &\Leftrightarrow e^{\sigma(D-\tau)} < 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} < (1-W) \\
 &\Leftrightarrow 1 - (1-W) e^{\sigma(D-\tau)} > 1 - (1-W) = W \Leftrightarrow \mathbf{exphomo}_{\tau,\sigma}(V_1, V_2, W) > W
 \end{aligned}$$

$$\begin{aligned}
 D = \tau &\Leftrightarrow e^{\sigma(D-\tau)} = 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} = (1-W) \\
 &\Leftrightarrow 1 - (1-W) e^{\sigma(D-\tau)} = 1 - (1-W) = W \Leftrightarrow \mathbf{exphomo}_{\tau,\sigma}(V_1, V_2, W) = W
 \end{aligned}$$

$$\begin{aligned}
 D > \tau &\Leftrightarrow e^{\sigma(D-\tau)} > 1 \Leftrightarrow (1-W) e^{\sigma(D-\tau)} > (1-W) \\
 &\Leftrightarrow 1 - (1-W) e^{\sigma(D-\tau)} < 1 - (1-W) = W \Leftrightarrow \mathbf{exphomo}_{\tau,\sigma}(V_1, V_2, W) < W
 \end{aligned}$$

This shows it has tipping point τ . It has no strict tipping point, as

$$\mathbf{exphomo}_{\tau,\sigma}(V_1, V_2, 1) = 1 \text{ for all } V_1, V_2.$$

The following proposition shows that weighted averages of functions with tipping point τ also have a tipping point τ , and the same for having a strict tipping point.

Proposition 3

A weighted average (with positive weights) of homophily combination functions with tipping point τ also has tipping point τ , and with strict tipping point τ , also has strict tipping point τ .

Proof

This can be verified in a straightforward manner.

Suppose

$$c(V_1, V_2, W) = \gamma_1 c_1(V_1, V_2, W) + \dots + \gamma_m c_m(V_1, V_2, W)$$

with $\gamma_1 + \dots + \gamma_m = 1$. Suppose $0 < W < 1$. Then $|V_1 - V_2| < \tau \Rightarrow c_i(V_1, V_2, W) > W$ for all i , and therefore

$$c(V_1, V_2, W) > \gamma_1 W + \dots + \gamma_m W = W$$

Similarly

$$\begin{aligned} |V_1 - V_2| = \tau &\Rightarrow c(V_1, V_2, W) = W \\ |V_1 - V_2| > \tau &\Rightarrow c(V_1, V_2, W) < W \end{aligned}$$

Now suppose $c(V_1, V_2, W) > W$, then $|V_1 - V_2| = \tau$ or $|V_1 - V_2| > \tau$ cannot hold as they imply $c(V_1, V_2, W) = W$ or $c(V_1, V_2, W) < W$, therefore $|V_1 - V_2| < \tau$. The same for the other clauses. Moreover, suppose that the functions $c_i(V_1, V_2, W)$ all have strict tipping point τ . Then

If $|V_1 - V_2| < \tau$ then

$$c(V_1, V_2, 0) = \gamma_1 c_1(V_1, V_2, 0) + \dots + \gamma_m c_m(V_1, V_2, 0) > 0$$

If $|V_1 - V_2| > \tau$ then

$$c(V_1, V_2, 1) = \gamma_1 c_1(V_1, V_2, 1) + \dots + \gamma_m c_m(V_1, V_2, 1) < \gamma_1 + \dots + \gamma_m = 1$$

Therefore also $c(V_1, V_2, W)$ has a strict tipping point.

When more characteristics on network structure and adaptation combination functions are assumed, more refined results can be found, as will be shown in Section 6. Next, consider the combination functions used for social contagion of the states.

Definition 2 (Properties of combination functions for social contagion)

(a) A function $c(\cdot)$ is called *monotonically increasing* if

$$U_i \leq V_i \text{ for all } i \Rightarrow c(U_1, \dots, U_k) \leq c(V_1, \dots, V_k)$$

(b) A function $c(\cdot)$ is called *strictly monotonically increasing* if

$$U_i \leq V_i \text{ for all } i, \text{ and } U_j < V_j \text{ for at least one } j \Rightarrow c(U_1, \dots, U_k) < c(V_1, \dots, V_k)$$

(c) A function $c(\cdot)$ is called *scalar-free* if $c(\alpha V_1, \dots, \alpha V_k) = \alpha c(V_1, \dots, V_k)$ for all $\alpha > 0$

Definition 3 (Normalised network)

A network is *normalised* or uses normalised combination functions if for each state Y it holds $c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y}) = 1$, where X_1, \dots, X_k are the states with outgoing connections to Y .

Note that $c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y})$ is an expression in terms of the parameter(s) of the combination function and $\omega_{X_1,Y}, \dots, \omega_{X_k,Y}$. To require this expression to be equal to 1 provides a constraint on these parameters: an equation relating the parameter value(s) of the combination functions to the parameters $\omega_{X_1,Y}, \dots, \omega_{X_k,Y}$. To satisfy this property, often the parameter(s) can be given suitable values. For example, for a Euclidean combination function $\lambda_Y = \omega_{X_1,Y}^n + \dots + \omega_{X_k,Y}^n$ will provide a normalised network. This can be done in general:

- (1) **normalisation by adjusting the combination functions.** If any combination function $c_Y(\dots)$ is replaced by $c'_Y(\dots)$ defined as

$$c'_Y(V_1, \dots, V_k) = c_Y(V_1, \dots, V_k) / c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y})$$

then the network becomes normalised: indeed $c'_A(\omega_{X_1,Y}, \dots, \omega_{X_k,Y}) = 1$

- (2) **normalisation by adjusting the connection weights (for scalar-free combination functions).** For scalar-free combination functions also normalisation is possible by adapting the connection weights; define:

$$\omega'_{X_i,Y} = \omega_{X_i,Y} / c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y})$$

Then the network becomes normalised; indeed it holds:

$$c_Y(\omega'_{X_1,Y}, \dots, \omega'_{X_k,Y}) = c(\omega_{X_1,Y} / c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y}), \dots, \omega_{X_k,Y} / c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y})) = 1$$

Definition 4 (Symmetric network and symmetric combination function)

- a) A network is called *weakly symmetric* if for all states X, Y it holds $\omega_{X,Y} > 0$ and $\omega_{Y,X} > 0$. It is fully symmetric if for all states X, Y it holds $\omega_{X,Y} = \omega_{Y,X}$.
- b) The homophily combination function $c_{\omega_{X,Y}}(V_1, V_2, W)$ is called *symmetric* if $c_{\omega_{X,Y}}(V_1, V_2, W) = c_{\omega_{Y,X}}(V_2, V_1, W)$.

If the combination function $c_{\omega_{X,Y}}(V_1, V_2, W)$ is symmetric, the network is fully symmetric if the initial values for $\omega_{X,Y}$ and $\omega_{Y,X}$ are equal.

Proposition 4

- a) When the homophily function $c(V_1, V_2, W)$ is symmetric, and initially the network is fully symmetric, then the network is continually fully symmetric.
- b) For every $n > 0$ a Euclidean combination function of n^{th} degree is strictly monotonically increasing, scalar-free, and symmetric

Proof

- a) This follows from the fact that in this case the difference equation for $\omega_{X,Y}$ is symmetric in X and Y
- b) A Euclidean combination function is composed of strictly monotonic functions as each function $V_i \rightarrow V_i^n$ is monotonic for positive n and positive values V_i , and so are $W \rightarrow W/\lambda$ and $W \rightarrow W^{1/n}$. From

$$\begin{aligned}
 \mathbf{eucl}_{n,\lambda}(\alpha V_1, \dots, \alpha V_k) &= \sqrt[n]{\frac{(\alpha V_1)^n + \dots + (\alpha V_k)^n}{\lambda}} \\
 &= \sqrt[n]{\frac{\alpha^n V_1^n + \dots + \alpha^n V_k^n}{\lambda}} \\
 &= \alpha \sqrt[n]{\frac{V_1^n + \dots + V_k^n}{\lambda}} \\
 &= \alpha \mathbf{eucl}_{n,\lambda}(V_1, \dots, V_k)
 \end{aligned}$$

it follows that it is scalar-free. The rest directly follows.

Relating adaptive network structure to bonding behaviour

In this section, an analysis is presented of the asymptotic behaviour of an adaptive network based on the coevolution of social contagion and bonding by homophily. Here it is found out how the network and adaptation properties defined in Section 5 (such as a tipping point τ), affect the emerging behaviour of the adaptive network. In particular, it is addressed how these properties imply that over time the network ends up in certain states (for both the nodes and the connections) that represent community formation. It will turn out that under certain conditions in terms of the properties defined in Section 5, in the end, a number of communities emerge such that:

- Nodes within one community have similar state values
- Connections between nodes within a community are very strong
- Connections between nodes from different communities are very weak

Note that these phenomena were already observed in the simulation examples in Section 4, but now it will be proved why under certain conditions they always have to occur. To formalise this, the following general notions are important; e.g., (Brauer and Nohel 1969; Hirsch 1984; Lotka 1956).

Definition 5 (Stationary point and equilibrium)

A state Y has a *stationary point* at t if $\mathbf{d}Y(t)/\mathbf{d}t = 0$, is increasing at t if $\mathbf{d}Y(t)/\mathbf{d}t > 0$, and is decreasing at t if $\mathbf{d}Y(t)/\mathbf{d}t < 0$, and similarly for adaptive connection weights ω . The network is in *equilibrium* at t if every state Y and every connection weight ω in the model has a stationary point at t . The equilibrium is *attracting* if any small perturbations of its values lead to convergence to the equilibrium values.

Considering the specific differential equation format for a temporal-causal network model shown in Section 2, and assuming nonzero speed factors the following more specific criteria for stationary points, and for increasing and decreasing trends, in terms of the combination functions and connection weights are easily found:

Lemma 2 (Criteria for a stationary, increasing and decreasing)

Let Y be a state and X_1, \dots, X_k the states with outgoing connections to state Y . Then

- (i) Y has a stationary point at $t \iff c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) = Y(t)$
- (ii) Y is strictly increasing at $t \iff c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) > Y(t)$
- (iii) Y is strictly decreasing at $t \iff c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) < Y(t)$

Similarly for connection weights ω .

Theorem 1 presents some of the results found for the relation between the emerging equilibrium values for states and for connection weights. It is shown how the distance of the equilibrium values of two states relates to the equilibrium value of their connections. Note that for now no specific assumption is made on the combination function for social contagion. To prove Theorem 1, the following lemma is a useful means. It shows that when the function $c(V_1, V_2, W)$ satisfies the tipping point criterion, a connection weight 0 can only be reached for states X and Y when $|X(t) - Y(t)| > \tau$ and a connection weight 1 can only be reached for states X and Y when $|X(t) - Y(t)| < \tau$.

Lemma 3 Suppose the function $c(V_1, V_2, W)$ has tipping point τ for V_1 and V_2 . Then

- (i) The value 0 for $\omega_{X,Y}$ can only be reached from $\omega_{X,Y}(t)$ with $0 < \omega_{X,Y}(t) < 1$ if $|X(t) - Y(t)| > \tau$
- (ii) The value 1 for $\omega_{X,Y}$ can only be reached from $\omega_{X,Y}(t)$ with $0 < \omega_{X,Y}(t) < 1$ if $|X(t) - Y(t)| < \tau$

Proof

- (i) The proof is by contraposition. Suppose $0 < \omega_{X,Y}(t) < 1$ holds and $|X(t) - Y(t)| > \tau$ does not hold. Then $|X(t) - Y(t)| \leq \tau$, and by Definition 1a)(i) and (ii) it follows that $c(V_1, V_2, W) \geq W$, and therefore from the difference equation it follows that $\omega_{X,Y}(t + \Delta t) \geq \omega_{X,Y}(t)$, so $\omega_{X,Y}(t)$ will not become lower and in particular will not reach 0.
- (ii) is similar using Definition 1a)(ii) and (iii).

This lemma already reveals that connection weights $\omega_{X,Y}$ converging to 0 have some relation to the state values of X and Y having distance more than τ , and connection weights $\omega_{X,Y}$ converging to 1 have some relation to the state values of X and Y having a distance less than τ . These relations between connection weights and state values during a convergence process are made more precise for reaching an equilibrium state in Theorem 1.

Theorem 1 (Relations between equilibrium values for states and for connection weights)

Suppose the function $c(V_1, V_2, W)$ has tipping point τ for V_1 and V_2 and an attracting equilibrium state is given with values \underline{X} for the states X and $\underline{\omega}_{X,Y}$ for the connection weights $\omega_{X,Y}$. Then the following hold:

- a) If $|\underline{X} - \underline{Y}| < \tau$, then the equilibrium value $\underline{\omega}_{X,Y}$ is 1; in particular this holds when $\underline{X} = \underline{Y}$. Therefore, if $\underline{\omega}_{X,Y} < 1$, then $|\underline{X} - \underline{Y}| \geq \tau$, and, in particular, $\underline{X} \neq \underline{Y}$.

- b) If $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| > \tau$, then the equilibrium value $\underline{\omega}_{X,Y}$ is 0. Therefore, if $\underline{\omega}_{X,Y} > 0$, then $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| \leq \tau$.
- c) $0 < \underline{\omega}_{X,Y} < 1$ implies $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| = \tau$.

Proof

- a) Suppose two states are given with equilibrium values $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ with distance less than τ : $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| < \tau$. Given this, from the equilibrium equation $\mathbf{c}_{\omega_{X,Y}}(\underline{\mathbf{X}}, \underline{\mathbf{Y}}, \underline{\omega}_{X,Y}) = \underline{\omega}_{X,Y}$, by Definition 1(ii) it follows that $0 < \underline{\omega}_{X,Y} < 1$ cannot be true, and therefore $\underline{\omega}_{X,Y} = 0$ or $\underline{\omega}_{X,Y} = 1$. By Lemma 3(i) and the equilibrium being attracting it follows that $\underline{\omega}_{X,Y} = 0$ can be excluded, so $\underline{\omega}_{X,Y} = 1$.

The other statement, that if $\underline{\omega}_{X,Y} < 1$, then $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| \geq \tau$, follows by logical contraposition.

- b) For $|\underline{\mathbf{X}}-\underline{\mathbf{Y}}| > \tau$ this is similar, using Definition 1(iii). The last statement follows from the contraposition of the previous one.
- c) This immediately follows from a) and b).

This Theorem 1 explains some of the observations made in Section 4, in particular, that in all of these simulations the connection weights all end up either in value 0 (the state equilibrium values differ at least τ) or in value 1 (the state equilibrium values are equal).

When also some assumptions are made for social contagion, for example, as explored in (Treur, 2018b) more refined results can be found. The following is a basic Lemma for normalised networks with social contagion combination functions that are monotonically increasing and scalar-free.

Lemma 4

Let a normalised network with nonnegative connections be given with combination functions that are monotonically increasing and scalar-free; then the following hold:

- a) (i) If for some state Y at time t for all nodes X with $\omega_{X,Y} > 0$ it holds $X(t) \leq Y(t)$, then $Y(t)$ is decreasing at t : $\mathbf{d}Y(t)/\mathbf{d}t \leq 0$.
- (ii) If, moreover, the combination function is strictly increasing and a state X exists with $X(t) < Y(t)$ and $\omega_{X,Y} > 0$, then $Y(t)$ is strictly decreasing at t : $\mathbf{d}Y(t)/\mathbf{d}t < 0$.
- b) (i) If for some state Y at time t for all nodes X with $\omega_{X,Y} > 0$ it holds $X(t) \geq Y(t)$, then $Y(t)$ is increasing at t : $\mathbf{d}Y(t)/\mathbf{d}t \geq 0$.
- (ii) If, moreover, the combination function is strictly increasing and a state X exists with $X(t) > Y(t)$ and $\omega_{X,Y} > 0$, then $Y(t)$ is strictly increasing at t : $\mathbf{d}Y(t)/\mathbf{d}t > 0$.

Proof

The proofs for a) and b) are similar. Therefore only the proof for a) is given.

- a) (i) This proof shows that lower values of the states with incoming connections can never increase the state value of state Y . More specifically, assume for all states X_i with (positive) outgoing connections to Y it holds $X_i(t) \leq Y(t)$. Therefore

$$\begin{aligned}
 c_Y(\omega_{X_1,Y}X_1(t), \dots, \omega_{X_k,Y}X_k(t)) &\leq c_Y(\omega_{X_1,Y}Y(t), \dots, \omega_{X_k,Y}Y(t)) \\
 &= c_Y(\omega_{X_1,Y}, \dots, \omega_{X_k,Y})Y(t) \\
 &= Y(t)
 \end{aligned}$$

and by Lemma 2(i) and (iii) this implies $\mathbf{d}Y(t)/\mathbf{d}t \leq 0$.

(ii) This proof shows that a strictly lower values of one of the states with incoming connections will actually decrease the state value of state Y . More specifically, from being strictly monotonous it follows

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

and by Lemma 2(iii) this implies $\mathbf{d}Y(t)/\mathbf{d}t < 0$.

b) This can be proven in a similar manner.

Now the more specific theorem is obtained for the connection weights if the combination functions for social contagion for the states are assumed strictly monotonically increasing and scalar-free. It states that the connection weights in an attracting equilibrium are all 0 or 1, and when τ is a strict tipping point, that depends just on whether they connect states with the same or a different equilibrium state value.

Theorem 2 (Equilibrium values $\underline{\omega}_{X,Y}$ all 0 or 1)

Suppose the network is weakly symmetric and normalised, and the combination functions for the states are strictly monotonically increasing and scalar-free. Suppose that $c(V_1, V_2, W)$ has a tipping point τ . Then

- a) In an attracting equilibrium state for any states X, Y from $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$ it follows $\underline{\omega}_{X,Y} = 0$.
- b) In an attracting equilibrium state for any states X, Y with $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ it holds $\underline{\omega}_{X,Y} = 0$ or $\underline{\omega}_{X,Y} = 1$.
- c) If $c(V_1, V_2, W)$ has a strict tipping point τ , then in an equilibrium state for any X, Y with $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ it holds $\underline{\omega}_{X,Y} = 1$.

Proof

- a) The proof goes by reductio ad absurdum (reduce to absurdity); it shows that the opposite of the claimed statement cannot be true by deriving a contradiction from this opposite statement. So, suppose (oppositely) that in an attracting equilibrium state, states X and Y exist such that $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$ and $\underline{\omega}_{X,Y}, \underline{\omega}_{Y,X} > 0$. Take a state X with this property with highest value $\underline{\mathbf{X}}$. Then for all states Z with $\underline{\mathbf{Z}} > \underline{\mathbf{X}}$ it holds $\underline{\omega}_{X,Z} = \underline{\omega}_{Z,X} = 0$. Therefore all states X_i with a nonzero (positive) outgoing connection weight to state X satisfy $\underline{\mathbf{X}}_i \leq \underline{\mathbf{X}}$. Moreover, one of these X_i is state Y with $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$, so, as X has the highest value, it holds $\underline{\mathbf{Y}} < \underline{\mathbf{X}}$. Now apply Lemma 4a)(ii) to this state X . It follows that $\mathbf{d}X(t)/\mathbf{d}t < 0$; therefore $X(t)$ cannot be not in equilibrium. This contradicts the premise that the network is in equilibrium.

Therefore no nodes X and Y exist such that $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$ and $\underline{\omega}_{X,Y}, \underline{\omega}_{Y,X} > 0$. This implies that $\underline{\omega}_{X,Y} = 0$ and $\underline{\omega}_{Y,X} = 0$ for all nodes X and Y with $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$.

- b) Also this proof goes by reductio ad absurdum (reduce to absurdity); also here it is shown that the opposite of the claimed statement cannot be true by deriving a contradiction from this opposite statement. So, suppose (oppositely) $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ and $0 < \omega_{X,Y}(t) < 1$. Then by Definition 1a)(i) from $X(t) = Y(t)$ it follows that $c_{\omega_{X,Y}}(X(t), Y(t), \omega_{X,Y}(t)) > \omega_{X,Y}(t)$. From this by Lemma 2(ii) it follows that $d\omega_{X,Y}(t)/dt > 0$: $\omega_{X,Y}(t)$ is strictly increasing and is not in equilibrium. This contradicts the premise that the network is in equilibrium. Therefore in the equilibrium state when $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ it holds $\underline{\omega}_{X,Y} = 0$ or $\underline{\omega}_{X,Y} = 1$.
- c) From b) it is already known that $\underline{\omega}_{X,Y} = 0$ or $\underline{\omega}_{X,Y} = 1$. The former option $\underline{\omega}_{X,Y} = 0$ has to be excluded now. Also this goes by reductio ad absurdum (reduce to absurdity); also here it is shown that the option $\underline{\omega}_{X,Y} = 0$ cannot be true by deriving a contradiction when this option is assumed. So, suppose $\underline{\omega}_{X,Y} = 0$. If $c(V_1, V_2, W)$ is strict, and $|V_1 - V_2| < \tau$ then by Definition 1b)(i) it holds $c(V_1, V_2, 0) > 0$, so by Lemma 1(ii) it follows that when $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$, the value $\underline{\omega}_{X,Y} = 0$ cannot be an equilibrium value, which contradicts the premise that the network is in equilibrium. Therefore in an equilibrium for any X, Y with $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ it holds $\underline{\omega}_{X,Y} = 1$.

The following theorem shows the implications for emerging communities or clusters and distances between different equilibrium values for nodes.

Theorem 3 (Partition and equilibrium values of nodes)

Suppose the network is weakly symmetric and normalised, the combination function for the states is strictly monotonically increasing and scalar-free, and the combination function for the connections uses tipping point τ and is strict and symmetric. Then in any attracting equilibrium state a partition of the set of states into disjoint subsets C_1, \dots, C_p occurs such that:

- (i) For each C_i the equilibrium values for all the states in C_i are equal: $\underline{\mathbf{X}} = \underline{\mathbf{Y}}$ for all $X, Y \in C_i$.
- (ii) Every C_i forms a fully connected network with weights 1: $\underline{\omega}_{X,Y} = 1$ for all $X, Y \in C_i$.
- (iii) Every two nodes in different C_i have connection weight 0: when $i \neq j$, then $X \in C_i$ and $Y \in C_j$ implies $\underline{\omega}_{X,Y} = 0$.
- (iv) Any two distinct equilibrium values of states $\underline{\mathbf{X}} \neq \underline{\mathbf{Y}}$ have distance $\geq \tau$. Therefore there are at most $p \leq 1 + 1/\tau$ communities C_i and equilibrium values $\underline{\mathbf{X}}$.

Proof

Suppose in the equilibrium there are p distinct state values V_1, \dots, V_p ; then define the sets

$$C_i = \{ X \mid \underline{\mathbf{X}} = V_i \}$$

It can easily be verified in a straightforward manner that these sets fulfill what is claimed:

- (i) By definition all state values in one C_i are equal

- (ii) From Theorem 2c) it follows that all states with equal values have connections 1, therefore any C_i is fully connected.
- (iii) This follows from Theorem 2a).
- (iv) Suppose for some X, Y it holds $|\underline{X} - \underline{Y}| < \tau$. Then by Definition 1(i) it follows $c_{\omega_{X,Y}}(V_1, V_2, 0) > 0$. Therefore 0 cannot be the equilibrium value $\underline{\omega}_{X,Y}$; from Theorem 2a) it follows that $\underline{X} = \underline{Y}$, and therefore X and Y are in one C_i . This implies that the state values in different C_i have distance $\geq \tau$.

Characterising the attracting equilibrium states in terms of strongly connected components

To analyse connectivity, within Graph Theory the notion of a strongly connected component has been identified. The main parts of the following definitions can be found, for example, in (Harary et al. 1965), Ch. 3, or (Kuich 1970), Section 6. Note that in the current paper only nonnegative connection weights are considered.

Definition 6 (reachability and strongly connected components)

- (a) State Y is *forward reachable* from state X if there is a directed path from X to Y with nonzero connection weights and speed factors.
- (b) A network N is *strongly connected* if every two states are mutually forward reachable within N .
- (c) A state is called *independent* if it is not forward reachable by any other state.
- (d) A *subnetwork* of a network N is a network whose states and connections are states and connections of N .
- (e) A *strongly connected component* C of a network N is a strongly connected subnetwork of N such that it is maximal: no larger strongly connected subnetwork of N contains it as a subnetwork.

Strongly connected components C can be identified by choosing any node X of N and adding all nodes that are on any cycle through X . Note also that when a node X is not on any cycle, then it will form a singleton strongly connected component C by itself; this applies in particular to all nodes of N with indegree or outdegree zero. There are efficient algorithms available to determine the strongly connected components of a network or graph; for example, see (Bloem et al. 2006; Chen 2009; Fleischer et al. 2000; Gentilini et al. 2003; Łacki 2013; Li et al. 2014; Tarjan 1972; Wijs et al. 2016).

The strongly connected components form a partition of the nodes of the graph or network. By this partition a more abstract, simpler view of the network can be created, called condensation graph, in which each component becomes one abstract node; this is defined as follows.

Definition 7 (condensation graph)

The *condensation graph* $C(N)$ of a network N with respect to its strongly connected components is a graph whose nodes are the strongly connected components of N and whose connections are determined as follows: there is a connection from node C_i to node C_j in $C(N)$ if and only if in N there is at least one connection from a node in the strongly connected component C_i to a node in the strongly connected component C_j .

An important result is that a condensation graph $C(N)$ is always an acyclic graph. In a sense, all cycles are locked up inside the components, and the connections between these components do not contain any cycles. The following theorem summarizes this; it was adopted from (Harary et al. 1965), Ch. 3, Theorems 3.6 and 3.8, or (Kuich 1970), Section 6.

Theorem 4 (Acyclic condensation graph)

- a) For any network N its condensation graph $C(N)$ is acyclic, and has at least one state of outdegree zero and at least one state of indegree zero.
- b) The network N is acyclic itself if and only if it is graph-isomorphic to $C(N)$. In this case, the nodes in $C(N)$ all are singleton sets $\{X\}$ containing one state X from N .
- c) The network N is strongly connected itself if and only if $C(N)$ only has one node; this node is the set of all states of N .

Note that in an adaptive network, the strongly connected components and the condensation graph usually change over time. For example, connection weights that were 0 can become nonzero, or nonzero connection weights can converge to 0. From Theorem 3 it immediately follows that in the equilibrium state the C_i defined there are the strongly connected components of the network. Then the results from Theorem 3 can be rephrased in the above terms in the following way.

Theorem 5 (Strongly connected components in an attracting equilibrium)

Suppose the network is weakly symmetric and normalised, the combination functions for social contagion between the nodes is strictly monotonically increasing and scalar-free, and the homophily combination function for the connections uses tipping point τ and is strict and symmetric. Then in any attracting equilibrium state the following hold:

- (i) There are at most $p \leq 1 + 1/\tau$ strongly connected components.
- (ii) Each strongly connected component is fully connected and all states in it have a common equilibrium state value.
- (iii) There are no nonzero connections between states from different strongly connected components, and the equilibrium values of these states have distance $\geq \tau$.
- (iv) The condensation graph $C(N)$ is totally disconnected: it has no connections at all.

The following converse holds as well; this shows how equilibrium states can be characterised by specific properties of the strongly connected components.

Theorem 6 (Strongly connected components characterisation)

Suppose the network is weakly symmetric, the combination function for social contagion between the nodes is strictly monotonically increasing, normalised and scalar-free, and the homophily combination function for the connections uses tipping point τ and is strict and symmetric. Suppose at some time point t the following hold:

- (i) Each strongly connected component C is fully connected and all states in C have a common state value.

- (ii) All connections between states from different strongly connected components have weight 0 and the equilibrium values of these states have distance $> \tau$.

Then the network is in an equilibrium state.

Proof

First in a) it is proven that the state values are stationary; next, in b) it is proven that the connection weights are stationary. Having both stationary, the network is in equilibrium.

- a) Consider within any component C any state Y which has only nonzero incoming connections from states X_1, \dots, X_k . Due to (ii) these necessarily belong to the same component C . As within C the state values are equal to one value V and each connection has weight 1 the following holds

$$\begin{aligned} \text{aggimpact}_Y(t) &= c_Y(X_1(t), \dots, X_k(t)) \\ &= c_Y(V, \dots, V) \\ &= V \cdot c_Y(1, \dots, 1) \\ &= V \\ &= Y(t) \end{aligned}$$

and therefore by Lemma 2(i) it holds $dY(t)/dt = 0$, so $Y(t)$ is stationary.

- b) Next, it is proven that the connection weights are stationary. Consider the connection weight $\omega_{X,Y}$ for the connection from states X to Y in the same component C . Suppose as a perturbation from 1 it holds $\omega_{X,Y}(t) < 1$. Given that $|X(t) - Y(t)| < \tau$, from Definition 1(i) it follows that $c_{\omega_{X,Y}}(X(t), Y(t), \omega_{X,Y}(t)) > \omega_{X,Y}(t)$, and therefore $d\omega_{X,Y}(t)/dt > 0$, so it would move upward to 1. Therefore $\omega_{X,Y}$ is stationary. A similar argument for states X and Y in different components shows that $\omega_{X,Y}$ would move downward to 0, and therefore is stationary.

As both the states and the connection weights have been proven stationary, it has been found that the network is in equilibrium.

Note that the situation as characterised is a quite specific situation with trivial condensation graph, if compared, for example, to a general situation as addressed in (Treur 2018d), where usually nontrivial condensation graphs occur.

Overview of simulation experiments

In recent years for a number of the functions from Table 2 simulation experiments have been performed for real-world domains and related to empirical data. An overview of them is shown in Table 4. Some of the experiences are the following.

In general, it was possible to get similar patterns as in the empirical data by using dedicated Parameter Tuning methods such as Simulated Annealing. As may be expected the remaining Root Mean Square error differed with the different studies and domains, and also depends on how fine-grained the scoring scales for the data were; for example when a 3 points scale was used for empirical data, already because of that in the $[0, 1]$ interval a variation of 0.15 should be expected within the empirical data

Table 4 Overview of simulation experiments for different combination functions

Combination functions		Domain and data	Reference and venue
Homophily	Contagion		
slhomo _{σ,τ} (..)	ssum _λ (..)	Friendships in a classroom Glasgow Empirical Data 2016	(Blankendaal et al. 2016) ECAI'16
		Social media: blogger appreciation Instagram data	(Kozyreva et al. 2018) SocInfo'18
		Social media: music appreciation Twitter data	(van Gerwen et al. 2019) DCAI'18
	ssum _λ (..) alogistic _{σ,τ} (..)	Segregation of queer community Questionnaire data	(Heijmans et al. 2019) ICICT'19
	alogistic _{σ,τ} (..)	Social media: Zwarte Piet debate Twitter data	(Roller et al. 2017) COMPLEXNETWORKS'17
sqhomo _{σ,τ} (..)	ssum _λ (..)	Segregation of immigrants Literature data	(Kappert et al. 2018) ICCC'18
	alogistic _{σ,τ} (..)	Friendships in a classroom Glasgow Empirical Data 2016	(van Beukel et al. 2017, van Beukel et al. 2019) PAAMS'17, Neurocomputing 2019
aqhomo _{σ,τ} (..)	ssum _λ (..)	Scalefree vs random networks Literature data	(Sharpanskykh and Treur 2013, 2014) ICCC'13, Neurocomputing 2014
log2homo _{σ,τ} (..)	ssum _λ (..)	Social media: physical activity Twitter data	(van Dijk and Treur 2018) ICCC'18
	alogistic _{σ,τ} (..)	Friendships in a classroom Knecht 2008	(Boomgaard et al. 2018) SocInfo'18

themselves, and for a 5 points scale 0.1. Then an average deviation between model and empirical data will be at least in that order of magnitude, not smaller, say 0.15 to 0.25. This indeed was usually the case in the examples discussed in Table 4.

Using the simple linear or simple quadratic functions **slhomo**_{σ,τ}(..) or **sqhomo**_{σ,τ}(..) the change of connection weights is very slow when the weights are close to 0 or 1 due to the factor $W(1 - W)$ in the function. The advantage of this slowing down effect is that the connection weight values stay within the [0, 1] interval in a natural manner. But the downside of this factor is that when they are exactly 0 or 1 they will even freeze and not be able to change anyway, and the same when 0 or 1 are used as initial values. This is because **slhomo**_{σ,τ}(..) or **sqhomo**_{σ,τ}(..) have a tipping point, but no strict tipping point; also see the remark after Definition 1. The same holds for **log2homo**_{σ,τ}(..). So, when these functions are used, it is better to initialise all connection weight values between 0.1 and 0.9 instead of in the full [0, 1] interval. This is different for the advanced linear and quadratic homophily functions **alhomo**_{σ,τ}(..) or **aqhomo**_{σ,τ}(..). For them approaching 0 or 1 is still slow, but leaving 0 or 1 can be fast: they do have a strict tipping point. Note, however, that this is not necessarily always a good property. Maybe when a very strong connection has been formed, even some differences that occur may not affect the connection immediately. So, in some cases, or for some types of persons the slow change close to 0 or 1 as shown by **slhomo**_{σ,τ}(..) or **sqhomo**_{σ,τ}(..) may even be more plausible.

Note that in (Kozyreva et al. 2018) a slightly different multicriteria variant of the simple linear homophily function was used that takes into account multiple states in the similarity measure: instead of $|V_1 - V_2|$ for 1 criterion, for k criteria the following Euclidean distance formula is used to measure similarity:

$$\sqrt{(V_{1,1} - V_{1,2})^2 + \dots + (V_{k,1} - V_{k,2})^2}$$

In (Blankendaal et al. 2016) and (van Beukel et al. 2017) for bonding not only a homophily principle was used, but also a ‘more becomes more’ principle which can be considered a variant of what sometimes is called ‘the rich get richer’ (Simon 1955; Bornholdt and Ebel 2001), ‘cumulative advantage’ (Price 1976), ‘the Matthew effect’ (Merton 1968), or ‘preferential attachment’ (Barabási and Albert 1999; Newman 2003). For this ‘more becomes more’ principle, in (Blankendaal et al. 2016) a scaled sum combination function was used for the weights of the connections to a given other person Y with scaling factor the number k of such connections (resulting in the average of these connection weights), and in (van Beukel et al. 2017; van Beukel et al. 2019) and advanced logistic sum:

$$\begin{aligned} & \mathbf{ssum}_k(W_1, \dots, W_k) \\ & \mathbf{alogistic}_{\sigma, \tau}(W_1, \dots, W_k) \end{aligned}$$

where W_1, \dots, W_k refer to the weights $\omega_{X_1, Y}, \dots, \omega_{X_k, Y}$ of all connections of the other person Y . Note that the first scaled sum option $\mathbf{ssum}_k(\dots)$ adapts to the other person’s average connection weight independent of the number of connections, whereas the second $\mathbf{alogistic}_{\sigma, \tau}(\dots)$ option is more additive as more connections of the other person provide higher values. In both approaches, the two combination functions for both the homophily and the more becomes more principle were combined as a weighted sum.

Discussion

In this paper, it was analysed how community formation can be related to characteristics of the adaptive network’s structure, in particular in an adaptive network for bonding based on similarity. Here network structure includes the structure (and settings) of the adaptation principles incorporated. This has been addressed for adaptive social networks for bonding based on homophily (McPherson et al. 2001; Pearson et al. 2006). Relevant properties of the network and the adaptation principle have been identified, such as a tipping point for similarity. As one of the results, it has been found how the emergence of communities strongly depends on the value of this tipping point. It has been shown, for example, that some properties of the structure of the network and the adaptation principle entail that the connection weights all converge to 0 (for persons in different communities) or 1 (for persons within one community). As another main result, it has been found that the formation of communities strongly depends on the value τ of this tipping point; there can be no more communities than $1 + 1/\tau$ within the interval $[0, 1]$.

The presented results do not concern results for just one type of network or function, as more often is found. Instead, they were formulated and proven at a more general level and therefore can be applied not just to specific networks but to classes of networks satisfying the identified relevant properties of network structure and

adaptation characteristics. Note, however, that the focus in the current paper is on deterministic behaviour; therefore stochastic models such as, for example, the one reported in a nontechnical manner in (Axelrod 1997), are not covered by this analysis; e.g., see the description here:

For those who prefer symbolic statements, here is a complete description of how the culture, c , at a site can change. Select a random site (s), a random neighbor of that site (n), and a random feature (f). Let $G(s,n)$ be the set of features, g , such that the cultural traits are unequal—that is, $c(s,g) \neq c(n,g)$. If $c(s,f) = c(n,f)$ and G is not empty, then select a random feature, g , in $G(s,n)$ and set $c(s,g)$ to $c(n,g)$. This implementation of the model takes advantage of that fact that the probability that a random feature, f , will have the same trait at two sites equals the cultural similarity between those two sites. (Axelrod 1997), p. 208, footnote 4

It may be an interesting research focus for the future to explore whether and how the analysis results found here have counterparts for stochastic network models. Besides, in (Axelrod 1997) also regional differences are addressed. In a more extensive application of the models discussed in the current paper, that may be an interesting ingredient to address as well.

The results found also have been related to the strongly connected components of the network (Harary et al. 1965), Ch. 3, or (Kuich 1970), Section 6). A characterisation in terms of properties of the strongly connected components and their relations was found for states of an adaptive network which are attracting equilibrium states.

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

Not applicable.

Authors' contributions

The author read and approved the final manuscript.

Ethics approval and consent to participate

I approve the ethical code.

Consent for publication

I give my consent for publication.

Competing interests

The author declares that he has no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 November 2018 Accepted: 8 April 2019

Published online: 25 June 2019

References

- Ashby WR (1960) Design for a brain. Chapman and Hall, London (second extended edition). First edition, 1952
- Axelrod R (1997) The dissemination of culture: a model with local convergence and global polarization. *J Confl Resolut* 41(2):203–226
- Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512
- Blankendaal R, Parinussa S, Treur J (2016) A temporal-causal modelling approach to integrated contagion and network change in social networks. In: Proc. of the 22nd European conference on artificial intelligence, ECAI'16, vol 285. IOS Press, Frontiers in Artificial Intelligence and Applications, Amsterdam. pp 1388–1396

- Bloem R, Gabow HN, Somenzi F (2006) An algorithm for strongly connected component analysis in $n \log n$ symbolic steps. *Form Meth Syst Des* 28:37–56
- Boomgaard G, Lavitt F, Treur J (2018) Computational analysis of social contagion and Homophily based on an adaptive social network model proceedings of the 10th international conference on social informatics, SocInfo'18. Lecture notes in computer science 11185. Springer Publishers, Cham. pp 86–101
- Bornholdt S, Ebel H (2001) World wide Webscaling exponent from Simon's 1955 model. *Phys Rev E* 64(2001):035104
- Brauer F, Nohel JA (1969) The qualitative theory of ordinary differential equations. W.A. Benjamin Inc., New York
- Chen Y (2009) General spanning trees and reachability query evaluation. In: Desai BC (ed) Proc. of the 2nd Canadian conference on computer science and software engineering, C3S2E'09. ACM Press, New York City. pp 243–252
- Fleischer LK, Hendrickson B, Pinar A (2000) On identifying strongly connected components in parallel. In: Rolim J (ed) Parallel and distributed processing. IPDPS 2000. Lecture notes in computer science, vol 1800. Springer, Cham. pp 505–511
- Gentilini R, Piazza C, Policriti A (2003) Computing strongly connected components in a linear number of symbolic steps. In: Proc. SODA'03, pp 573–582
- Glasgow Empirical Data (2016) https://www.stats.ox.ac.uk/~snijders/siena/Glasgow_data.htm
- Gross T, Sayama H (eds) (2009) Adaptive networks: theory, models and applications. Springer, Cham.
- Harary F, Norman RZ, Cartwright D (1965) Structural models: an introduction to the theory of directed graphs. Wiley, New York
- Heijmans P, van Stijn J, Treur J (2019) Modeling cultural segregation of the queer community through an adaptive social network model. In: Proceedings of the fourth international congress on information and communication technology, ICICT'19. Advances in intelligent systems and computing, Springer
- Hirsch MW (1984) The dynamical systems approach to differential equations. *Bull Am Math Soc* 11:1–64
- Holme P, Newman MEJ (2006) Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys Rev E* 74(5):056108
- Kappert C, Rus R, Treur J (2018) On the emergence of segregation in society: network-oriented analysis of the effect of evolving friendships. In: Nguyen NT, Pimenidis E, Khan Z, Trawinski B (eds) Computational collective intelligence: 10th international conference, ICCCI 2018, proceedings, vol 1. Springer Verlag, Cham. pp 178–191 Lecture notes in artificial intelligence, vol. 11055
- Knecht, A. (2008). Empirical data: collected by Andrea Knecht. https://www.stats.ox.ac.uk/~snijders/siena/siena_datasets.htm
- Kozyreva O, Pechina A, Treur J (2018) Network-oriented modeling of multi-criteria homophily and opinion dynamics in social media. In: Koltsova O, Ignatov DI, Staab S (eds) Social informatics: proceedings of the 10th international conference on social informatics, SocInfo'18, vol 1. Springer Verlag, Cham. pp 322–335 Lecture notes in AI, vol. 11185
- Kuich W (1970) On the entropy of context-free languages. *Inf Control* 16:173–200
- Kuipers BJ (1984) Commonsense reasoning about causality: deriving behavior from structure. *Artif Intell* 24:169–203
- Kuipers BJ, Kassirer JP (1983) How to discover a knowledge representation for causal reasoning by studying an expert physician. In: Proc. of the eight international joint conference on artificial intelligence, IJCAI'83. William Kaufman, Los Altos
- Łacki J (2013) Improved deterministic algorithms for decremental reachability and strongly connected components. *ACM Trans Algorithms* 9(3):Article 27
- Levy DA, Nail PR (1993) Contagion: a theoretical and empirical review and reconceptualization. *Genet Soc Gen Psychol Monogr* 119(2):233–284
- Li G, Zhu Z, Cong Z, Yang F (2014) Efficient decomposition of strongly connected components on GPUs. *J Syst Architect* 60(1):1–10
- Lotka AJ (1956) Elements of physical biology, 2nd edn. Williams and Wilkins Co. (1924), Dover Publications, New York
- McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. *Annu Rev Sociol* 27(1):415–444
- Merton RK (1968) The Matthew effect in science. *Science* 159(1968):56–63
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167–256
- Pearl J (2000) Causality. Cambridge University Press, Cambridge
- Pearson M, Steglich C, Snijders T (2006) Homophily and assimilation among sport-active adolescent substance users. *Connections* 27(1):47–63
- Price DJ d S (1976) A general theory of bibliometric and other cumulative advantage processes. *J Am Soc Inform Sci* 27(1976):292–306
- Roller R, Blommestijn SQ, Treur J (2017) An adaptive computational network model for multi-emotional social interaction. In: Proc. of the 6th international conference on complex networks and their applications, COMPLEXNETWORKS'17. Studies in computational intelligence Springer Verlag
- Sharpanskykh A, Treur J (2013) Modelling and analysis of social contagion processes with dynamic networks. In: Bădică C, Nguyen NT, Brezovan M (eds) Computational collective intelligence. Technologies and applications. ICCCI 2013. Lecture notes in computer science, vol 8083. Springer, Berlin, pp 40–50
- Sharpanskykh A, Treur J (2014) Modelling and analysis of social contagion in dynamic networks. *Neurocomputing* 146(2014):140–150
- Simon HA (1955) On a class of skew distribution functions. *Biometrika* 42(1955):425–440
- Tarjan R (1972) Depth-first search and linear graph algorithms. *SIAM J Comput* 1(2):146–160
- Treur J (2016a) Verification of temporal-causal network models by mathematical analysis. *Vietnam J Comput Sci* 3:207–221
- Treur J (2016b) Network-oriented modeling: addressing complexity of cognitive, affective and social interactions. Springer Publishers, Cham
- Treur J (2017) Modelling and analysis of the dynamics of adaptive temporal-causal network models for evolving social interactions. *Comput Soc Netw* 4:1–20
- Treur J (2019) The ins and outs of network-oriented modeling: from biological networks and mental networks to social networks and beyond. In: Transactions on computational collective intelligence 32, 120–139 (2019). Text of Keynote Lecture at the 10th International Conference on Computational Collective Intelligence, ICCCI'18
- Treur J (2018b) Relating emerging network behaviour to network structure. In: Aiello LM, Cherifi C, Cherifi H, Lambiotte R, Lió P, Rocha LM (eds) Proc. of the 7th international conference on complex networks and their applications, ComplexNetworks'18. Studies in computational intelligence 812. Springer Publishers, Cham. pp 619–634
- Treur J (2018c) Relating an adaptive social network's structure to its emerging behaviour based on homophily. In: Aiello LM, Cherifi C, Cherifi H, Lambiotte R, Lió P, Rocha LM (eds) Proc. of the 7th international conference on complex networks

- and their applications, *ComplexNetworks'18. Studies in computational intelligence*, vol 812. Springer Publishers, Cham. pp 635–651
- Treur J (2018d) Mathematical analysis of a network's asymptotic behaviour based on its strongly connected components. In: *Proc. of the 7th international conference on complex networks and their applications, ComplexNetworks'18*, vol. 1. *Studies in computational intelligence*, vol 812. Springer Publishers, Cham. pp 663–679
- Turnbull L, Hütt M-T, Ioannides AA, Kininmonth S, Poepl R, Tockner K, Bracken LJ, Keesstra S, Liu L, Masselink R, Parsons AJ (2018) Connectivity and complex systems: learning from a multi-disciplinary perspective. *Appl Netw Sci* 3(47). <https://doi.org/10.1007/s41109-018-0067-2>
- van Beukel S, Goos S, Treur J (2017) Understanding homophily and more-becomes-more through adaptive temporal-causal network models. In: De la Prieta F (ed) *Trends in cyber-physical multi-agent systems. The PAAMS collection - 15th Int. Conf. PAAMS'17. Advances in intelligent systems and computing*, vol 619. Springer, Cham. pp 16–29
- van Beukel S, Goos S, Treur J (2019) An adaptive temporal-causal network model for social networks based on the Homophily and more-becomes-more principle. *Neurocomputing* 2019 in press
- van Dijk M, Treur J (2018) Physical activity contagion and homophily in an adaptive social network model. In: Nguyen NT, Pimenidis E, Khan Z, Trawinski B (eds) *Computational collective intelligence: 10th international conference, ICCCI 2018, proceedings*, vol 1. Springer Verlag, Cham. pp 87–98 *Lecture Notes in AI*, vol. 11055
- van Gerwen S, van Meurs A, Treur J (2019) An adaptive temporal-causal network for representing changing opinions on music releases. In: De La Prieta F, Omatu S, Fernández-Caballero A (eds) *Distributed computing and artificial intelligence, 15th international conference. DCAI 2018. Advances in intelligent systems and computing*, vol 800. Springer, Cham, pp 357–367
- Vazquez F (2013) Opinion dynamics on coevolving networks. In: Mukherjee A, Choudhury M, Peruani F, Ganguly N, Mitra B (eds) *Dynamics on and of complex networks, volume 2, modeling and simulation in science, engineering and technology*. Springer, New York, pp 89–107
- Vazquez F, Gonzalez-Avella JC, Eguíluz VM, San Miguel M (2007) Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states. *Phys Rev E* 76:046120
- Wijs A, Katoen JP, Bošnjak D (2016) Efficient GPU algorithms for parallel decomposition of graphs into strongly connected and maximal end components. *Form Meth Syst Des* 48:274–300

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
