The desire to avoid cognitive dissonance drives community formation in a social network model – Supplementary Information

Zoltán Kovács[1, 2], **Anna Zafeiris**[1, 2, +], **and Gergely Palla**[3,1,+,*]

[1]Department of Biological Physics, Eötvös University, Pázmány P. stny 1/A, 1117, Budapest, Hungary
[2]MTA-ELTE Lendület "Momentum" Innovation Research Group, Eötvös University, Institute of Archaeological Sciences, Budapest, 1088, Hungary
[3]Health Services Management Training Centre, Semmelweis University, Kútvölgyi út 1., 1125 Budapest, Hungary
[*]gergely.palla@emk.semmelweis.hu
[+]these authors contributed equally to this work

S1 Flowcharts of simulation process

To supplement the model description, we present also flowcharts depicting the logic of the simulations used in the main paper. Fig.S1 shows the flowchart detailing the simulation process in the social network, Fig.S2 shows the flowchart detailing the communication process in the social network and Fig. S3 shows the flowchart detailing the stabilisation process of agents.

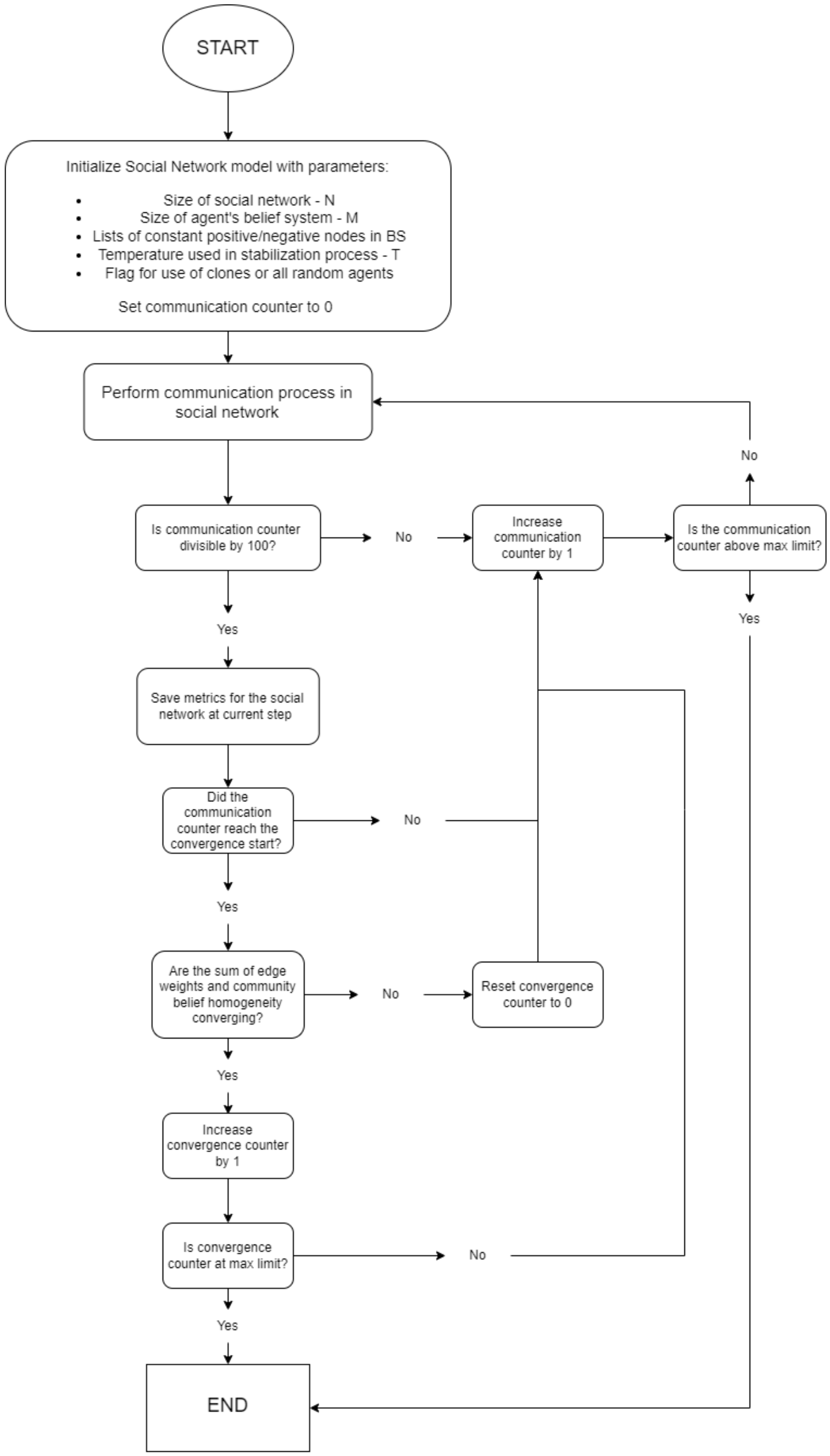


Figure S1. Flowchart detailing the simulation process in the social network

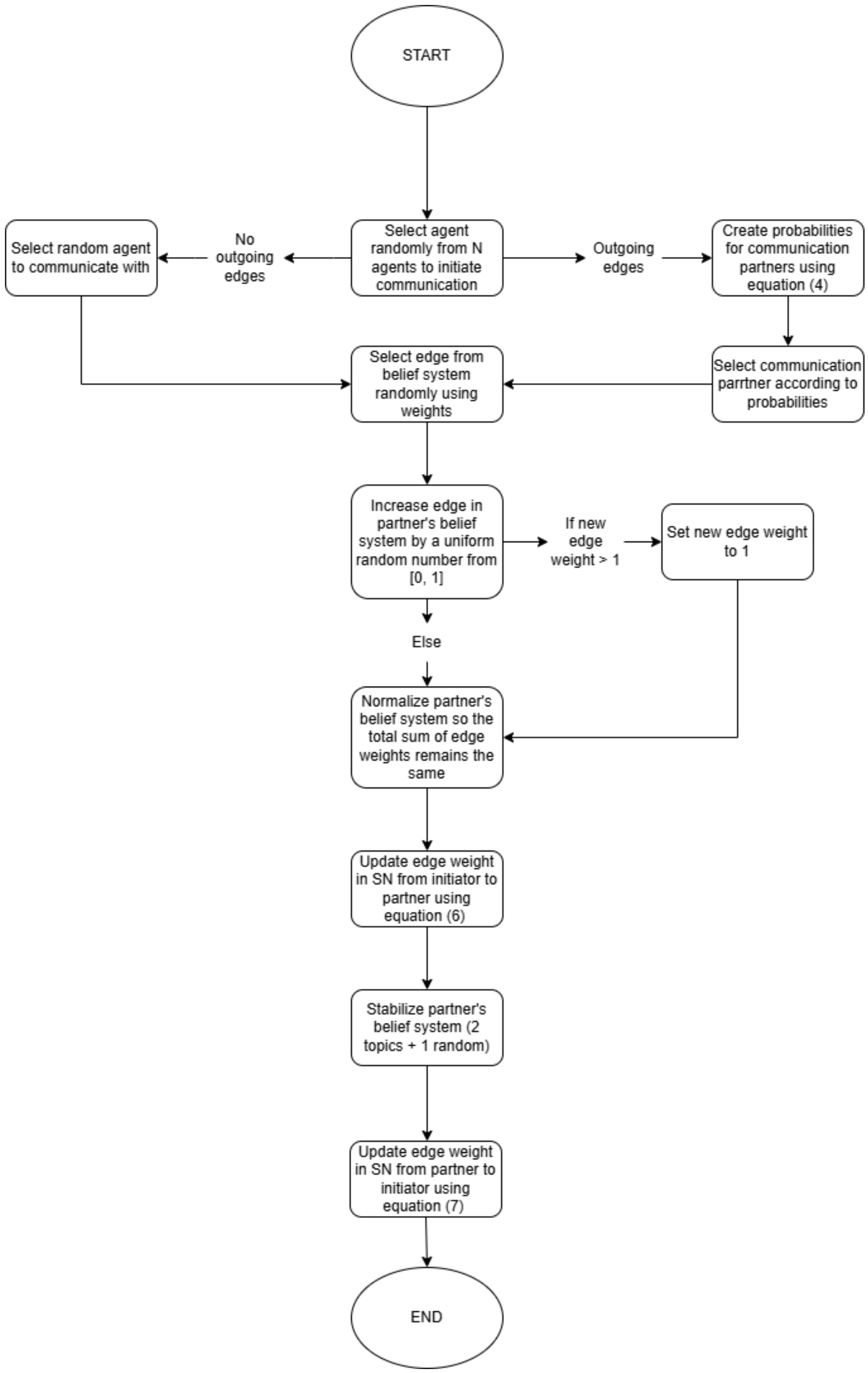


Figure S2. Flowchart detailing the communication process in the social network

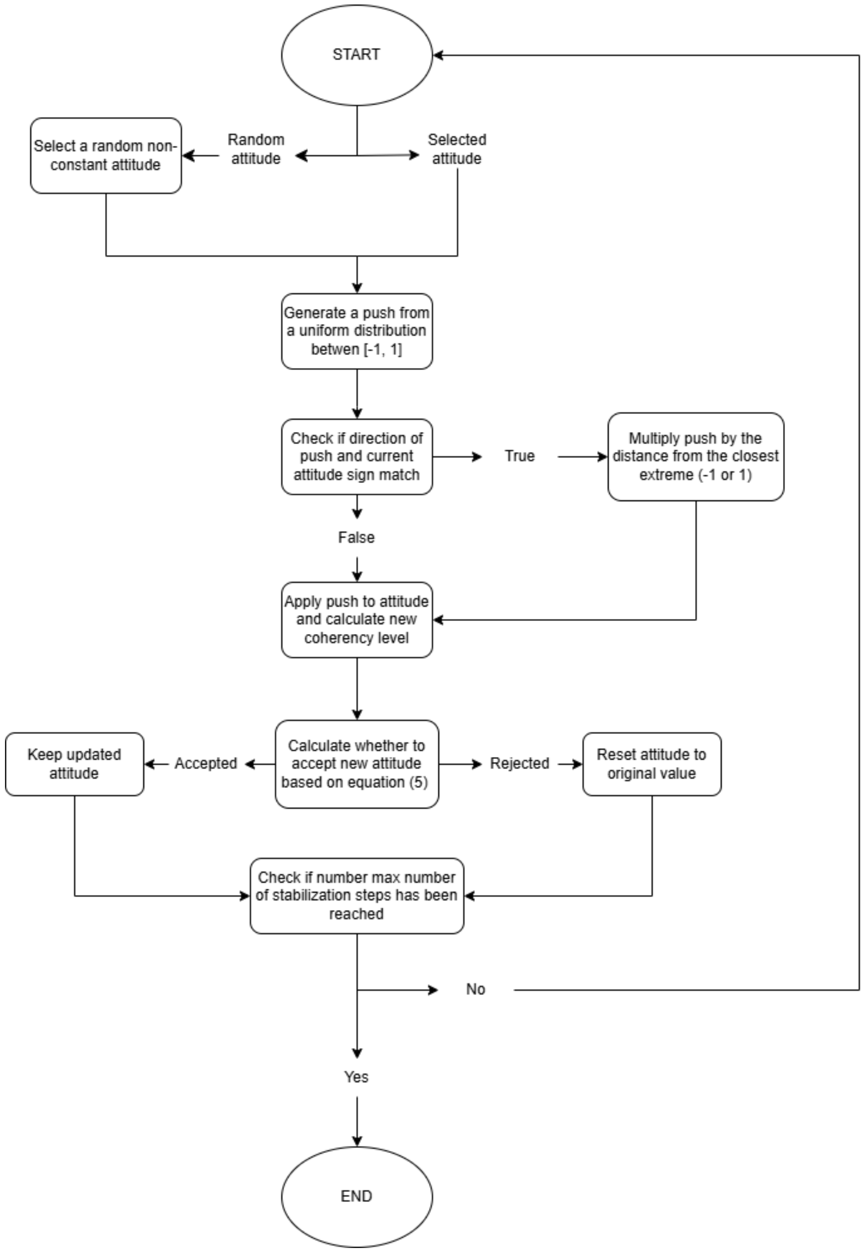


Figure S3. Flowchart detailing the stabilization process of agents

S2 250 agent simulation results

The Figs. S4–S7 show the results of simulations with 250 agents. Due to the larger size of the social network, the run-time of the simulations was longer, so we used fewer dissonance penalty values and total iterations. The results show similar behaviour to those obtained on 100 agent simulations.

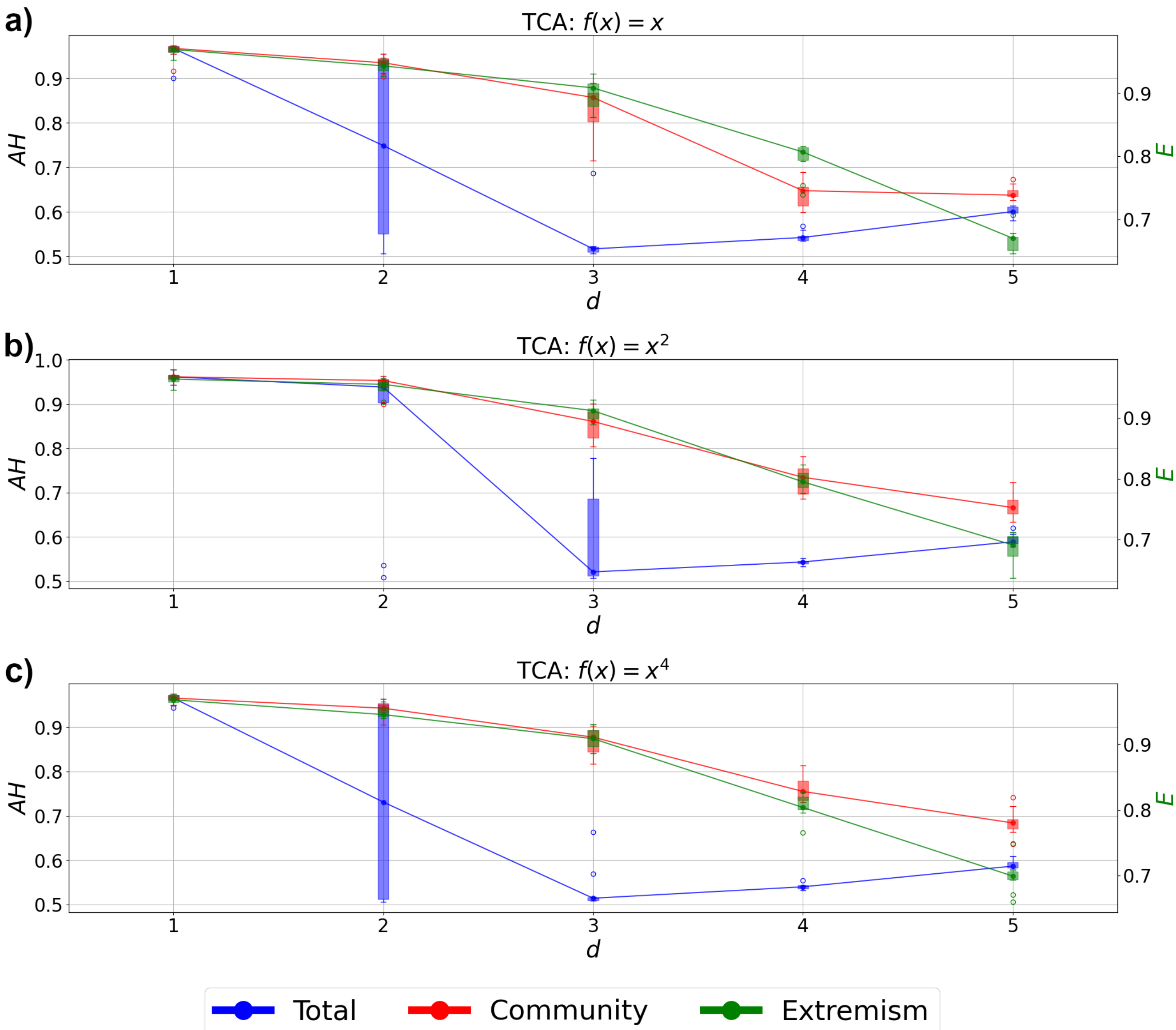


Figure S4. The total and intra-community homogeneity of the agent's attitudes. While the total homogeneity drops sharply around a dissonance penalty of 3, the communities stay close in attitudes. This shows that in our model, the agents try to form social groups with others who have similar attitudes towards most topics. This figure is analogous to Fig.4 in the main text.

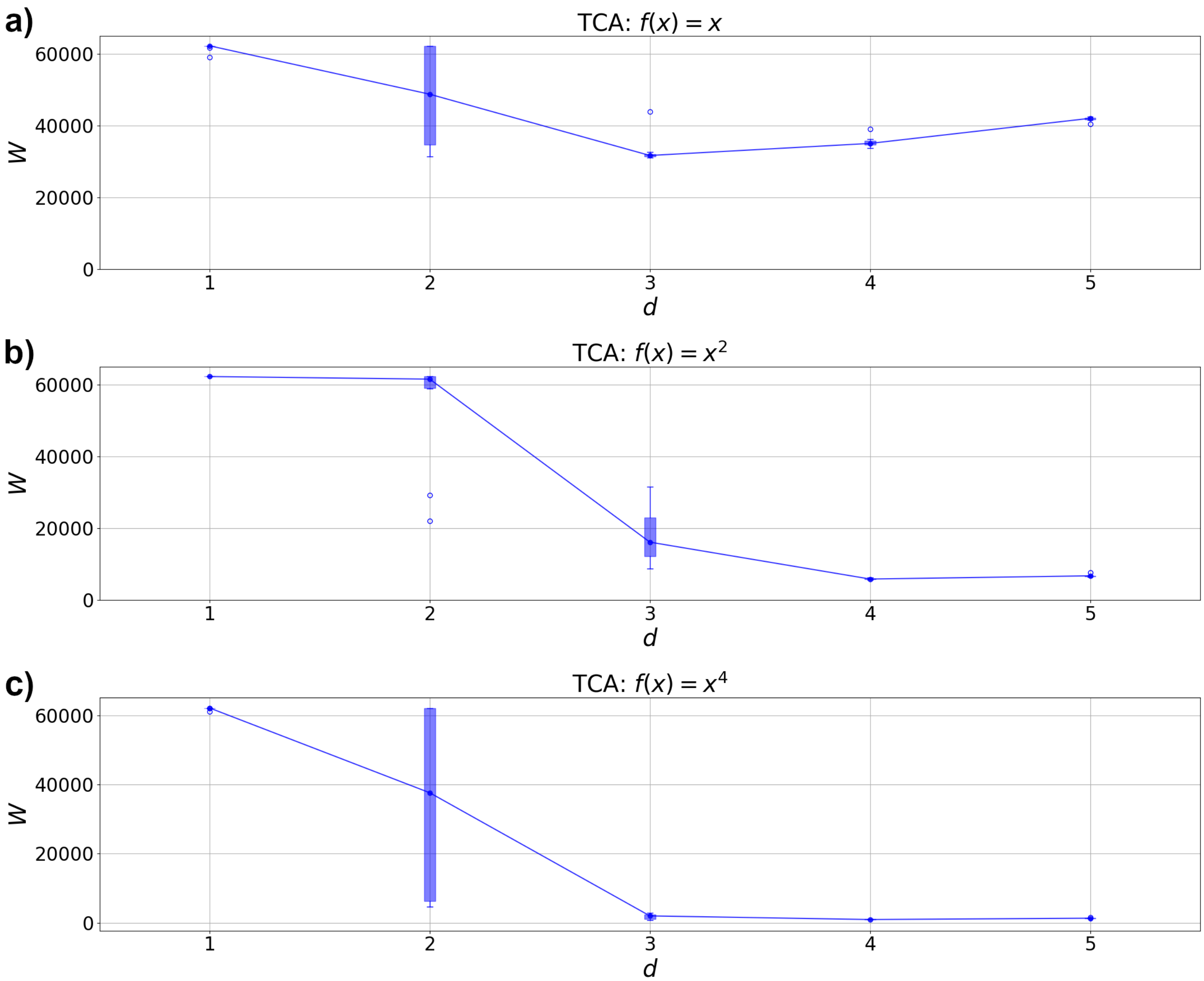


Figure S5. Sum of edge weights after convergence for different dissonance penalties. We can see that for the $f(x) = x$ case around half of the possible edges still remain, but for the two stricter TCA-s, a vast majority of the edges disappear, resulting in a sparse social network. This figure is analogous to Fig.5 in the main text.

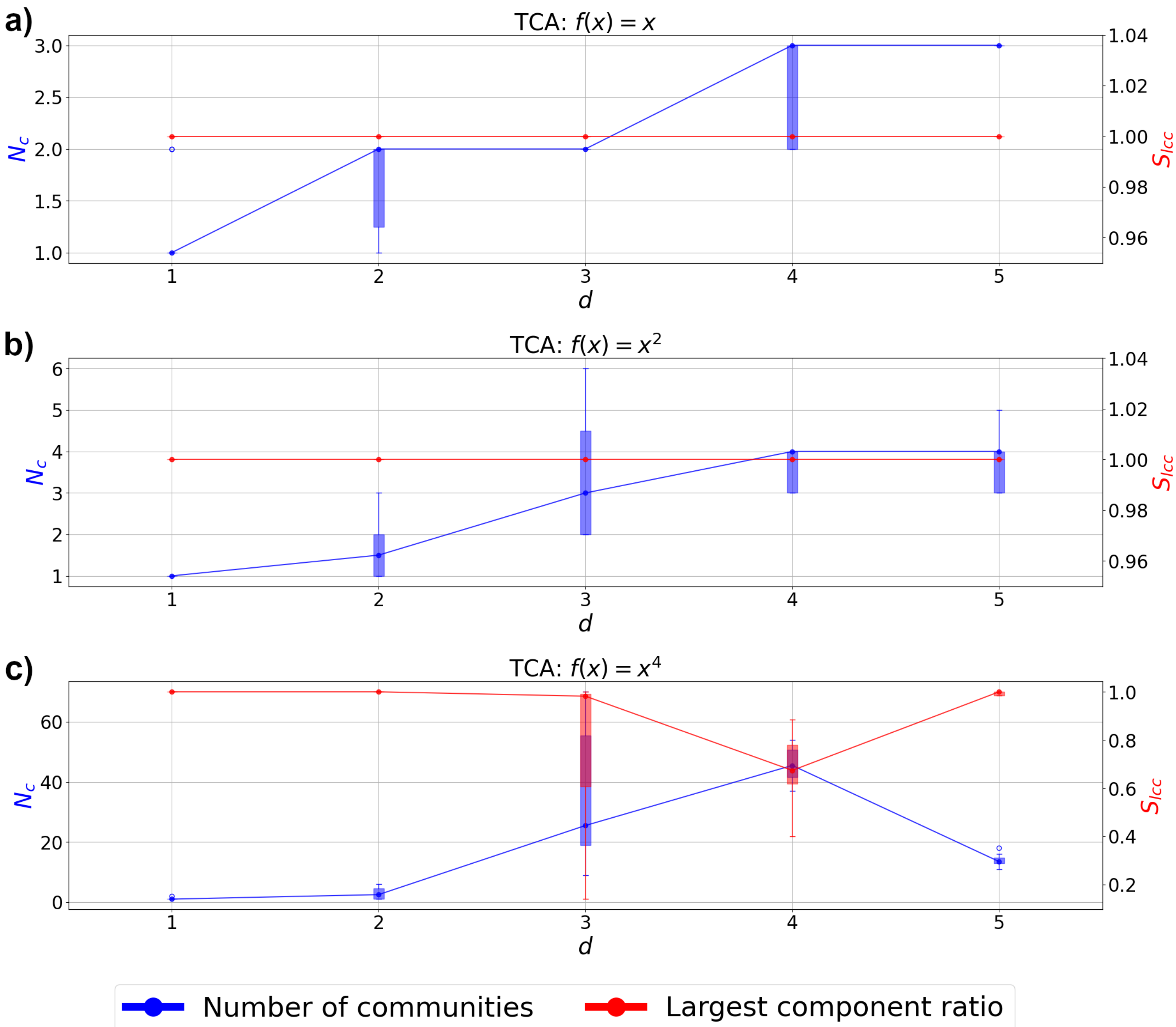


Figure S6. The number of communities and the ratio of the largest component to the total nodes in the social network. The fully connected social network starts to break down at around a dissonance penalty of 3, which results in on average 2 groups for $f(x) = x$ TCA, and 20 to 30 for the stricter TCA-s. This simultaneously means that the giant component slowly disappears, and the communities become isolated from each other. This figure is analogous to Fig.6 in the main text.

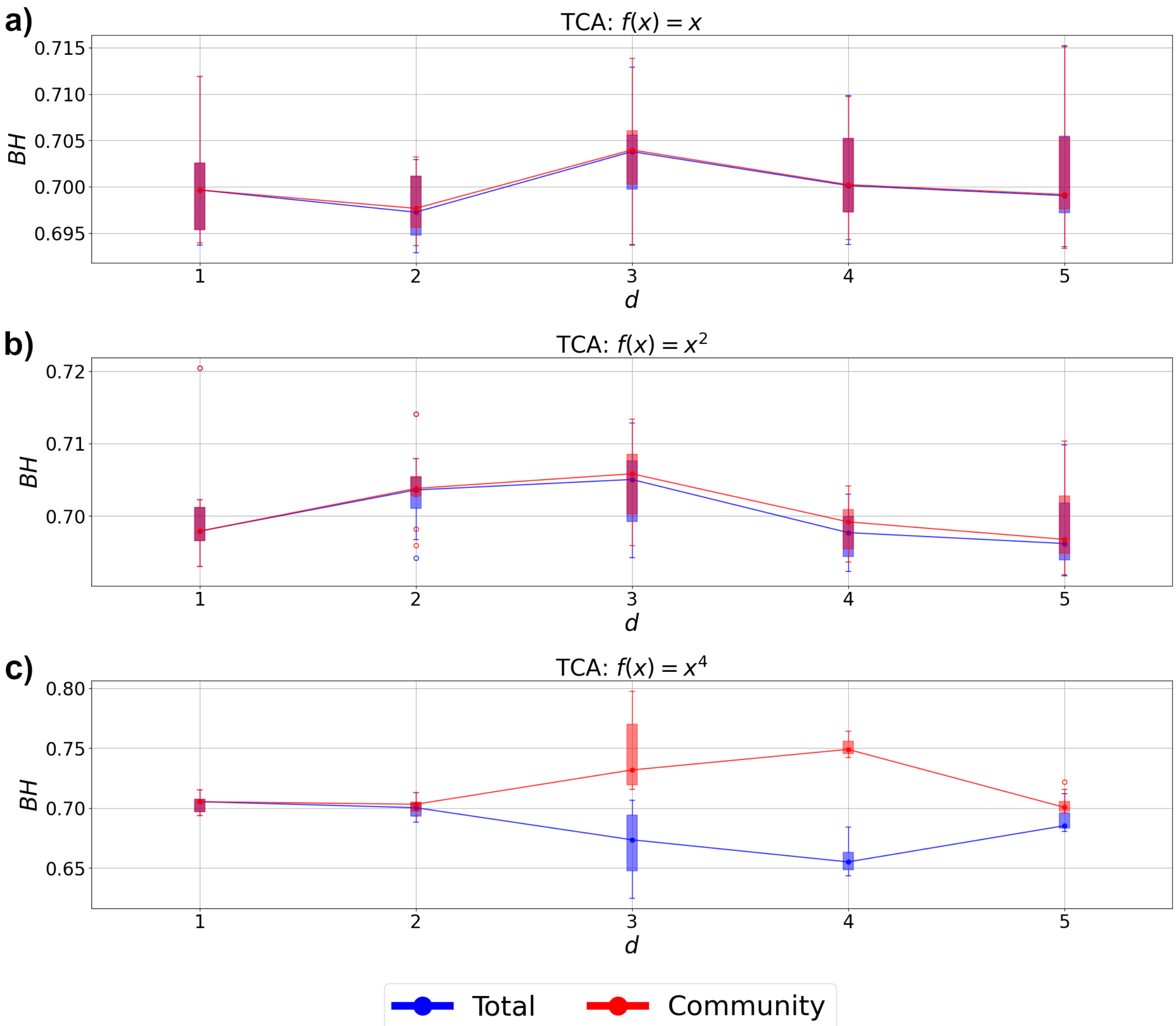


Figure S7. The total and intra-community homogeneity of the belief systems. We can also see the difference between the TCA: $f(x) = x$ case, and the stricter TCA-s. When the social network becomes sparse, the belief systems inside the communities also become more similar. This figure is analogous to Fig.8 in the main text.

S3 Single agent simulations

To isolate and examine the fundamental dynamics of our model, we ran simulations involving single, isolated agents. These simulations serve as an analogy to the communication process. At each iteration step, a link in the agent's belief system was chosen randomly and its weight was increased, mimicking the influence of external information or communication. We then continuously monitored changes in the agent's internal parameters, including attitudes and coherency levels.

S3.1 Flips in the belief system

The most fundamental process driving the changes in the social network in our simulations is the ability of agents to undergo an attitude flip, potentially changing the sign of their overall attitude. Our findings show that these flips occur with higher frequency as the dissonance penalty increases.

To rigorously examine this phenomenon, we monitored the total number of flips and the inter-flip waiting times across 500,000 simulation steps, utilising 1,000 independent agents across a variety of dissonance penalty values.

Our results show that the number of flips on average starts to go above 0 at a dissonance penalty of 2, and continues to increase afterwards. The full results can be seen in Fig.S8.

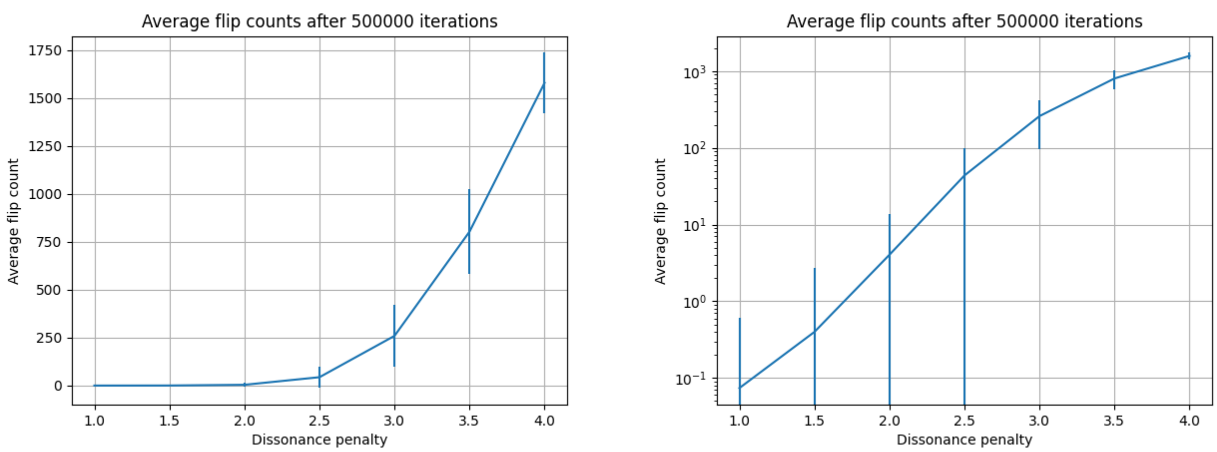


Figure S8. Number of attitude flips for different dissonance penalties, averaged over 1000 agents, with linear and logarithmic scales. The number of flips starts to increase at $d = 2$.

We also monitored the wait times between attitude flips. Our key observations were that the wait time initially follows a log-normal distribution for lower dissonance penalties, which then transitions into an exponential distribution at $d = 4$. These results can be seen in Fig.S9.

S3.2 Attitude changes at flips

To quantify the internal impact of an attitude flip, we measured the change in the agent's coherency level immediately following the flip event. This was calculated as the difference in coherency (coherency after flip - coherency before flip).

For all parameter sets examined, the distribution of these coherency changes was consistently Gaussian (Normal). The expected value (mean) of the distribution was observed to be positive, indicating that attitude flips generally increase the agent's internal coherency.

A relatively small percentage of attitude flips resulted in a negative change in coherency. This minor deviation is likely attributable to the inherent stochasticity of the stabilization process which causes the flip. These results, illustrating the coherency change distribution, are presented in Fig.S10,

S3.3 Time spent in optimal state in the belief system

A core assumption of our model is that agents adjust their belief system to maximize internal coherency. Since higher dissonance penalties facilitate easier attitude changes, we hypothesized that agents would align their current attitude sign with the theoretically optimal sign more frequently. To measure this, we introduced two fictional "reference" agents, one with all attitudes at +1, except the constant negative and the other with all attitudes at -1, except the constant positive.By taking the associations (link weights) from the observed agent's evolving belief system and calculating the coherency levels of these two reference agents, we determined which of the two theoretical states was instantaneously more optimal for the agent. We then monitored the fraction of total simulation time the observed agent spent with its current attitude sign aligned with this more optimal theoretical state. As shown in Fig.S11, this fraction begins to increase at higher dissonance penalties, confirming that increased internal flexibility allows agents to better track and align with the coherency-maximizing state.

S4 Two agent simulations

To further explore the fundamental mechanics of our model, we conducted simulations involving a minimal social network consisting of only two agents. In this setup, the agents engage in communication with each other at every iteration step.

Across all simulation runs, we consistently observed that the two belief systems converged over time, leading to a state of extreme homogeneity. This rapid and complete synchronization highlights the strong influence of continuous, reciprocal communication on belief alignment within a minimal network structure. An illustrative example of this process, showing the saturation of belief system homogeneity, is presented in Fig.S12.

This high homogeneity results from a process where a substantial number of shared link weights within the belief system decay to near zero.

Since the probability of an agent choosing a link for communication is dependent on its current weight, once a link's weight becomes negligible in the belief systems of both agents, it is highly unlikely to be reinforced or chosen again. This effectively removes the link from the communication process. The belief system distributions in these case take on a similar shape for both agents, however, they still aren't identical due to the randomness in the communication process. An example of these can be seen in Fig.S13.

Another aspect of two agent communication that we examined is whether the attitudes of the two agents converge over time. We calculated the difference in the non-constant attitudes between the two agents for multiple dissonance penalties. When dissonance penalty was 1, the agents almost never changed their initial belief system, so they were either in agreement from the beginning, which meant a distance of 0 or completely opposed to each other, which meant a distance of 2. The average in this case was 1, with a large standard deviation, meaning there was no significant effect of communication here. However, for larger dissonance penalty values, this distance started to decrease, since agents would flip their beliefs more often. However, there were still multiple cases where a difference in attitudes continued to exist. The results of these simulations can be found in Fig.S14.

S5 Investigating effect of triadic closure functions

An important parameter in our simulations was the triadic closure affinity function. This influenced the likelihood that an agent would communicate with a friend of their friends instead of their direct friends, as defined in eq.3 in the main text. We used three different power functions for this purpose: x, x^2 and x^4, and observed different behaviours for each.

To help clarify their exact meaning, I calculated their effect assuming that in our network, each node has n friends, and these friends increase the probability of communication with $n-1$ individuals. The results of this calculation can be found in fig.S15. An important result is that for the cases where the exponent was greater than 1, the relative probability of communication with friends of friends decreases after the average number of friends increases above 2. For the $TCA = x^1$ case, it starts to approach 1. It is important to note that in most systems, the friends of our friends may already be our friends directly, and as such instead of expanding our social network, we may just have an even higher chance of communicating with our well-connected friends.

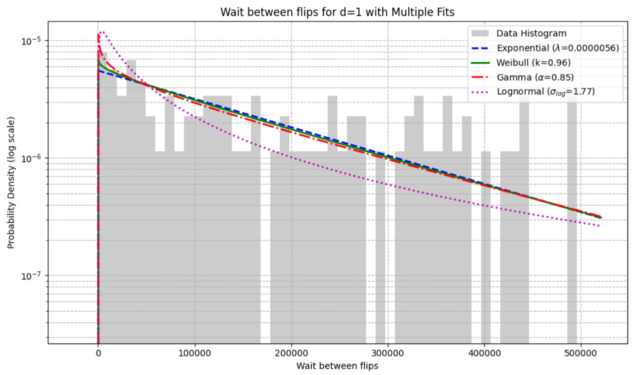
During our research, we also investigated what would happen if we turned off triadic closure completely, and agents only communicated with their direct friends. What we found in this case is that the resulting social network became unrealistic, with "chains" of agents forming instead of internally well-connected communities. An example of this is shown in Fig.S16.

S6 Simulations with fixed social network

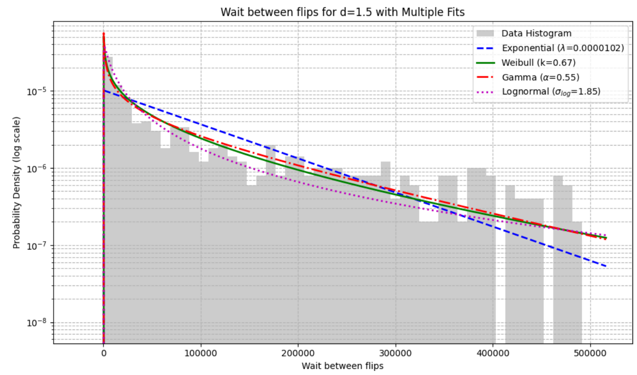
To see whether changes in the social network are responsible for the divergence of attitudes or belief systems, we performed a simulation where the social network was fixed. In these results, we could see that both the attitudes and the belief system weights started to diverge, similar to the evolving social network case. These results are presented in Fig.S17

S7 Social network and belief system visualization

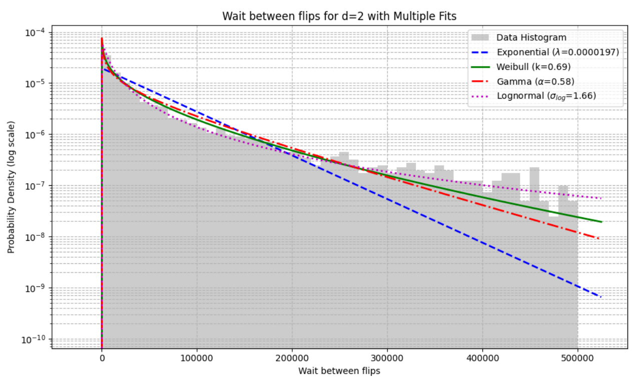
To help readers understand our results, we created a graphical representation of our social network, where the nodes of the network are replaced with the individual agent's belief systems. In the belief system, the two constant nodes are on the left and right respectively, and marked by squares. The other beliefs are represented by circles. Nodes are colored according to the attitudes of the agents, with red corresponding to negative and green corresponding to positive values.



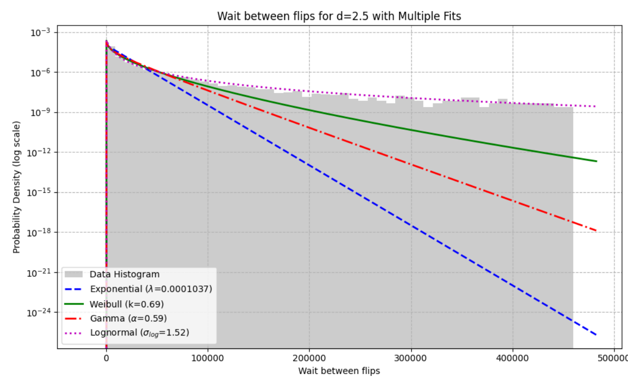
(a) Dissonance penalty = 1.0



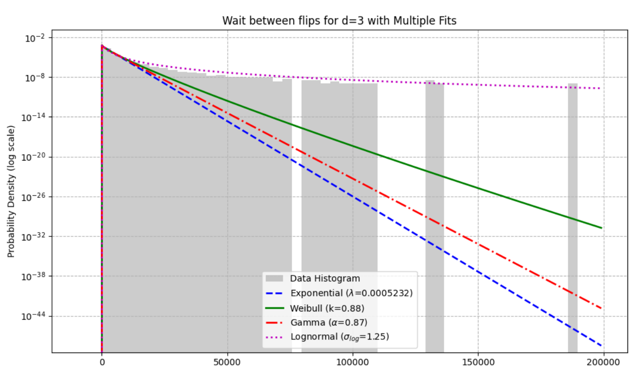
(b) Dissonance penalty = 1.5



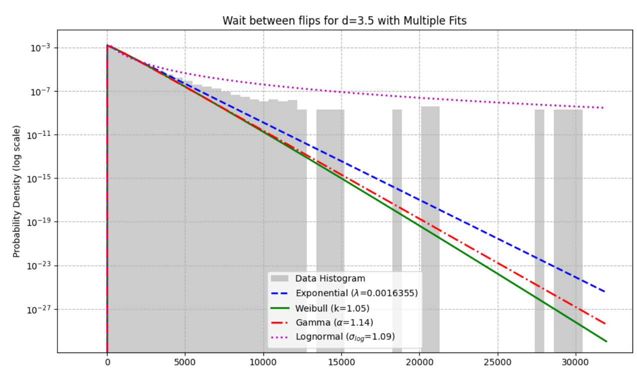
(c) Dissonance penalty = 2.0



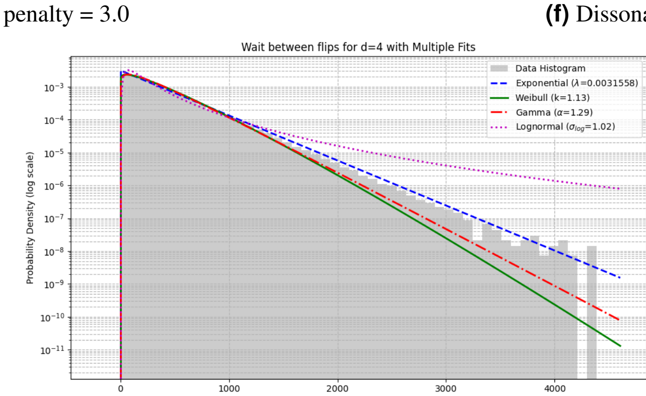
(d) Dissonance penalty = 2.5



(e) Dissonance penalty = 3.0

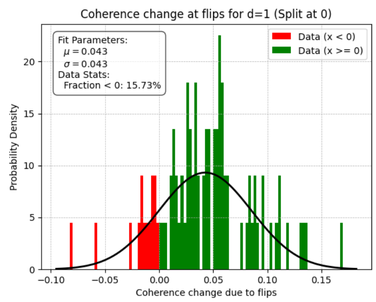


(f) Dissonance penalty = 3.5

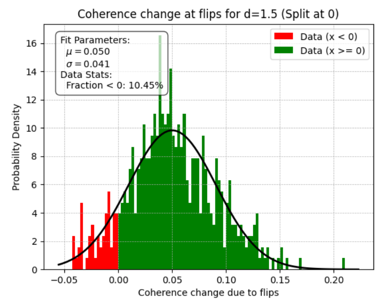


(g) Dissonance penalty = 4.0

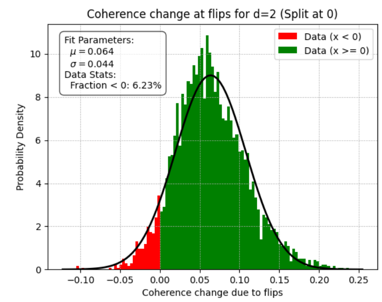
Figure S9. Distribution of flip wait times, with fitted distributions. The distribution slowly transforms from a log-normal to an exponential form.



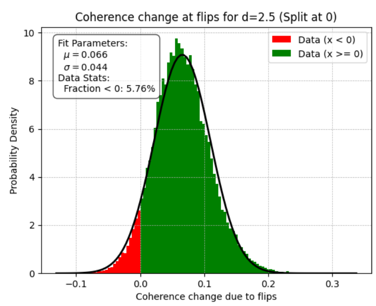
(a) Dissonance penalty = 1.0



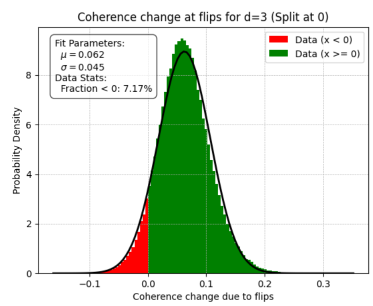
(b) Dissonance penalty = 1.5



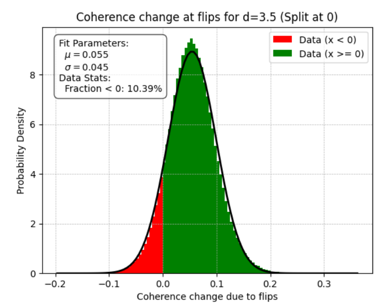
(c) Dissonance penalty = 2.0



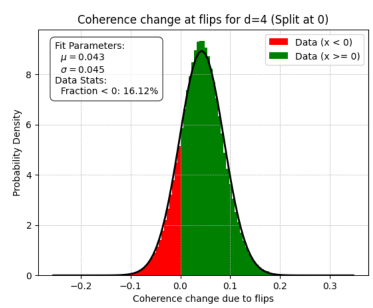
(d) Dissonance penalty = 2.5



(e) Dissonance penalty = 3.0



(f) Dissonance penalty = 3.5



(g) Dissonance penalty = 4.0

Figure S10. Distribution of coherence changes after opinion flips. The result is a normal distribution in each case, with the mean being at positive coherency. The fraction of "suboptimal" flips increases with dissonance penalty, where agents are more likely to flip their opinions.

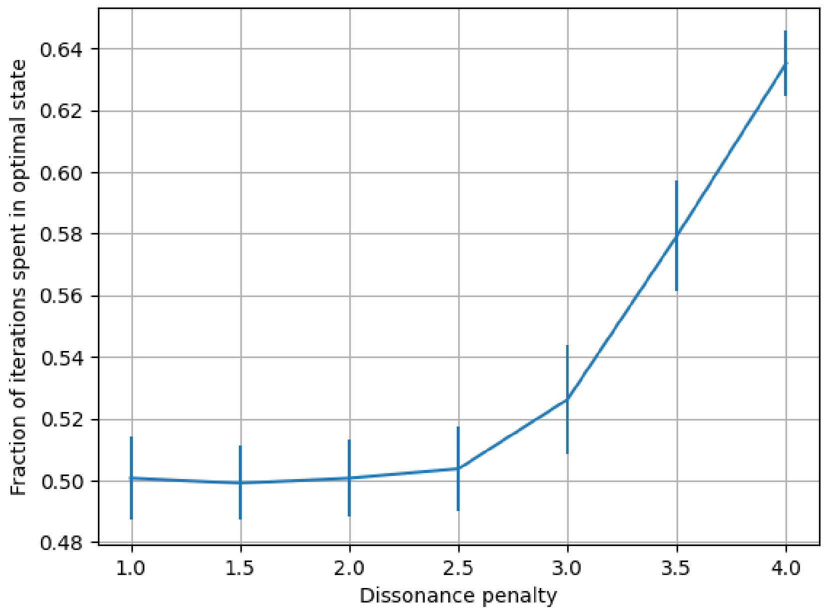


Figure S11. Fraction of iterations spent in the theoretically optimal attitude sign. At smaller dissonance penalties, the value is slightly above the completely random 0.5, but it starts to steadily increase for higher values.

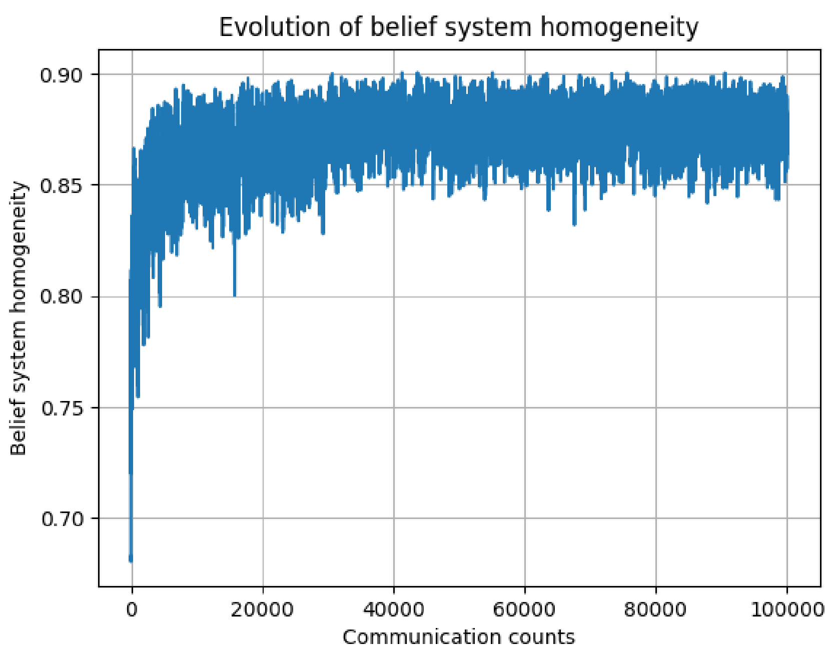


Figure S12. Evolution of the belief system homogeneity in a two agent social network. Over multiple communication iterations, the two belief systems start to become extremely similar to each other, since only a small number of topics are discussed.

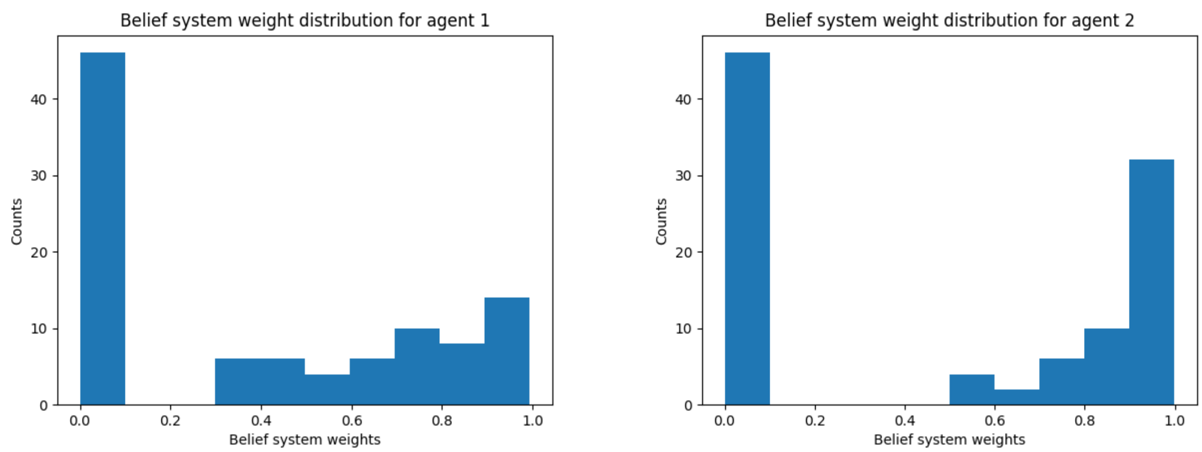


Figure S13. Belief system weight distribution for two agents communicating only with each other. The large peak of links with weights close to 0 is around the same size for both agents, representing the topics which don't occur during the communication process. The weights of the rest of the topics can either have a large peak close to 1, or have a more uniform distribution.

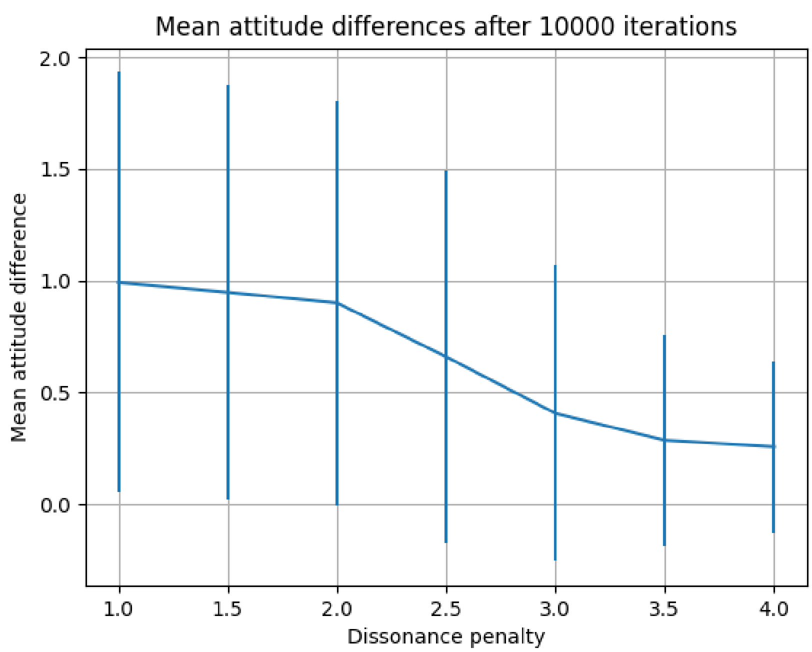


Figure S14. Mean attitude differences after 10000 communication steps for two agents. Starting from a completely random case at d=1, the differences start to shrink. This corresponds to the fact that agents who have a higher intolerance for cognitive dissonance tend to reach agreement more often, since they can flip their attitudes more easily into the optimal state based on coherence.

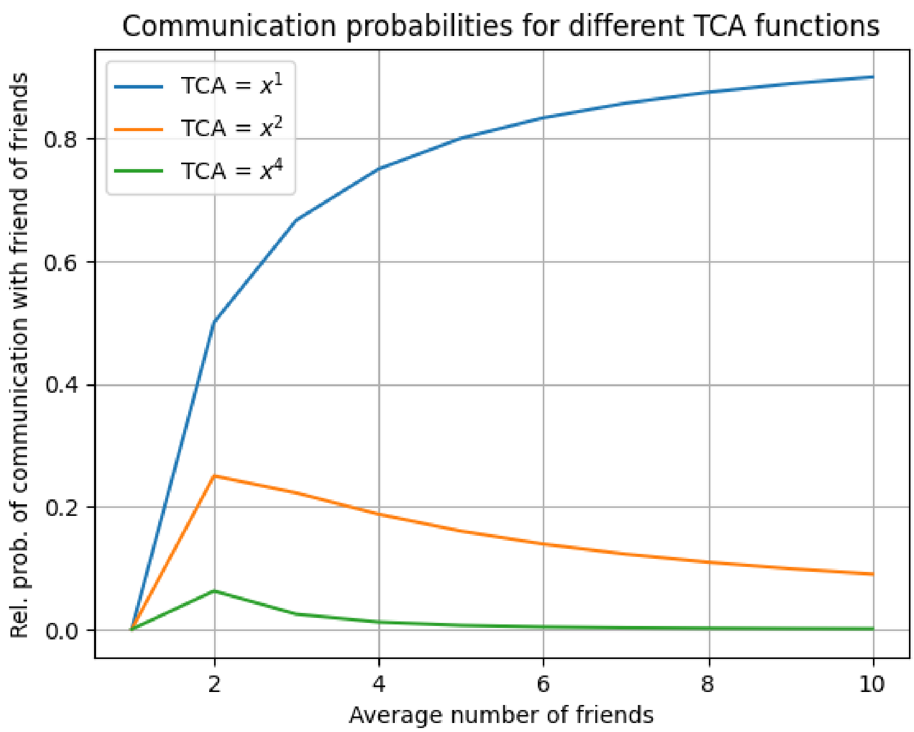


Figure S15. Relative communication probability with friends of friends for each TCA function.

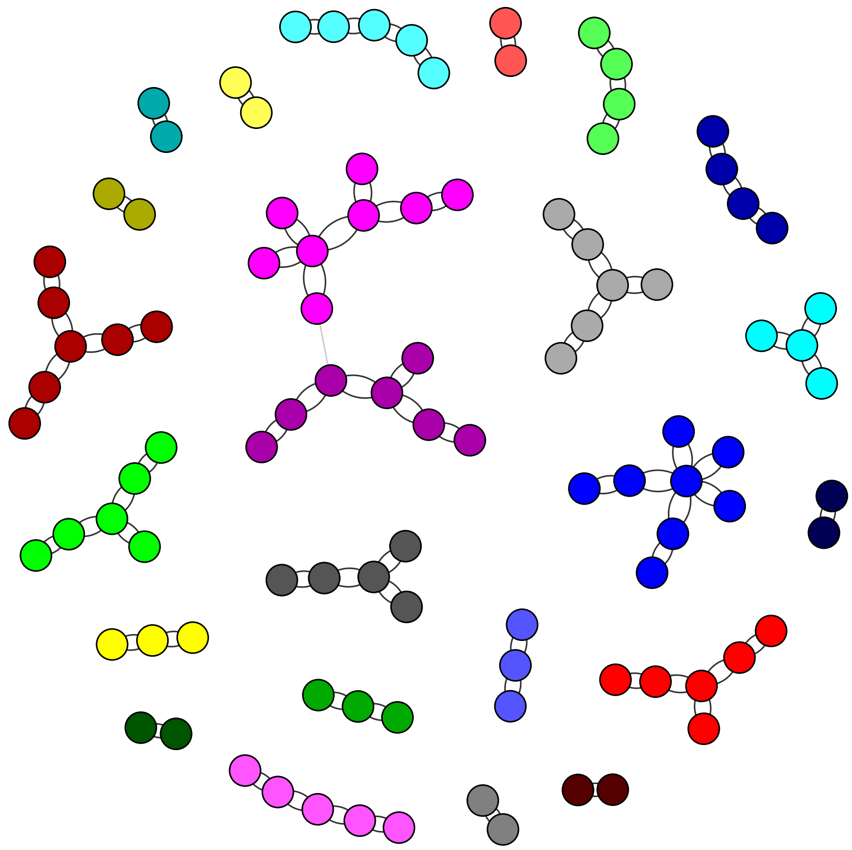


Figure S16. Social network in a simulation where communication with friends of friends was disallowed. In this case, agents tend to communicate with only their direct friend, and don't expand their social network further.

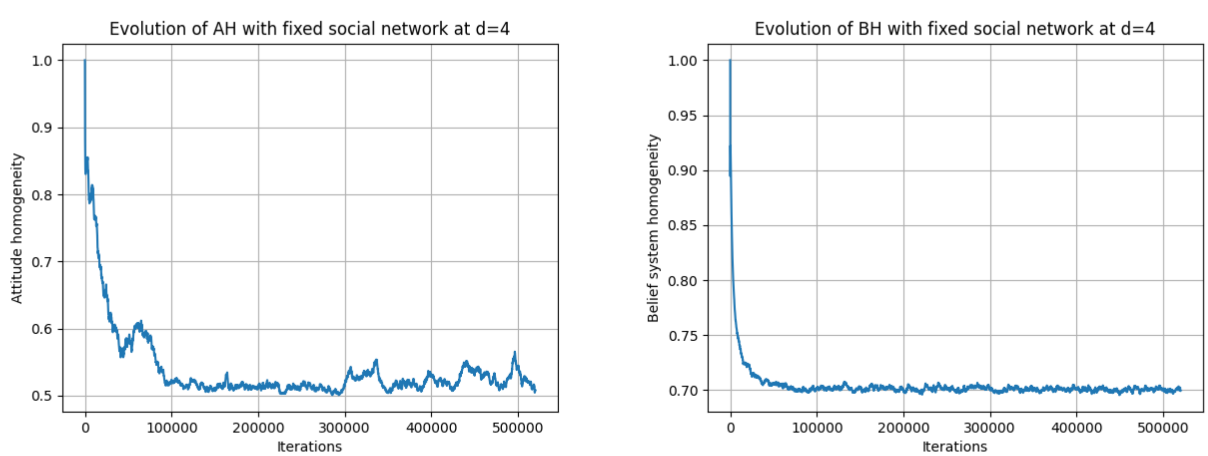


Figure S17. Time evolution of attitude homogeneity and belief system homogeneity over the communication process in a single social network with fixed links. Over time, both homogeneities start to diverge, similar to our evolving social networks.