

# Alone in the crowd: A computational social network model uncovering the clustering mechanisms of loneliness.

Bas D.L. Châtel<sup>1,2</sup>, Rick Quax<sup>2</sup>, Geeske Peeters<sup>1,3</sup>, Jeroen Janssen<sup>1</sup>, Rense Corten<sup>4</sup>, Marcel G. M. Olde Rikkert<sup>1,5</sup>, Vítor V. Vasconcelos<sup>2,6</sup>

<sup>1</sup>Department of Geriatric Medicine, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>2</sup>Computational Science Lab, Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands

<sup>3</sup>Radboudumc Alzheimer Center, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>4</sup>Department of Sociology/ICS, Utrecht University, Utrecht, The Netherlands

<sup>5</sup>Department of Geriatric Medicine, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>6</sup>Centre for Urban Mental Health, University of Amsterdam, Amsterdam, The Netherlands

Correspondence should be addressed to [bas.chatel@radboudumc.nl](mailto:bas.chatel@radboudumc.nl)

**Abstract:** Loneliness is a complex multicausal system with many adverse (mental) health effects. Most interpretations of loneliness describe subjective loneliness as an individual pathology, leading to a lack of focus on network interactions. However, research has found that, like other conditions, loneliness tends to cluster in social networks, possibly due to social processes such as homophily (i.e., the increased propensity to befriend like-minded people) and induction (i.e., influencing others) (?). Induction consists of three pathways. First, the behavioral pathway states that loneliness can lead individuals to act less trustingly and more hostile toward others, potentially harming relationships and perpetuating loneliness. Second, the cognitive pathway states that loneliness can arise from a discrepancy between the expectation and perception of one's connection to their social network. Finally, the emotional contagion pathway states that individuals may experience a convergence of emotions through nonverbal communication. This study aims to model these influences using uses an agent-based model and determine the sufficiency or necessity for obtaining loneliness clustering in to simulate these influences and explore interplay in forming loneliness clusters within social networks. We found that the Homophily is encoded by initializing static networks with predefined modularity based on fixed labels, enabling the analysis of the effects of homophily on various combinations of inductive processes over time. The model could replicate empirical findings of three degrees of influence identified by ? were universal for ; however, they were found to be prevalent across most networks, a finding supported by ? . Our aligning with the results of ? . Therefore, these influence metrics prove uninformative in disentangling causal mechanisms. Additionally, we discovered that inductive pathways only contribute to loneliness clustering when a certain level of homophily already exists within the network. Lastly, our results indicate that each inductive pathway displays distinct dynamics and that the cognitive route is the most dominant predictor of clustering. Furthermore, we found that inductive pathways only lead to loneliness clustering when there is a certain degree of homophily. These findings suggest that the social environment context of an individual could be crucial in developing loneliness, highlighting the need for interventions targeting individuals and their social networks.

**Keywords:** Human behavior, Mental health, Social influence, Homophily, Agent-based modeling

## ● Introduction

- 1.1 Loneliness, defined as the unwelcome feeling of a gap between the social connections we want and the ones we have, is a complex multi-causal system (???) with numerous negative adverse (mental) health effects (???). For example, it is associated with an increased risk of developing clinical dementia (?) , depression (?) , and behaviors harmful to health (?) . In the Netherlands, 20% of 55-year-olds and different interpretations regarding its nature (?) . Most interpretations of subjective loneliness, as opposed to 62% of 95-year-olds experience moderate to severe loneliness (?) .

- 1.2 While subjective loneliness is often viewed as an individual-centric issue (?) distinct from social isolation (i.e., the objective absence of social relationships), describe loneliness as an individual pathology (?). However, loneliness is not only a personal attribute but also driven by it is important to recognize that network-level influences, as seen by factors influence loneliness. Research suggests that lonely individuals tend to connect with other lonely individuals, indicating the role of network influences in loneliness. This phenomenon extends beyond direct connections, as lonely individuals increase not only the likelihood of their friends experiencing loneliness (first-degree) but also their friends' friends (second-degree) and even their friends' friends' friends (third-degree). This concept, known as the tendency of lonely individuals to be connected with other lonely individuals. This "three degrees of influence," was proposed by Christakis and Fowler in their work on the spread of various traits within social networks (????). The clustering of loneliness could amplify the adverse effects of loneliness can exacerbate its detrimental effects on physical and mental health (?) and has been observed in various well-being (?), as observed across different populations (??). Recognizing the network-level dynamics of loneliness sheds light on its potential amplification and highlights its significance in public health and social research.
- 1.3 Despite existing literature, the causal mechanisms underpinning the clustering of loneliness are still not fully understood. One possible explanation directs us to the theoretical framework provided by ? describes three sociopsychological processes: homophily (homophily, induction, and shared environment).
- Homophily refers to the tendency of individuals to form connections with others that have similar attributes or characteristics ), shared environment (connected individuals have similar exposures to factors contributing to loneliness), and induction (??).
  - Induction refers to people influencing one another )? such that lonely individuals may contribute to the spread of loneliness in their immediate environment. Induction consists of three pathways: cognitive, behavioral, and emotional contagion. The cognitive pathway suggests that loneliness arises from a mismatch between individuals' social network expectations and their actual social experiences (??). The behavioral pathway posits that lonely individuals may behave in ways that decrease relationship quality with others, thereby increasing induction in their surroundings (??). The emotional contagion pathway proposes that nonverbal communication between individuals may lead to the convergence of emotions . Lonely individuals may contribute to the spread of loneliness in their immediate environment through their shy, awkward, and hostile behavior towards socially ambiguous cues. converging emotions (??).
  - Shared environment refers to connected individuals having similar exposures to factors contributing to loneliness (??).

This theoretical framework remains untested. Identifying and understanding Understanding these mechanisms may facilitate targeted interventions to promote social integration and potentially lead to a sustainable reduction of loneliness on a network level.

Our research aims to model the three inductive pathways under differing strengths of homophily and examine their Our research explores the relationship between homophily and induction processes in social networks. To this aim, we first assess whether the model can replicate the "three degrees of influence" as a way of measuring clustering. We assess the sufficiency or necessity to induce of inductive pathways in inducing subjective loneliness clustering in social networks. We utilize a computational model to determine whether each pathway alone could be sufficient to exhibit loneliness clustering or whether their combination is necessary. For example, inductive influences could be particularly relevant when the population displays a certain while considering varying levels of homophily.

We utilize an agent-based model to simulate the interaction between individuals within a shared environment. With this simplifying assumption that the simulated population exists within the same environment, we can more easily test the interaction between homophily and induction.

We hypothesize that the level of homophily plays a significant role in loneliness clustering. For instance, when a population exhibits a high level of homophily, as lonely individuals befriending those of similar loneliness levels are more likely to sustain and exacerbate each other's loneliness. Conversely, for low homophily lonely individuals tend to befriend others with similar levels of loneliness, leading to a reinforcing cycle that intensifies their loneliness. In contrast, when homophily is low, individuals are less assort and better more integrated into a relatively healthier environment, reducing their capacity to impact their surroundings negatively. This decreased cumulative effect could weaken negative impact on their surroundings and diminishing the likelihood of clustering of loneliness loneliness clustering.

## ● Methods

- 1.1 By examining these dynamics, we hope to understand better the factors contributing to loneliness clustering in social networks. Ultimately, our findings can inform strategies and interventions aimed at reducing loneliness and promoting healthier connections among individuals.

## ● Model

- 2.1 This section provides a functional description of the model needed to understand the results. Readers interested in more technical details are referred to Additional details are available in the Extended Methods in Appendix ??A. An overview and justification of the assumed model parameters can be found in table ??2.

### Agent and network initiation

- 2.2 We present an agent-based model operating on a directed network comprising two subpopulations, each characterized by the initial loneliness level of its individuals (high and low initial energy). Loneliness is quantified as social energy ( $e$ ) bound between 0 and 1 ( $0 \leq e \leq 1$ ), representing a person's current social activity or engagement level. Lower social energy corresponds to greater loneliness. We present an agent-based model operating on a directed network comprising two subpopulations with  $N = 1000$  individuals in total, each characterized by the initial loneliness level of its individuals (high and low initial social energy). Both subpopulations are assumed to be within the same shared environment. This simplifying assumption removes the need to explicitly model the shared environment process, enabling a more straightforward analysis of the interplay between homophily and induction processes proposed in the theoretical framework by ?.
- 2.3 The population is initialized with a predefined initial assortativity (

### Homophily

- 2.4 Homophily is the observation that individuals with similar traits are more likely connected, with traits, e.g., ethnicity and race, contributing to shaping the social network ?. Here, we take the social network as a consequence of this homophilic process, and we initialize the population with a predefined level of homophily, i.e., a measure of homophily) encoding the clustering strength of the subpopulations in terms of agent relationships. We use the Pearson correlation coefficient to compute the association between the energies of the relationship-initiating and relationship-receiving agents. An fixed structure. We do this by attributing two arbitrary labels, associated with a homophilic trait, to the individuals and tuning the modularity of the network based on those fixed labels (i.e., the respective subpopulations). We scale the modularity by a factor of 2 such that an initial value of -1 indicates complete avoidance of one's subpopulation, 1 indicates exclusive interaction within one's subpopulation, and 0 denotes random mixing between the two subpopulations.
- 2.5 To incorporate the principle of homophily into our model, we examine various levels of initial assortativity in the starting population. First generate networks of varying degrees of homophilic-trait modularity, we initialize the population with an extreme assortativity ( $\rho = -1$  extreme modularity ( $Q = 1$ )) by creating two identical, unconnected components in the network, each characterized only by their difference in initial energy. The relationships in these components are then rewired to converge towards the desired assortativity. This rewiring is done by creating an index pairing of identical relationships between both components and randomly rewiring the pairs homophilic-trait tag. Then, we rewire connections to affect modularity. Since the two subgraphs are identical, we randomly select a connection (same in both subgraphs) and rewire the pairs, turning two homophilic connections into two heterophilic ones. This process introduces entropy in a stepwise manner stepwise and, given that the network is large enough, will reach all assortativity-modularity levels between 1 and -1. The value  $\rho = -1$  is reached when the complete list is rewired, and the network is fully bipartite. For a visual representation of this process, we refer the reader to Figure 5 in Appendix A.

- 2.6 Finally, to understand whether low social energy can cluster, we map the initial levels to the homophilic trait (thus, individuals with different traits start with distinct social energy levels).

## State ~~Update~~update

- 2.7 At each timestep, all agents simultaneously update their energy levels based on the three sociopsychological induction processes proposed by ?. The relative influence of each pathway can be adjusted, allowing for their strengths to range from 0 (i.e., the pathway does not influence the agents' energy) to 1 (i.e., the pathway is solely responsible for changes in the agents' energy) in a linear combination. The ~~updates to an agent's energy level are determined by their~~ sociopsychological state and ~~their~~ the interactions with other agents in the network ~~determine an agent's energy level update.~~

### Cognitive pathway

- 2.8 The cognitive pathway ~~proposes states~~ that loneliness can arise from a discrepancy between the expectation and the perception of one's connection to their social network. This perception is affected by interactions with others and can change over time -
- 2.9 To approximate and better understand (?????). To model this process, we ~~model consider~~ the interplay between social energy and perceived connectivity with one's social network. Perceived social connectivity reflects how connected a person feels to their social environment.

2 Our model considers that perceived social connectivity ~~changes over time due to social energy and the natural decay of social relationships. If a person's social energy is lower than their perceived social connectivity multiplied by a decay rate decays over time while social energy rebuilds it. Suppose a person has no social energy to engage with others. Then, their perceived social connectivity will gradually decrease over time. This decay rate is represented by a constant parameter called  $\beta$ , at a fixed decay rate.~~ A higher decay rate means a person's perceived social connectivity will decrease more rapidly without social energy. Social energy will rebuild this. The intuition here is that one needs to put social energy into upholding social relationships; without it, there will be a gradual reduction in connectivity.

- 2.11 ~~Pereived (perceived) connectivity. This perceived~~ social connectivity also feeds back into social energy. The ~~social energy at any given time is determined by the~~ difference between a person's perceived social connectivity and the average connectivity of their neighbors (i.e., the expectation of how their social connectivity should be). ~~This difference is normalized based on the number of incoming connections a person has to other nodes in the network.~~ determines the rate of change in social energy at any given time.

### Behavioral pathway

- 2.12 Loneliness can lead individuals to act less trusting and more hostile toward others, potentially harming relationships and perpetuating loneliness (??). We model this by calculating neighboring nodes' energy contributions based on their number of outgoing connections. Nodes with fewer connections can allocate relatively more energy to a friend. Furthermore, we incorporate a threshold to determine when an agent is considered lonely and may exhibit less trusting and more hostile behavior. This threshold determines when one starts to have a negative impact on its surroundings. Finally, the contribution of others is weighted by the social energy of the agent itself, illustrating that the energy levels of both the individual and their neighbors influence one another. An individual with low ~~social~~ energy becomes less receptive to others and displays "less trusting and more hostile" behaviors; the social energy of an individual then acts as a regulator, down-regulating the effect of others on the individual.

### Emotional contagion pathway

- 2.13 Individuals may experience a convergence of emotions through ~~nonverbal communication (non)verbal communication~~ (??). For example, facial expressions, vocalizations, postures, and movements can influence others in one's immediate environment to become more lonely, leading to a convergence of

emotions. Therefore, we assume that individuals will converge toward the mean energy for the emotional contagion pathway upon interaction.

## Analysis

### 2.14 We assess

#### Measuring clustering

- 2.15 To quantify loneliness clustering and the inductive processes of loneliness, we test whether our model can replicate the empirical findings. We assess whether any of the pathways replicates the "three degrees of influence" rule found in the body of work from Christakis and Fowler on obesity, smoking, happiness, and loneliness (????). This rule suggests that lonely individuals increase not only the likelihood of their friends being lonely (i.e., first degree) but also their friends' friends (i.e., second degree) and even their friends (i.e., third degree). To replicate this emergent property, we calculate the propensity of individuals with a specific trait to self-organizing at different degrees of separation compared to the random distribution of that trait within the network (i.e., the ratio of lonely to non-lonely individuals in (????)). Our metric for measuring influence was the maximum degree of influence (mDOI), the network. We use the method by largest network distance where the correlation of individuals self-organizing in the same trait is greater than zero. This method is derived from ? as this is less computationally intensive than the original methods used by ? and would theoretically yield the same results. We define the degree of separation at which this propensity reaches a negative value as the maximum degree of separation for self-organization. An example of this decline in correlation over distance can be seen in Figure 1, where the Pearson correlation is calculated over the social energy of each focal individual compared to their neighbors at different distances.
- 2.16 Next, we conduct a sensitivity analysis to examine the impact of varying levels of homophily on the system's convergence over time across the three pathways. Finally, we aim to identify the most influential pathways by analyzing different linear combinations of pathway strengths, enabling us to understand better the interplay between the various pathways and their relative contributions to the clustering of loneliness in the social network.

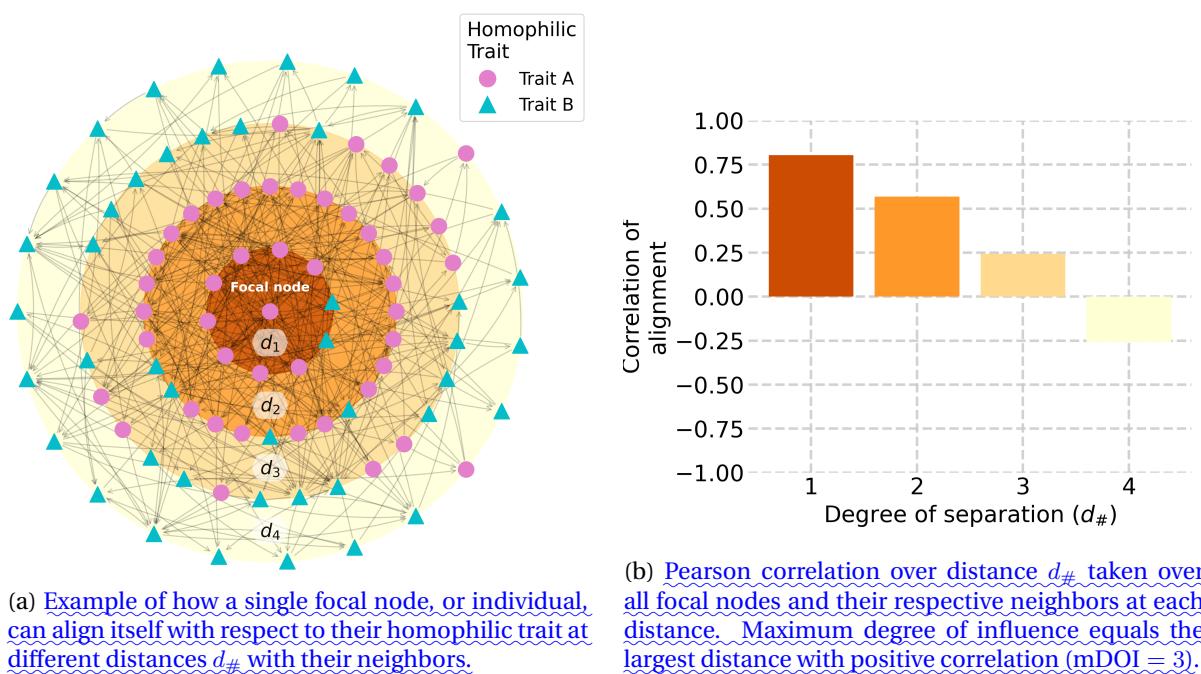


Figure 1: Example of how the maximum degree of influence (mDOI) is calculated.

## ● Results

- 3.1** We examined the interactions between homophily and induction to understand the causal mechanisms contributing to the clustering of loneliness in social networks. We assessed different configurations. Specifically, we examined how the processes of homophily and induction interact, based on the framework proposed by [1]. Firstly, we evaluated "peer influence" at different degrees, we assessed different configurations to investigate the reproducibility of the three degrees of influence observed empirically. We tested the maximum distance at which agents remained positively correlated or self-organized based on their social energy (i.e., the maximum degree of influence) across varying levels of initial homophily. Next, we analyzed the temporal dynamics of separation and then analyzed the model's temporal dynamics, including the convergence of assortativity and the convergence of the average energy in each subpopulation energy within subpopulations to quantify the differences among the inductive pathways and identify the pathway with the greatest influence on the system. This comprehensive approach provides insights into the underlying mechanisms shaping loneliness clustering in social networks.

### Degrees of separation influence

- 3.2** We investigated the extent of "peer influence" on social networks across various social distances. Our metric for measuring influence was the maximum degree of influence, the largest network distance where the likelihood of two individuals self-organizing in the same trait is greater than zero. An example

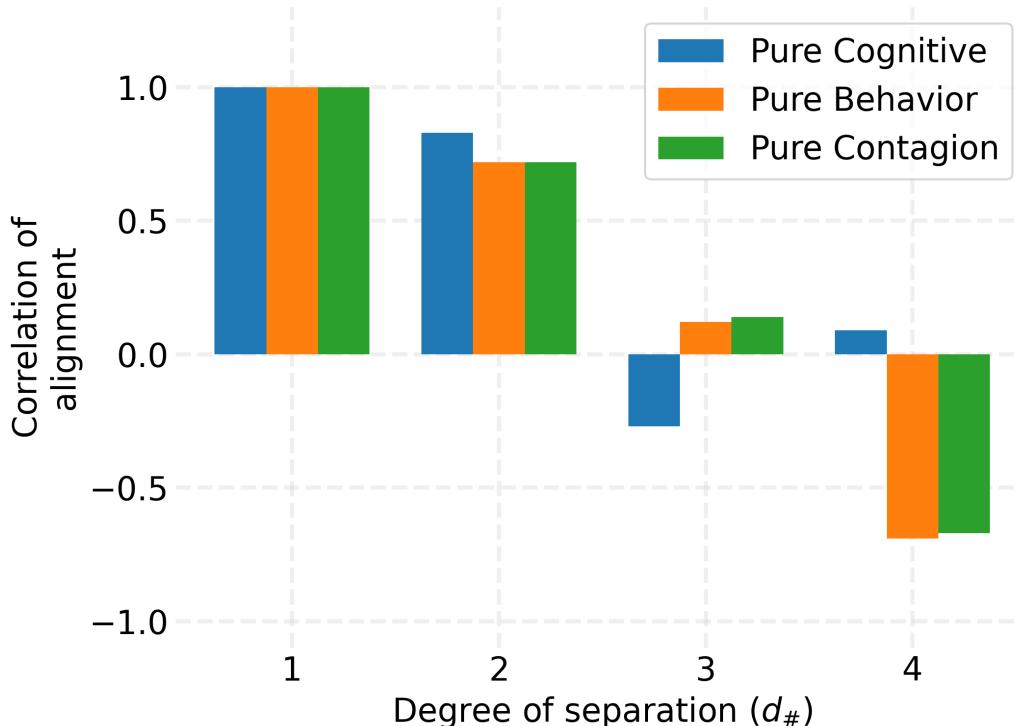


Figure 2: Peer influence—The correlation of self-organization on social networks ( $Q_0 = 0.8$ ) across varying social distances for all three inductive pathways on their own. The cognitive pathway has a single graph  $mDOI = 2$ , while the behavioral and emotional contagion pathways have a  $mDOI = 3$ , showing that the model is able to replicate the empirical findings of three degrees of influence.

100 timesteps 1.000 timesteps 10.000 timesteps Figure ?? shows the propensity to self-organize on social networks across varying social distances normalized on the first degree of separation. This graph considers a starting assortativity of .8, purely influenced by the contagion pathway, and has a maximum degree of separation of 3. Figures ??, ??, and ?? illustrate the maximum degree of separation for different starting assortativity values and pathways. Error bars indicate standard errors of the mean based on 20 simulations.

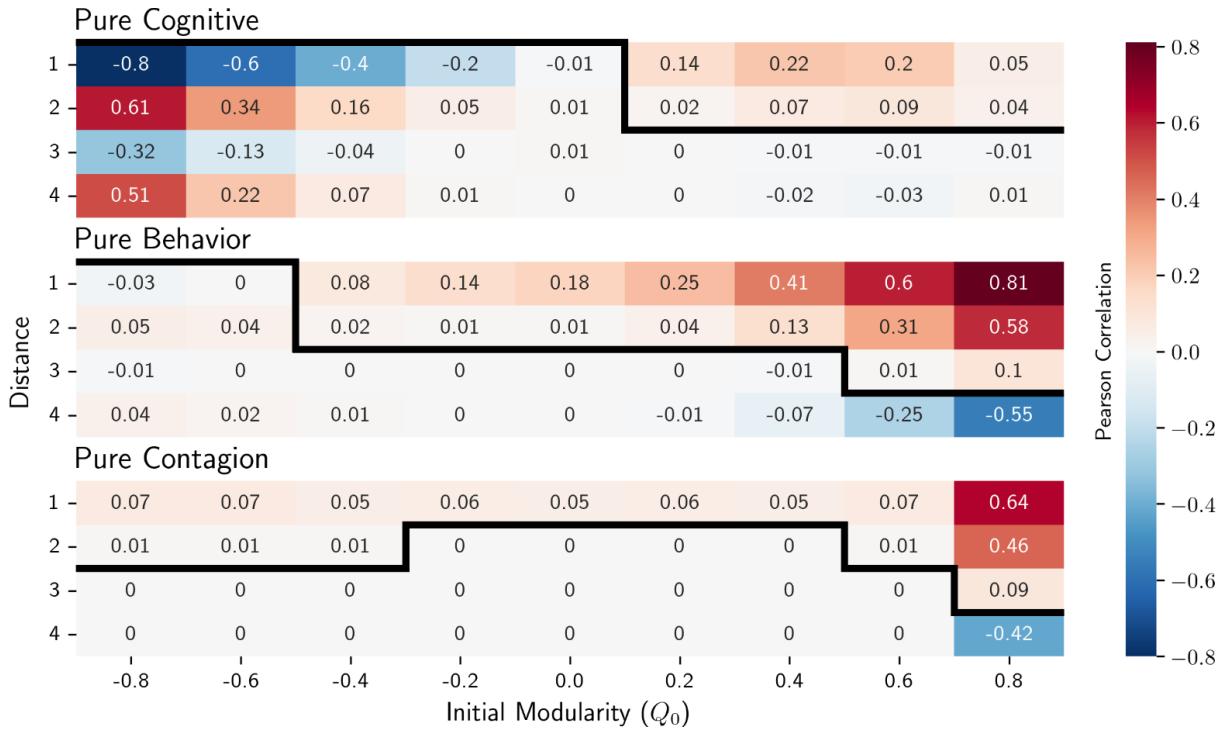


Figure 3: Pearson correlation matrix illustrating the correlations over distance for each induction pathway at various initial modularity levels. The black line represents the maximum degree of influence ( $\rho_m$ mDOI) and converged ( $\rho_t$ ) energy dynamics for the different levels of homophily and pathways. Each correlation value represents the mean correlation from 20 simulations.

of this decline in propensity over degrees of separation can be seen in Figure ??, where the maximum degree of separation is assessed whether we could replicate the empirical finding of 3 for a network with starting assortativity of .8 and purely influenced by the contagion pathway degrees of influence. Figure 2 depicts the Pearson correlations over distance for each inductive pathway, normalized on the first degree of separation (i.e., distance = 1). This illustrates that the model can replicate between two (i.e., the cognitive pathway) or three degrees of influence on a highly homophilic network ( $Q_0 = 0.8$ ).

- 3.3 We investigated how the variation of mDOI (maximum degree of separation varies influence) with different initial assortativity levels of modularity (representing homophily) for each induction pathway (Figure ??). We assessed the convergence of this metric over time, which provides. Figure 3 displays the correlations over distance for each initial modularity and induction pathway. The mDOI values are indicated by a dark black line, highlighting the size of the effect and the point at which the correlations intersect the zero line. Our findings reveal the presence of positive correlations over distance, with mDOI values ranging from 1 to 3 across all pathways. The correlations tend to increase as the initial modularity increases, while negative or zero initial modularity results in small correlations. For further insights into the correlation distributions over distance for each pathway across modularity, the stability of the clustering behavior in the network. We observe that the contagion pathway initially increases after initialization and then, given enough time, converges towards zero. The other pathways behave more consistently and have a positive maximum degree of separation when the starting assortativity is positive metric over simulation time, the model's sensitivity to network size, and the model's sensitivity to noise, please refer to Appendix B1, B2, B3, and B4, respectively.

## Convergence dynamics

### Energy dynamics

- 3.4 We recover the result that the ego's influence extends beyond direct connections and found up to three degrees of influence. The generalized nature of the degree of influence mDOI metric under positive

initial assortativity levels of homophily makes it challenging to differentiate between the various pathways and determine the causal mechanisms behind loneliness clustering based on it. Therefore, to gain a comprehensive understanding of the system, we examine the convergence of assortativity and analyze the energy dynamics of each pathway and their combinations.

## Clustering

- 3.5 To evaluate whether clustering is sustainable or emerges under specific initial conditions, we analyzed the relationship between the initial Pearson correlation and the converged state (Figure ??). Our results indicate that the behavioral and contagion pathways push assortativity to random mixing at negative initial assortativity while the cognitive pathway remains constant. At positive initial assortativity, the cognitive pathway exhibits a reduced clustering coefficient, with a peak at initial assortativity of .4 and moving towards zero at .8. In contrast, the behavioral and contagion pathways show slightly lowered converged compared to the positive starting assortativity. To illustrate the calculation of a network's assortativity, please refer to Appendix ???. We offer scatterplots for each pathway at various intervals while maintaining a constant initial assortativity.

## Energy over time

- 3.6 We analyzed the energy dynamics of each inductive pathway, individually and mixed, by examining the mean energy dynamics per subpopulation over time. Figure 4 displays the dynamics over time for each pathway, with line colors representing different initial assortativity levels of homophily. The first row depicts the energy development for each pathway acting alone over time. Each pathway exhibits unique dynamics, with the contagion pathway fully converging to a limit, demonstrates distinct dynamics. The cognition and contagion pathways converge rapidly to equilibrium, but with notable differences. The contagion pathway tends to converge towards the mean social energy, while the cognition pathway converging quickly to equilibrium, and the behavioral pathway displaying results in a shift of agents within each subpopulation, transitioning from high to low energy and vice versa. On the other hand, the behavioral pathway exhibits more complex behavior that depends on the initial assortativity regarding how it converges, influenced by the initial levels of homophily, which determines its convergence pattern.
- 3.7 The second row presents a scenario where one pathway is dominant, with 80% strength weight in the energy update process, while the remaining two are at 10% strength. In this case, When dominant, the cognitive pathway dominates the dynamics, but the convergence rate is slowed down when the contagion pathway is dominant. When the behavioral pathway is dominant, the energetic subpopulation converges towards maximum energy instead of a value between 0 and 1. cognition route still drives dynamics, except for the initially energetic subpopulation which moves towards full energy. When looking at the dominant contagion pathway, the convergence rate is slowed. The third row illustrates the combination of two equally active pathways, with the cognitive pathway dominating the dynamics taking over the dynamics when equally mixed with the other two. When the contagion and behavioral pathways are equally mixed, the contagion pathway tends to pull the system towards a high-energy toward the average energy values, except for when the initial level of homophily is high ( $Q_0 = 0.8$ ). Here, there seems to be a tipping point in the level of homophily that tends to pull the system toward a fully energetic state.

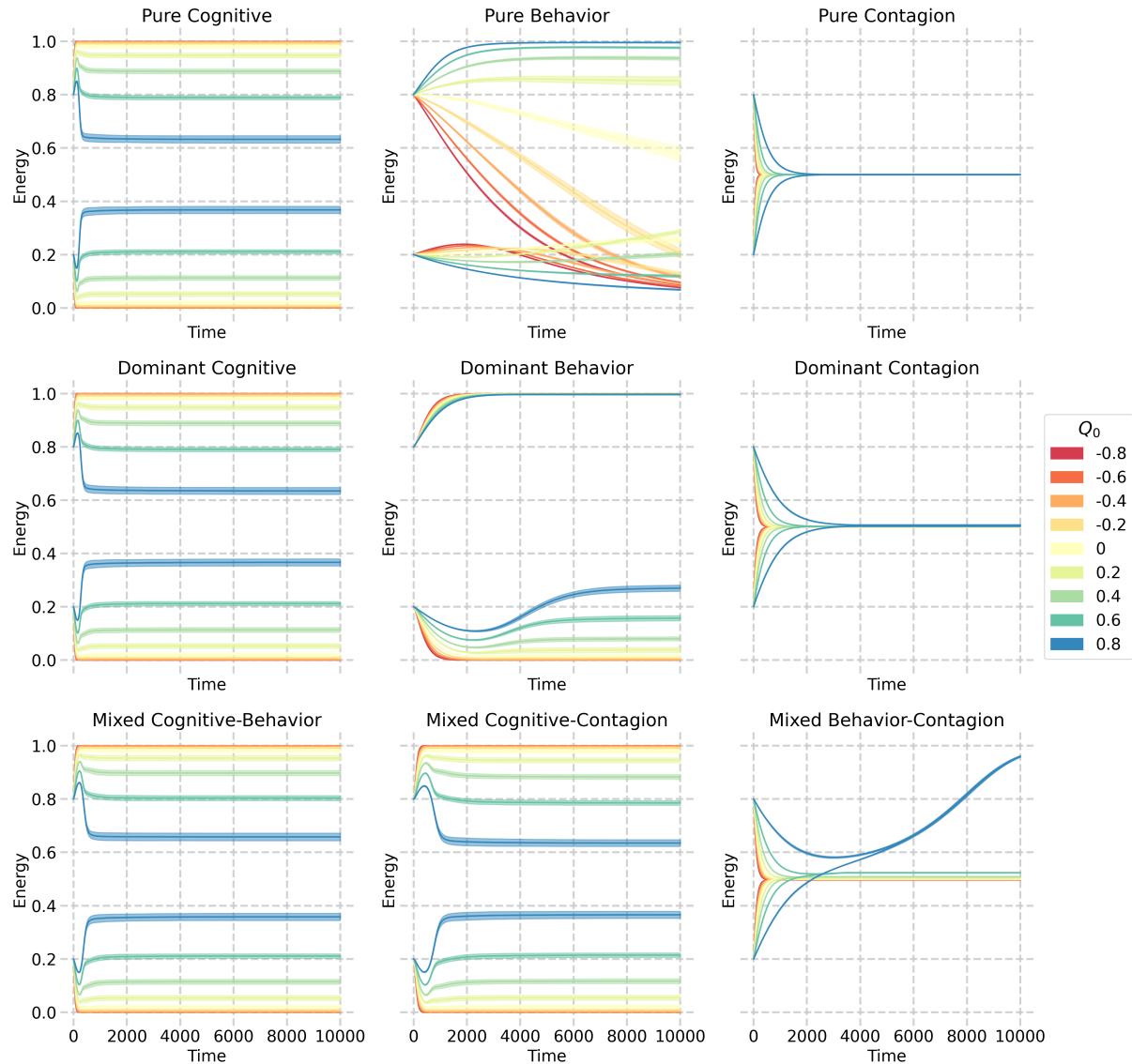


Figure 4: Mean energy dynamics per subpopulation over time for linear combinations of three inductive pathways. The line colors represent the starting **assortativity****modularity**. Each pane presents the mean and standard error of the means based on 20 simulations. The first row displays the purely active pathway, the second row shows a 50/50 mix of two pathways, and the third row depicts a scenario where one pathway is dominant at 80% while the other two are at 10%.

## ● Discussion

- 4.1 This study ~~investigated~~ explores the interplay between homophily and ~~the inductive pathways that lead to inductive pathways in the clustering of loneliness in within social networks~~. We found that ~~the three~~ Our findings replicate the degrees of influence found by ~~were universal for networks with positive assortativity for pathways not purely based on contagion processes~~. When these conditions are met, our results indicate that the cognitive and emotional contagion routes are the most important predictors of loneliness clustering, with the cognitive route being the most dominant. This universality of the three degrees of influence is a finding identified by ~~?~~, suggesting that these influences are stronger in networks with increasingly positive initial levels of homophily ( $Q_0 > 0$ ). These degrees of influence seem to emerge naturally, regardless of the underlying mechanisms or type of information being spread, which is further supported by ~~?~~, who used several standard models of information spreading to show that these patterns of correlation emerge naturally. Therefore, while these influence metrics are significant system features, they do not provide meaningful insights for disentangling causal relationships deriving from different mechanisms.
- 4.2 In addition, our study sheds light on the importance of homophily in predicting loneliness clustering. We found that inductive pathways were only effective in predicting loneliness clustering when there was a certain degree of homophily present. This highlights the need for future research to consider the role of homophily in understanding the social nature of mental pathology. Our findings indicate that the ability of inductive pathways to predict loneliness clustering is contingent upon a positive level of homophily. This aligns with prior studies showing how homophily fosters the formation of clusters or subgroups within larger social networks ~~?~~ leading to echo chambers, filter bubbles, and, ultimately, polarization ~~(?)~~. Such polarization can potentially reinforce specific mental states, such as depression and loneliness ~~(???)~~.
- 4.3 Furthermore, In our analysis of the energy dynamics of each inductive pathway revealed that each pathway individually exhibits unique dynamics. However, when mixing inductive pathways, the individual trajectories become hard to discern when the cognitive pathway is more active, as it seems more dominant. We provide a first step suggesting that perturbing a real-world system could be a fruitful approach to identifying which mechanisms are active. Upon disrupting the system and observing how it converges back, we could ~~the inductive pathways~~, we can categorize them as self-activating (i.e., the cognition route), complex contagion (i.e., the behavior pathway), and simple contagion systems (i.e., emotional contagion). The cognition route involves the emergence of loneliness based on individuals' expectations and perceptions of their social lives compared to others. Internal cognitive processes and beliefs act as self-activating factors in the experience of loneliness ~~(??)~~. This dynamic gives rise to a bistable system with two attractor states: a lonely state and a non-lonely state. Similar bistable dynamics are hypothesized to exist in other mental health issues, such as depression ~~(?)~~.
- 4.4 The behavior route involves inducing loneliness in others by exhibiting reduced trust and increased hostility once a loneliness threshold is surpassed ~~(?)~~. This aligns with complex contagion, which requires a critical mass of individuals within ones social network to adopt the behavior for extensive spread ~~(?)~~. The network topology can affect the critical mass required to spread competing opinions as it is lowered in modular networks compared to random networks that lead to a faster and more extensive spread of the behavior ~~(?)~~. Furthermore, the convergence speed is affected by the changes in modularity as wide bridges (i.e., many connections between communities) tend to speed up complex contagion as opposed to long bridges (i.e., few links formed with weak ties belonging to another community) having the opposite effect ~~(?)~~. Therefore, future research should include sampling more initial network topologies and examining their effects over different levels of modularity using the provided methodology.
- 4.5 Lastly, the emotional contagion pathway operates through a simple contagion process where non-verbal communication spreads and aligns mental states. Simple contagion involves spreading behaviors or beliefs with a single exposure or reinforcement. Emotional contagion occurs through the transmission of non-verbal cues and expressions, leading to a convergence of emotional states. This is consistent with the ~~?~~ model that describes information or behavior spread through social influence. We find that in this pathway, modularity tends to slow down the spread of the contagion by limiting connections between different communities which is in line with literature ~~(??)~~. Excessive modularity confines the contagion to individual communities, while inadequate modularity allows for rapid and widespread diffusion that saturates the entire network ~~(?)~~.
- 4.6 Lastly, the emotional contagion pathway operates through a simple contagion process where non-verbal communication spreads and aligns mental states, consistent with the ~~?~~ model of information spreading

through social influence. Higher modularity tends to slow down the spread of emotional contagion by limiting connections between different communities which is in line with literature (??). Excessive modularity confines the contagion to individual communities, while inadequate modularity allows for rapid and widespread diffusion that saturates the entire network (?).

- 4.7 We identify the cognition and emotional contagion routes as the main determinants of dynamical behavior over short and long periods. Each pathway exhibits unique energy dynamics, but the interaction between pathways complicates the identification of individual trajectories and the assessment of causal contributions. To gain insights into the underlying processes driving the clustering of loneliness. These findings suggest that the social environment could play a crucial role in the development of loneliness, highlighting the need for interventions that target both individuals and their social networks. contributing to loneliness clustering, future research could observe naturally perturbed real-world systems or investigate the formation of social networks based on subjective perceptions of loneliness, as explored by ?, where they aimed to investigate whether first-year college students form social networks based on subjective perceptions of loneliness.
- 4.8 It is important to note that our study Our study has several notable strengths. Firstly, we have developed a method that enables systematic sampling of homophily levels while maintaining the population's degree distributions. Future research should investigate the consequences of our homophilic network generation procedure, including changes in triadic closure and modularity through community detection, which might cause unforeseen interactions concerning homophily that could drive the observed dynamics ?. Additionally, the simplicity of our model allows for assessing the interplay between homophily and induction with its three different causal pathways. However, our study also has several limitations. First, we did not take into account heterogeneity within the agents in terms of different initial conditions and relationship strengths. Second, as we focused exclusively on heterogeneity related to connections and their influence on the evolution of dynamical states. We neglected other forms of heterogeneity that could be explored, such as different environmental influences as shared environments among individuals may play a significant role in cluster formation in social networks (??), and omitting this aspect could overemphasize the importance of homophily and induction. Investigating the role of heterogeneity in future work is essential, but caution must be exercised as it could obscure structural understanding (?). Furthermore, the three induction pathways in our model are linear combinations of one another each other, and a non-linear schema may provide a better fit. Third, we assumed that the entire population lived in a single environment, which allowed us to avoid the necessity of setting up distinct environments to test this process. Finally, our model is highly simplified, which limits our ability to make quantitative predictions about real-world social networks. might provide a more accurate fit.
- 4.9 Future research could address these limitations by exploring more complex models that better capture the dynamics of real-world social networks. For example, validation on the Framingham heart study using social network data containing loneliness data over time could help to calibrate and validate our model and its extensions. In addition, by understanding the inductive pathways and the role of homophily in social networks, future research could quantify these inductive influences on a per-person level. Interventions Finally, interventions that target specific individuals within social networks, such as agents who are always at maximum energy or targeted rewiring methods, could be tested to determine their effectiveness in reducing loneliness in the network. For example, this model could be used to test ? suggestion that interventions to reduce loneliness may benefit from targeting people on the periphery of a social network to help repair their social connections. Doing so may create a protective barrier against loneliness, keeping the whole network from unraveling.

4 Overall, our Our study provides insights into the social nature of mental pathology and the role of inductive pathways in the development of developing loneliness clustering in social networks. It underscores the possible importance of considering the social environment when studying mental pathology. Mental role of existing homophily in social networks as a supportive platform for inductive processes. As such, mental health cannot be fully understood by examining individual-level factors alone but must be situated within the broader social context in which it occurs. By doing so, we can develop a more nuanced understanding of the causes and consequences of mental health issues and develop more effective interventions to address them.

## ● Appendix A: Extended Methods

Henceforth, we will use the following notations in table 1 to denote the concepts used in this paper.

Table 1: Explanation of notations.

| Notation              | Symbol | Explanation  |
|-----------------------|--------|--|
| $i$                   |        | Node/agent identification index.   |
| $j$                   |        | Node/agent identification index of node initiating relationship towards $i$ .  |
| $e_i$                 |        | Energy level of agent $i$ , an abstraction of the amount of energy one has to engage in social activity and used as a proxy for loneliness in our model.                           |
| $k_i$                 |        | Connectivity level of agent $i$ , an abstraction of the how connected the agent feels to <u>its</u> <u>their</u> social surroundings.  |
| $\langle e_j \rangle$ |        | Average energy taken over the incoming neighborhood nodes.   |
| $\deg_i^+$            |        | Out-degree, or number of outgoing relationships of node $i$ .  |
| $\deg_i^-$            |        | In-degree, or number of incoming relationships of node $i$ .   |
| $\beta$               |        | <u>Connectivity decay rate.</u>  |
| $\rho$                |        | Pearson correlation, or the assortativity, of the network for the energy property. <u>Used to quantify induction effects by measuring the clustering of energy in the network.</u> |
| $\beta Q$             |        | <u>Connectivity decay rate</u> Modularity, the metric used to quantify homophily based on fixed labels of subpopulations.  |

## Operationalisation of homophily

The model incorporates homophily by initializing the network in converging a network to a fixed state of assortativity. Two scale-free communities are created, and the resulting network's relationships are either fully assortative or disassortative. The initial relationships are rewired randomly to modularity ( $-1 \leq Q \leq 1$ ). First, two identical Barabasi-Albert subpopulations are created that only differ in initial energy. The resulting network exhibits a complete separation of relationships between these subpopulations, maximizing modularity. To achieve the desired initial assortativity, increasing entropy within the system. This process slowly converges the network towards an assortativity value of zero, making the network random. modularity, the initial connections are indexed in each subpopulation and paired with identical counterparts in the other subpopulation. These pairs are then randomly rewired, introducing entropy into the system while preserving the degree distribution. This iterative process facilitates the network's progressive convergence towards modularity of -1, encompassing all intermediate values during the transition. Figure 5 provides a visual representation of this process, while figure 6 provides a representation depicting the interconnectedness of the subpopulations over different initial modularities.

The resulting initial topology remains fixed throughout the simulation, but the assortativity can change based on the energy property. Eliminating relationship formation, strengthening, and dissolution from the model reduces the need for additional nontrivial assumptions. Moreover, separating homophily from induction enables us to evaluate the dominant mechanism in the system and how these two mechanisms interact, as we can initiate networks with various assortativity coefficients can be initiated. Finally, assortativity is measured by calculating the Pearson correlation ( $\rho$ ) between the "from" and "to" sides of the relationships. modularity values.

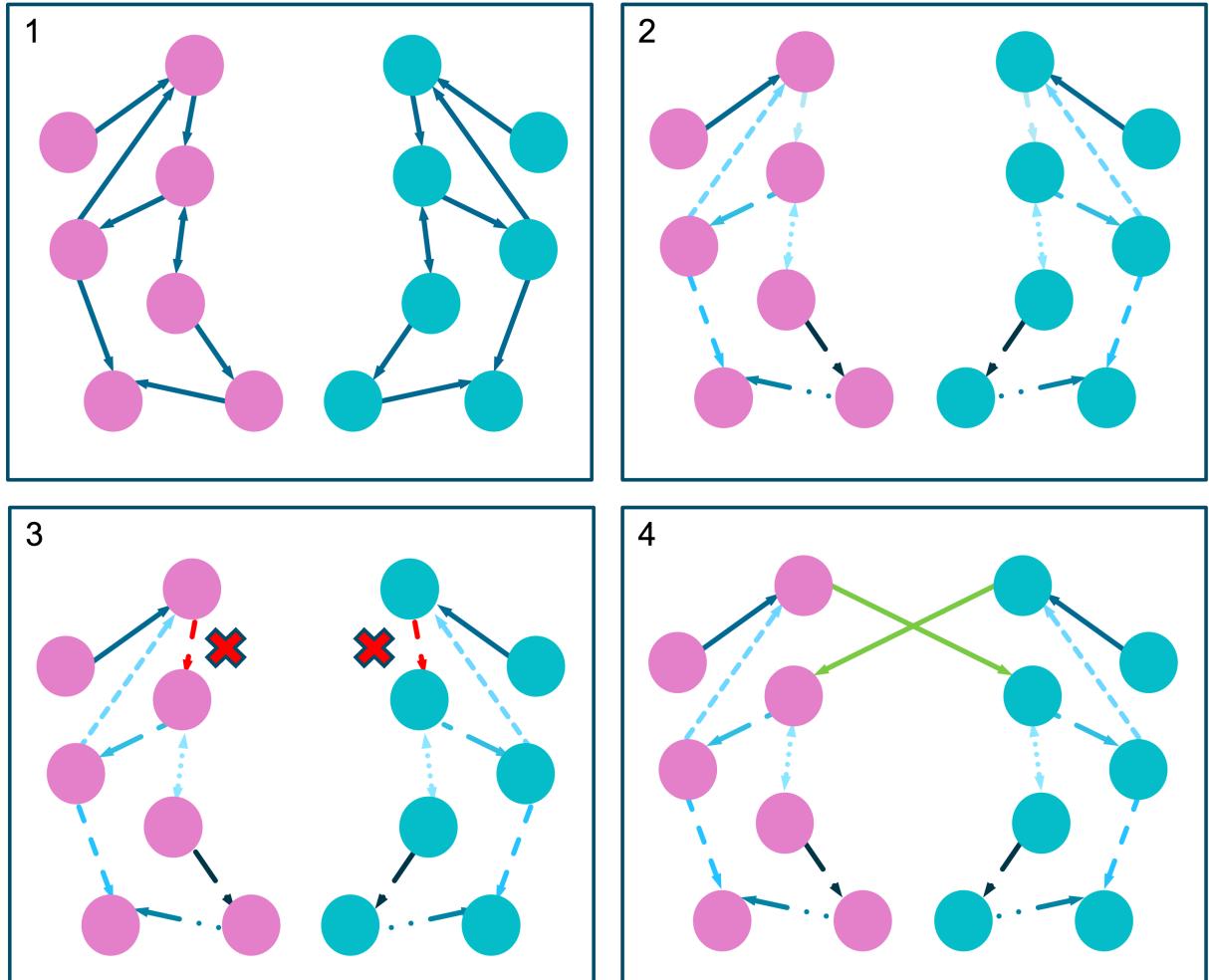


Figure 5: Schematic representation of a single iteration of the rewiring method. Two identical graphs are used as subpopulations ( $N=500$ ) with differing initial social energy. Identical links are paired up, and a random pair is selected for link removal. The removed link is then redirected to the other subpopulation.

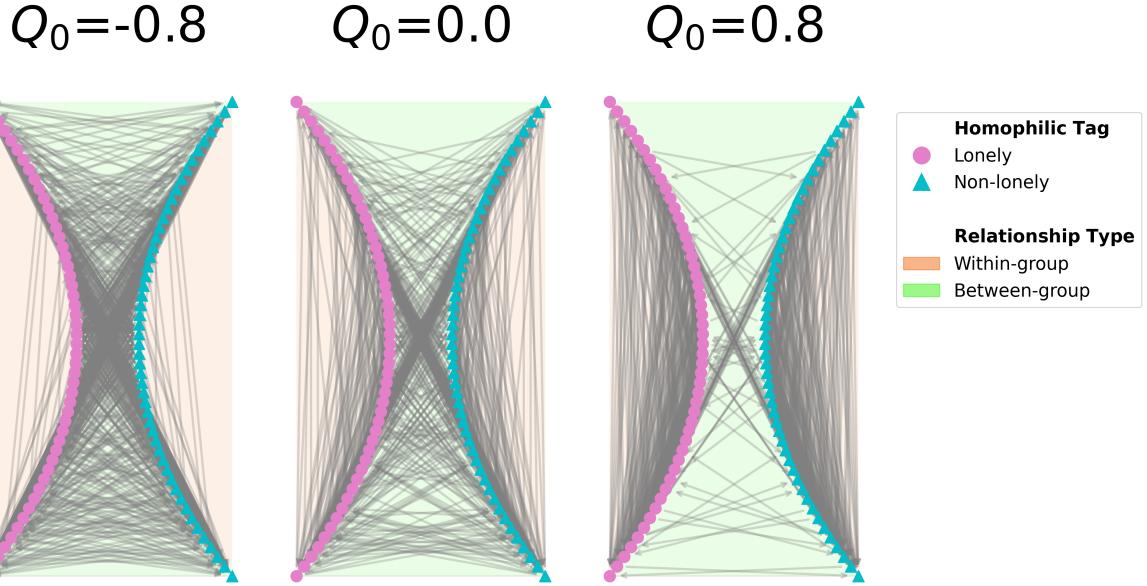


Figure 6: The network representation illustrates negative ( $Q_0 = -0.8$ ), neutral ( $Q_0 = 0$ ), and positive ( $Q_0 = 0.8$ ) modularity. Within the diagram, the red-shaded areas represent connections within groups, indicating relationships among nodes with similar attributes. Conversely, the blue-shaded areas represent connections between groups, denoting relationships among nodes with differing attributes. In addition to the shading, the nodes in the network exhibit different colors and shapes, signifying distinct homophilic tags.

## Operationalisations of the induction pathways

We describe the different pathways from an individual perspective, but individuals are embedded in a social network. Let  $a_{ij}$  be an adjacency matrix that identifies whether individual  $i$  is connected to individual  $j$ , for which case  $a_{ij} = 1$ , or they are not connected,  $a_{ij} = 0$ . At each time point,  $t$ , individuals update their internal state, including their social energy,  $e_i^t$ , based on different pathways. For convenience, we will ignore the time superscripts whenever it is unnecessary to specify the time point.

### Cognitive

The cognitive pathway states that loneliness is affected by a discrepancy between the expectation and the perception of one's social network. Interactions with others influence these expectations of what it means to have ideal social connectivity.

We have operationalized this process as an interplay between social energy  $e_i^t$  and perceived connectivity with one's own social network. We first describe the evolution of the,  $k_i^t$ . The change in that perceived social connectivity  $k_i^t$  for a node  $i$  over time, is given by

$$f_k(k_i^{t+1} - k_i^t) = e_i^t - k_i^t \beta. \quad (1)$$

The perceived social connectivity at time  $t + 1$  equals the social energy  $e_i^t$  at time  $t$  minus the product of the perceived social connectivity at time  $t$  and a constant parameter  $\beta$ . The social energy  $e_i^t$  represents the current level of social activity or engagement for node  $i$ , while the perceived social connectivity  $k_i^t$  represents the extent to which node  $i$  feels connected to its social environment.

In this system,  $\beta$  represents the rate at which a person's social connectivity decays over time without sufficient social energy to maintain it. The intuition is that social relationships can have a natural decay and that energy needs to be put into maintaining them. Specifically, suppose if a person's social energy

is lower than their perceived social connectivity multiplied by the decay rate (i.e.,  $\beta$ ). In that case low ( $e_i < k_i \beta$ ), their perceived social connectivity will gradually decrease as the social connections decay—decrease, and unless  $e_i$  increases,  $k_i$  will keep decreasing. A higher beta value implies a faster decay rate, meaning that a person's perceived social connectivity will decrease more rapidly over time without sufficient social energy. Conversely, a lower beta value implies a slower decay rate, meaning that a person's perceived social connectivity will decrease more slowly over time without social energy. If a person's social energy is high ( $e_i > k_i \beta$ ), perceived social connectivity will increase.

Social connectivity feeds back into social energy. The social energy at time  $t+1$  change in social energy of an agent  $i$  is determined by  $f_c^c$  in Eq.(2), the difference between the perceived social connectivity at time  $t$  and the mean perceived social connectivity of the nodes agent in node  $i$ 's neighborhood, denoted by  $\langle k_t^j \rangle$ . The term  $k_t^i - \langle k_t^j \rangle$  is  $\frac{\sum_j a_{ji} k_j}{\sum_j a_{ji}}$ . It represents the deviation of node  $i$ 's perceived social connectivity from the average connectivity of its neighbors (i.e., a proxy for expectations based on its neighbors).

$$f_c^c(i) = \frac{k_t^i - \langle k_t^j \rangle}{deg_i^-} k_i - \langle k_j \rangle_i \quad (2)$$

The term  $deg_i^-$  in the denominator represents the in-degree of node  $i$ , which is the number of incoming links to node  $i$  in a social network. This factor normalizes the social energy update for node  $i$  based on the number of connections it has to other nodes in the network.

## Behavioral

Individuals experiencing loneliness may display less trusting and more hostile behaviors toward others. These behaviors can negatively impact relationship satisfaction and may result in the loss or weakening of the relationship, subsequently inducing loneliness in others. To formalize this process, we propose the following formula: that the change in social energy of an agent  $i$  via the behavior pathway,  $f_b^b$ , follows

$$f_b^b(i) = \frac{e_j - \theta}{out_j} \left\langle \frac{e_j - \theta}{d_j^+} \right\rangle_i e_i, \quad (3)$$

In this equation,  $\langle \frac{e_j - \theta}{out_j} \rangle$  represents the fraction of where  $\langle \frac{e_j - \theta}{d_j^+} \rangle_i$  is the average social energy contribution from neighboring nodes  $j$  to  $i$ , calculated based on the number of others that  $j$  divides its time on (i.e., its out-degree),  $d_j^+ = \sum_l a_{jl}$ . An agent with a small out-degree can allocate more energy to node  $i$ , leading to a greater contribution. We subtract The value  $\theta$  as a threshold representing the represents a threshold point at which an agent is considered lonely and may exhibit less trusting and more hostile behavior. A value of  $e_j$  lower than  $\theta$  indicates a negative energy contribution from that neighbor, while a value of  $e_j$  higher than  $\theta$  indicates a positive influence on node  $i$ , contributing negatively to the social energy of its neighbors.

We introduce a weighting factor into the equation by multiplying  $\langle \frac{e_j - \theta}{out_j} \rangle$  with Further, the energy of the focal node  $i$ ,  $e_i$ . This weighting factor scales the effect above. This allows node  $i$  to modulate the impact of others based on its energy level, as individuals may exhibit "less trusting and more hostile" behaviors as energy levels decline. Consequently, the The value of  $e_i$  acts as a regulator, down-regulating the effect of others on the individual as the individual becomes less receptive to others.

## Emotional contagion

Individuals can experience a convergence of emotions through nonverbal communication. For example, loneliness can lead to shyness, social awkwardness, and hostility towards socially ambiguous cues, which could contribute to others in their immediate environment also experiencing loneliness. Emotional contagion is described as the "tendency for the facial expressions, vocalizations, postures, and

movements of interacting individuals to lead to a convergence of their emotions" (?). To model this process, we assume that agents converge towards the mean energy for the emotional contagion pathway upon interaction.

We propose the following equation to represent this process. The change in social energy of node  $i$  due to this process is given by the average social energy of the neighbors relative to that of the self:

$$f_{\text{ec}}(\underline{\underline{i}})^{\text{ec}} = \langle e_j \rangle_i - e_i \quad (4)$$

In this equation,  $\langle e_j \rangle_i = \frac{\sum_j a_{ji} e_j}{\sum_j a_{ji}}$  represents the mean energy of the neighborhood of node  $i$ . By subtracting the social energy of node  $i$  from the mean energy of its neighbors, we can calculate the average distance to the surrounding agents. We assume This implies that agents converge towards toward the mean energy of their neighborhood, as represented by  $\langle e_j \rangle_i$ , when they interact with others.

## Combining pathways

These pathways combine linearly to determine an individual's social energy level at time  $t+1$  ( $e_{t+1}^i$ ), which is also influenced by their energy level at time  $t$  ( $e_t^i$ ). The following equations summarize the system:

$$k_{t+1}^i - k_t^i = e_{t+1}^i - e_t^i - k_{t+1}^i \beta \quad (5)$$

$$e_{t+1}^i = (\left( p_1 f_c(e_t^i, k_t^i)^{\text{ct}} + p_2 f_b(e_t^i)^{\text{bt}} + p_3 f_{\text{ec}}^{\text{ect}} \right) + \text{norm}(e_t^i \mu, \sigma)) e_{t+1}^i (1 - e_{t+1}^i) p_3 \equiv (1 - (p_1 + p_2)) \quad (6)$$

Here, In our model, we have the equation  $p_1 + p_2 + p_3 \equiv 1$ , where  $p_1$ ,  $p_2$ , and  $p_3$  represent the three induction pathways. To account for inherent uncertainties and randomness, a noise term  $\text{norm}(\mu, \sigma)$  is added with  $\mu = 0$  and the  $\sigma = 0.02$ . Additionally, the function  $e(1 - e)$  in the equation is a logistic function that ensures the constrains the social energy level stays within the range of 0 and 1, ensuring a realistic representation of the system's dynamics.

## Parameter values

All parameter values used in the simulations can be found in table ??.

Table 2: An overview of parameter values used in the simulations.

| Parameter                    | Symbol  | Value | Comment   |
|------------------------------|---------|-------|---|
| Population size              | $N$     | 1000  | Number of agents in the simulation.   |
| Simulation duration          | $T$     | 4000  | Simulation duration in timesteps.   |
| Number of edges per new node | $m$     | 11    | Number of edges to attach from a new node to existing nodes in Barabasi-Albert algorithm. |
| Decay parameter              | $\beta$ | .5    | Decay strength of social connectivity.  |

Assortativity Initial Modularity  $\rho Q_0$

$$\begin{aligned} -1 &\leq \rho \leq 1 \\ -1 &\leq Q_0 \leq 1 \end{aligned}$$

Pearson correlation  
of the energy of  
"from" and "to"  
nodes as a measure  
of homophily Metric  
of homophily  
measuring the  
interconnection  
between  
subpopulations in  
the initial network.

---

## ● **Appendix B: Extended results**

### ● **Appendix B: Energies**

#### **Kernel Density Estimations of correlation distributions**

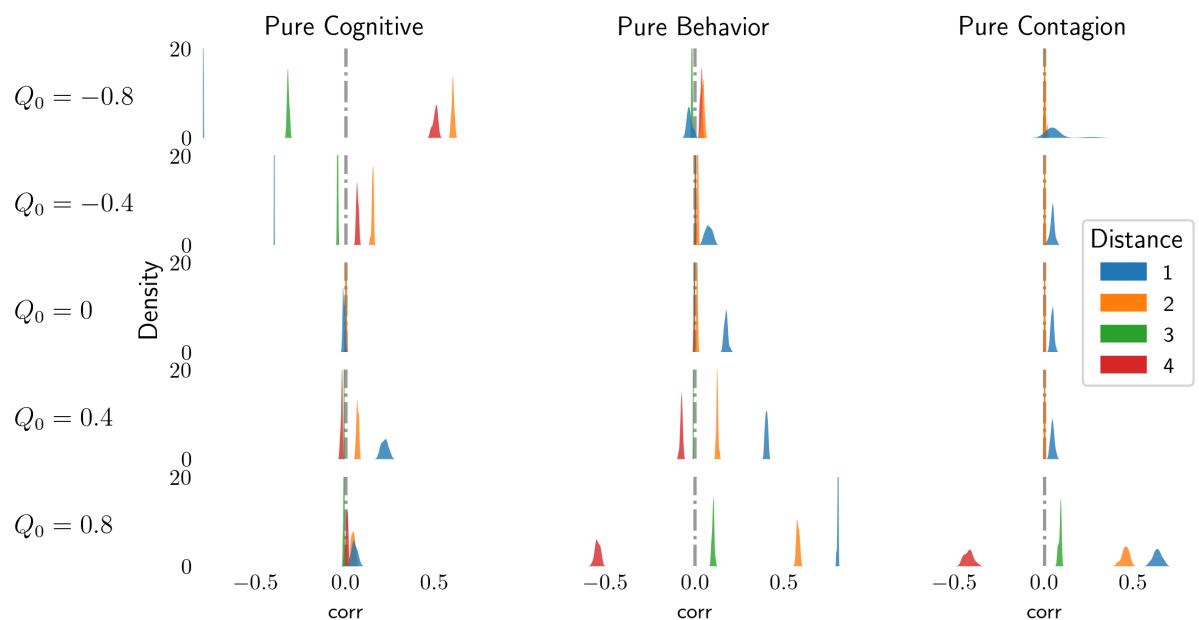


Figure 7: **Kernel Density Estimations of correlation distributions for each pathway over different initial modularities.** Each distance is plotted separately and colored accordingly. We can see that distribution estimations are narrow and increase as  $Q_0$  increases.

## Degrees of influence over time

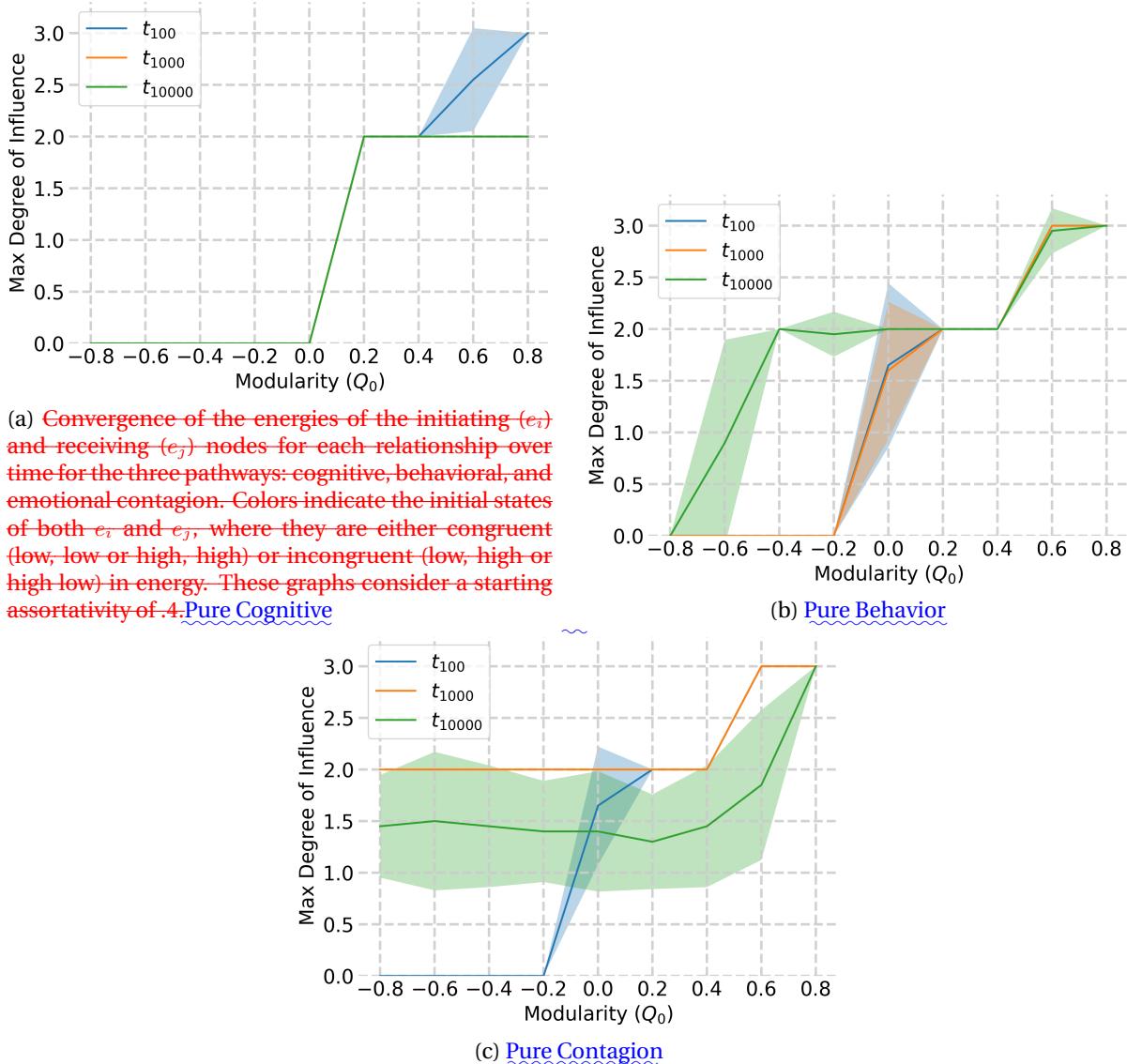


Figure 8: 8a), 8b), and 8c) illustrate the maximum degree of influence (mDOI) for different starting levels of homophily values and pathways over time (i.e., timestep 100, 1,000, and 10,000), providing insights into the stability of the clustering behavior in the network. The pure contagion and behavioral pathways display positive mDOI at negative initial modularities after 1000 timesteps, while the pure contagion pathway initially increases over simulation time while decreasing after convergence. Error bars indicate standard errors of the mean based on 20 simulations.

## Sensitivity of network size

<TODO> Just waiting for data to finish generating.

## Sensitivity of noise parameter

<TODO> Just waiting for data to finish generating.