



Department of Business and Management  
Bachelor's Degree in Management and Computer Science  
Chair of Social Network Analysis

## **Effects of Heterogeneity on Rumours' Diffusion in Networks**

Prof. Venel Xavier Mathieu Raymond

---

SUPERVISOR

Organtini Lorenzo

---

CANDIDATE

ID: 255501

Academic Year 2022/2023

## Table of Contents

<b>Acknowledgements</b> .....	1
<b>1. Introduction</b> .....	2
<b>2. Literature Review</b> .....	3
<b>3. The Stochastic SIR model</b> .....	5
<b>4. Statements and Comments on the Spread of Information</b> .....	8
4.1. Normal contagion in line-shaped networks of 3 nodes .....	8
4.2. Normal contagion in line-shaped networks of an arbitrarily odd large length .....	10
4.3. Normal contagion in star-shaped networks .....	12
4.4. Normal contagion in networks with reversing nodes .....	13
<b>5. Rumours</b> .....	15
<b>6. Statements and Comments on the Debunking of Rumours</b> .....	16
6.1. Diffusion of rumours in line-shaped networks of 3 nodes.....	16
6.2. Diffusion of rumours in line-shaped networks of arbitrarily large odd length.....	22
6.3. Diffusion of rumours in star-shaped networks .....	26
6.4. Simulations of line-shaped networks with reversing nodes .....	30
<b>7. Conclusions</b> .....	31
<b>8. The Code</b> .....	33
<b>References</b> .....	38

## Acknowledgements

First and foremost, I wish to express my deepest gratitude to my thesis supervisor, Professor Venel, for his support and guidance throughout the writing of this thesis. His willingness to assist me and his patience have been invaluable in the completion of this work.

To my mother and father, thank you for your continuous encouragement to better myself and for always believing in me. Without your teachings, I would not be who I am today. To my sister, thank you for your companionship and for bringing me joy. To Grandmother Marisa, thank you for being my number one fan; I know that you would be proud. To my grandmother Pia, thank you for your emotional support.

To all my friends, thank you for supporting me. A special thanks must be given to Alfredo, Emanuele, Sébastien and Max, with whom I spent this year in Paris. It has truly been the best year of my life so far.

Lastly, I want to thank my best friend, Paolo, for growing together with me and with whom I hope to create many more memories.

Thank you

## 1. Introduction

Rumours can be defined as pieces of unverified, or at least not entirely true, information that spread through a network. They can arise from a variety of causes and can spread quickly due to the complexities of social interactions. The fast propagation of information, due to the extensive use of the internet, has led to the widespread diffusion of those phenomena, which can have a significant impact on individuals, organisations and societies. Compared to true information, rumours can be, and are, debunked, as exposing false information is beneficial for society. As such, rumours can be modelled as two normal diffusion processes occurring at the same time, where one symbolises a false statement, while the other is the truth. Those two types of information are spread by different agents in a network as, usually, the ones who have an interest in spreading false statements do not want the truth to emerge. Moreover, it is safe to assume that the truth will start spreading through society only after the rumour is circulating and, more specifically, only after the agents who know the truth hear about the rumour. Understanding how diffusion processes in social networks work is crucial to drafting policies aimed at debunking or controlling the spread of rumours in order to mitigate their negative effects. In this study, we elaborate on the work of Akbarpour & Jackson [1] by focusing on modelling and analysing rumours propagation in line-shaped and star-shaped networks, using the Susceptible-Infected-Removed (SIR) model, with nodes that follow Markov chains to set activity levels that define whether or not they can transmit the information (both the rumour and the truth).

The SIR model is a widely used tool in epidemiology for studying contagion processes. It categorizes the population as Susceptible (S), Infected (I), and Recovered (R): susceptible nodes are the agents not yet infected; infected nodes the ones that are infected and transmit the phenomenon; while recovered nodes are those that were infected and have now gained immunity, but are not capable to spread the phenomenon anymore. The model is then used to study the spread of a single piece of information or disease that travels through a network. Therefore, to model the spread of rumours and their debunking process, the model must be adapted to feature two pieces of information, one false and the other true, travelling through the network. For this reason, we call *spreaders* those who actively spread the rumour, and *debunkers* those who actively try to debunk it.

Additionally, to make the model more realistic, we must take into consideration the fact that not all nodes are always able to transmit information or disease at the same time. A node is

an entity that can transmit and receive information - in our case, an individual in the population. The traditional SIR model presumes a homogenous population, where everyone has the same chance of getting infected or spreading the disease. However, this does not reflect the complexities of the real world. For example, consider an organization that operates in day and night shifts, with a brief period of overlap to ensure a smooth transition. A rumour spread by a day-shift employee can only reach a night-shift worker during that short window of interaction when they are both present in the building. To model this behaviour, we use the same framework developed by Akbarpour & Jackson [1] to define the activity levels of nodes in the social network. By adopting their approach, the model acknowledges the presence of activity patterns observed in human behaviour, such as the burst of activity followed by periods of inactivity, by assigning to nodes a Markov chain that dictates whether or not the node is able to spreading or receiving information. As a result, information can only propagate when both the transmitter and the receiver are active simultaneously. Moreover, by assigning Markov chains to nodes in the SIR model, we can get useful insights into how activity patterns influence the spread of rumours.

By taking advantage of the Markovian SIR model, we aim to understand how activity patterns play a role in the diffusion of rumours, contributing to the more effective management of information propagation in the digital age. This improved understanding can help in the development of more effective policies and actions to counteract the negative effects of rumours on society.

The outline of the work is the following: Section 2 reviews previous research in the field, to gather fundamental knowledge for future evaluations; Section 3 introduces the model that is used in this work, developed by Akbarpour & Jackson [1]; Section 4 shows the statements done by Akbarpour & Jackson [1] for the spread of information in a network; Section 5 and Section 6 presents our work on rumours; and Section 7 states the conclusions. Additionally, at the end of the document, there is a paragraph explaining the code used for simulations.

## **2. Literature Review**

In a society, agents are heterogeneous and do not behave in all the same way. People can assume many different behaviours, like the burst one, documented in the diffusion processes such as emails or web browsing [2]. Indeed, when sending or answering emails, people

either dedicate an entire time slot for the day, or they check them frequently for smaller intervals of time.

The focus of this study is on the activity pattern of nodes, and on how it impacts the diffusion of true and false information. Nodes can be either active or inactive, depending on their ability to spread or receive information, to more effectively capture the intricate complexity of our society. We assume also that individuals have the same average activity level, as generally everyone is active for the same amount of time during a day. Coming back to the email example presented earlier, the two behaviours can be represented as two activity patterns: one in which a node is active for long periods of time; and the other in which a node is active with a certain probability for each unit of time. The latter is an activity pattern that follows a Poisson behaviour, where an agent is active with a certain probability during each period. However, because people do not act randomly during the day, studying networks in which nodes follow activity patterns that are time or history-dependent, i.e. do not only adopt the Poisson behaviour, yields more accurate results than applying homogeneous models. This has been demonstrated in the diffusion of email worms by Vazquez et al (2007) [3] in which the discrepancy between the simulations and the reality can be explained by the failure of the Poisson activity pattern to represent how individuals communicate. For such networks, new measures are possible, such as temporal distances (and paths), which although correlated with the static distances, have a wide spread [4].

There are several models to describe activity patterns. In this study, we use the model of Akbarpour & Jackson [1], whose aim was to find the implications of having heterogeneous nodes in a network for the spread of information. The model used is a modified version of the SIR model, where each node follows a Markov chain to define its ability to spread the phenomenon. They define three types of nodes, with different activity patterns, that however have the same average activity level, to represent the fact that individuals are usually active for the same amount of time in a day but in different ways. The authors then concluded that having a mix of nodes which follows different activity patterns is beneficial to the transmission of the information. Similar work was done by Laijun et al (2011) [6], with a modified SIR model in an attempt to better represent human behaviours. In this network, nodes can also assume the role of *hibernators*, which can remember the rumour after forgetting it, and susceptible nodes can become recovered without passing through the infected state. Those two added features should represent individuals who start spreading the

news when they are reminded of it and individuals who, thanks to their background, do not fall for the rumour.

Stochastic models in rumours analysis have been used also in other ways than to determine activity levels. In their work, Isham et al (2009) [7], considered the change of state between susceptible, infected and recovered in a SIR model as a Markov chain. In this model, the probabilities to change state depend not only on the current state of a node itself, but also on the states of all the nodes to which it is connected. By applying this concept to the spread of rumours, they derived a set of equations for determining the size of a rumour on a homogeneous network.

However, studying the debunking of rumours rather than their spread can be more interesting as while it is difficult to prevent rumours to spread, it is useful to draft policies to debunk them quickly and efficiently. Understanding how to represent this process in a network can help to create better strategies to inform quickly the population. Merlino et al (2022) [8] focused their paper on policy implementation by observing the spread of a true and a false message in a network. They used an SIS model in which nodes can decide to verify the truth of a rumour and found that the truth-to-rumour ratio depends on the level of verification and degree of homophily in the network. They concluded that incentivising individuals to verify information is the most likely to succeed as it can be shown that increases in verification rates drastically increase the truth-to-rumour ratio. Other studies on competing information in a network were done by Tabasco [9] and Campbell et al (2019) [10], with similar findings although not considering debunkable rumours.

Regarding different models than the SIR, research was done on DK models, developed by Daley & Kendall [11], and on MK models, proposed by Maki & Thomson [12], which however are not able to describe rumour spreading on large-scale networks. Thresholds in the rumour spreading rate below which a rumour cannot be spread were defined by Nekovee et al (2006) [13] for ER random graphs (and others) with forgetting mechanisms, and are significantly different from results obtained with simpler models [14-17].

### **3. The Stochastic SIR model**

The model used to study the impact of the different timing of activity patterns of nodes is a variation of the SIR (susceptible, infected, recovered) model. In addition to the status of node (susceptible, infected, recovered), to each node in the model can be active or inactive.

The infection can only be spread if both the infected (the sender) and the susceptible (the receiver) nodes are active at the same time. Once a node is infected, it can spread the phenomenon for  $T$  periods, after which the node stops being contagious and becomes recovered. We assume that nodes are either active or inactive in any period, with the same probability to be active on average of  $\lambda \in (0,1)$ . To model activity pattern, we assume that every node follows a Markov chain determining if it is active tomorrow as a function of today. The different behaviours observed in a population are represented by changing the probabilities of switching state in the Markov chain. *Figure 1* shows the general Markov chain, valid for any node:

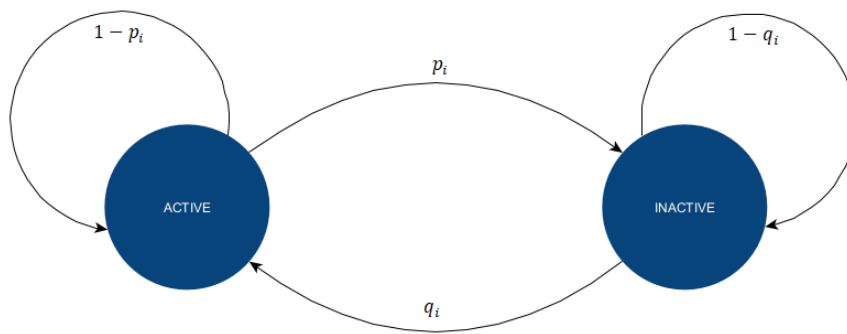


Figure 1: General Markov Chain

An agent  $i$  can switch from inactive to active with a probability of  $q_i$ , or stay inactive with a probability of  $1 - q_i$ . Similarly, it can switch from active to inactive with a probability of  $p_i$ , or stay active with a probability of  $1 - p_i$ . Because every node, regardless of its type, should be active with a probability of  $\lambda$  on average,  $\lambda$  should be the probability to be active at the steady state. At any point in time, this probability is the sum of the probability to stay active, given that in the previous period the node was already active, plus the probability to become active, given that in the previous period the node was inactive:

$$\lambda(t) = P(\text{Active} | \text{Active}) \cdot \lambda(t - 1) + P(\text{Active} | \text{Inactive}) \cdot (1 - \lambda(t - 1))$$

since at the steady state  $\lambda$  does not depend on time, we can rewrite the equation as:

$$\lambda = P(\text{Active} | \text{Active}) \cdot \lambda + P(\text{Active} | \text{Inactive}) \cdot (1 - \lambda) = (1 - p_i)\lambda + q_i(1 - \lambda)$$

Thus, rearranging the equation, the balance equation of the model is:

$$\lambda \cdot p_i = (1 - \lambda)q_i$$

Some useful re-writings of this equation are:



$$\frac{\lambda}{1-\lambda} = \frac{q_i}{p_i} \quad \lambda = \frac{1}{1 + \frac{p_i}{q_i}} \quad q_i = p_i \cdot \frac{\lambda}{1-\lambda}$$

The types of nodes mentioned above, differ in the values of the probabilities  $q_i$  and  $p_i$ , but all nodes have the same  $\lambda$  at the steady state. We consider three types of nodes in order to create heterogeneous populations. A *sticky* node, named after the sticky behaviour, has both  $q_i$  and  $p_i$  close to 0, and thus its Markov chain can be heuristically approximated to a Markov chain where it is not possible to switch state [Figure 2]:



Figure 2: Markov Chain of a Sticky node

In this way, this type of node is almost perfectly auto-correlated over time.

A *reversing* node, named after its tendency to change state, has the maximal possible  $q_i$  and  $p_i$ , so that it has the maximal negative auto-correlation. As a result, if  $\lambda \leq \frac{1}{2}$  then  $p_i = 1$  and  $q_i = \frac{\lambda}{1-\lambda}$ . On the other hand, if  $\lambda \geq \frac{1}{2}$  then  $p_i = \frac{1-\lambda}{\lambda}$  and  $q_i = 1$ , consequentially in the case in which  $\lambda = \frac{1}{2}$ , the agent will switch state every period. Thus, the state of a *reversing* agent is as negatively serially correlated as possible, moving between active and inactive states as frequently as possible.

Finally, a *poisson* node, named after the Poisson distribution that it follows, is one that has  $p_i = 1 - \lambda$  and  $q_i = \lambda$  and thus is active every period with a probability of  $\lambda$ .

Notice that, because of the way the Markov chain is defined, knowing the state of the node at time  $X$  does not give any information on the state of the node at time  $X + 1$ . Moreover,  $\lambda$  does not assume the values of 0 and 1 as, for those two values, *sticky*, *poisson* and *reversing* nodes act in the same way.

To understand how heterogeneity of nodes influences the spread of information or rumours, those types of nodes will be concatenated to form line-shaped or star-shaped networks through which a phenomenon will spread.

#### 4. Statements and Comments on the Spread of Information

When considering networks of nodes that follow an activity pattern, new important results arise. Mainly, the optimal configuration of nodes that maximizes the diffusion of a phenomenon, has mixed types of nodes. Indeed, heterogeneity can substantially increase the speed and the reach of diffusion processes. To prove this claim, we will recreate and comment on four of the five propositions illustrated by Akbarpour & Jackson [1].

We show the results of those propositions with some insights on their proof, using the same notation they have used. Thus, the nodes' type will be indicated by the capital letter *S* for *sticky*, *P* for *poisson* and *R* for *reversing*.

In the following, statements 1 to 4 are taken from the work of Akbarpour & Jackson [1].

##### 4.1. Normal contagion in line-shaped networks of 3 nodes

Statement 1: In a line of three nodes, which once infected can transmit for a number of periods  $T \in \mathbb{N}$ , the configuration of *poisson* and *sticky* nodes that maximises both the probability that all nodes become infected and the expected number of infected is:

- *PSP* if  $\lambda < \lambda^*$  and;
- *PPP* if  $\lambda > \lambda^*$

Where  $\lambda^*$  is the unique solution in (0,1) for

$$\lambda = f(\lambda) = \left( \frac{1 - (1 - \lambda^2)^T}{1 - (1 - \lambda)^T} \right)^2 \quad (1),$$

if we are maximizing the probability of total infection and

$$\lambda = \frac{[1 - (1 - \lambda^2)^T][2 - (1 - \lambda^2)^T]}{[1 - (1 - \lambda)^T][2 - (1 - \lambda)^T]} \quad (2),$$

if we are maximizing the expected number of infected nodes. □

In a line of three nodes, if the agents have a high activity level, the best configuration is to have everybody following the *poisson* behaviour. On the other hand, heterogeneous sequences are preferred if the agents have low activity level.

To find the optimal sequences that maximise the spread of a disease, notice that it is always better to have a *poisson* node as the first node of a line. If the second node is *sticky*, the probability of it becoming infected is  $\lambda^2$ , if the first one is *sticky*, and  $\lambda(1 - (1 - \lambda)^T) > \lambda^2$  if the first one is *poisson*. The same result holds if the second node is *poisson*. Thus, we can rule out all combinations starting with a *sticky* node.

Then, we can look at the probability of a second node, adjacent to an infected one, to become infected. The possible combinations of nodes are: *PS*; *SP*; *SS* and; *PP*. For *PS*, the infection spreads only if, the *sticky* node is active, and the *poisson* node is active in at least one of the  $T$  periods in which it is contagious. The probability to be active for the *sticky* is  $\lambda$  as if it is active in the first period it will remain active. The probability for the *poisson* node to be active at least once is just 1 minus the probability of never being active, which is  $(1 - \lambda)^T$ . Thus:

$$PS = \lambda \cdot (1 - (1 - \lambda)^T)$$

The same reasoning is applied to the configuration *SP*, however as we are considering that the first node is infected, we must consider the probability to be active as 1 because the node needed to be active for it to become infected. Therefore:

$$SP = 1 - (1 - \lambda)^T$$

By setting the probability to be active of the first *sticky* node to 1, the probability of the configuration *SS* will therefore be:

$$SS = \lambda$$

Lastly, the probability for the combination *PP* is 1 minus the probability that the infection cannot spread. This happens either when the first is active and the second is not or the other way around, or when both are inactive at the same time. Thus:

$$PP = 1 - (\lambda(1 - \lambda) + \lambda(1 - \lambda) + (1 - \lambda)^2)^T = 1 - (2\lambda(1 - \lambda) + (1 - \lambda)^2)^T = 1 - (1 - \lambda^2)^T$$

Note that,  $SP > PP > PS$  and  $SP > SS > PS$ . Now,  $PP > PS$  implies that  $PPP > PPS$  and  $SP > SS$  implies that  $PSP > PSS$  since a configuration of three nodes can be split

into two subunits of two nodes that behave accordingly to the probabilities calculated above. Moreover, the inequalities hold as the first subunits of two nodes are the same for both  $PPP > PPS$  and  $PSP > PSS$ , making the comparison effectively only on the second subunit. Because other orderings of three nodes yield a lower probability of total infection, due to *sticky* nodes having a lower probability to spread the disease to a second node, the maximal layout in terms of overall infection is either  $PPP$  or  $PSP$ . Equation (1) is the result of the equation  $PPP = PSP$  as to compare the two configurations. The solution of this equation is unique as the function  $f(\lambda)$  has a unique fixed point in  $(0,1)$ . Indeed, consider the functions  $g(\lambda) = (1 - (1 - \lambda)^T)^2$  and  $h(\lambda) = \frac{g(\lambda)}{\lambda}$ . Because  $g(\lambda)$  is strictly increasing at the beginning and then becomes strictly decreasing with a unique inflexion point, as  $g'(0) = g'(1) = 0$ ,  $g'(\lambda) > 0$  for  $\lambda \in (0,1)$ ,  $g''(0) > 0$  and  $g''(1) = 0$ , with  $g'''(\lambda) < 0$ , then  $h(\lambda)$  is monotonically increasing before the inflexion point, and monotonically increasing after. Therefore, as also  $\lim_{\lambda \rightarrow 0} h(\lambda) = 0$  and  $h(1) = 0$ , the equation  $h(\lambda) = h(\lambda^2)$ , equivalent to  $\lambda = f(\lambda)$ , will have a unique solution, denoted as  $\lambda^*$ .

The proof for the expected number of infected is similar to this one. The main result of this first proposition is that *poisson* nodes are good receivers while *sticky* nodes are good senders. Indeed, the conditional probability  $SP$  is the highest among the others.

#### 4.2. Normal contagion in line-shaped networks of an arbitrarily odd large length

This statement is an extension of the previous one.

Statement 2: In a line composed of an arbitrarily odd large number of nodes, who once infected can transmit for a number of periods  $T \in \mathbb{N}$ , the configuration of *poisson* and *sticky* nodes that maximises the probability of total infection is:

- $PSSSS...SP$  if  $\lambda < \lambda^{**}$ ;
- $PSPSP...SP$  if  $\lambda^{**} < \lambda < \lambda^*$  and;
- $PPPPP...PP$  if  $\lambda > \lambda^*$

Where  $\lambda^*$  was defined in the previous proposition,  $\lambda^{**}$  is the interior solution of  $\lambda = [1 - (1 - \lambda)^T]^2$  and  $0 < \lambda^{**} < \lambda^* < 1$ . □

The proof for this statement is just an extension of the first statement, so we will not detail the proof. However, it is interesting to run some simulations of this case in order to verify the proposition. Those simulations are programmed in Python where nodes are objects and the nodes' types are classes. We compare the results of the infection processes for lines of five nodes in which all of the nodes are *poisson* to one in which nodes alternate between *poisson* and *sticky* (starting from *poisson* as they are better senders). The nodes stay contagious for 2 turns, after which they remain infected but can no longer spread the infection. 50,000 iterations were run over a range of 30 lambdas. To show the result, the ratio of the probabilities to get more than 50% of the population infected or a full infection for the alternating and *poisson* sequences are used, as it allows us to see which combination performs better and by how much.

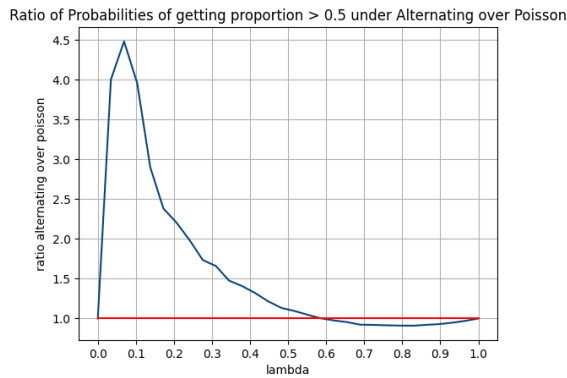


Figure 3

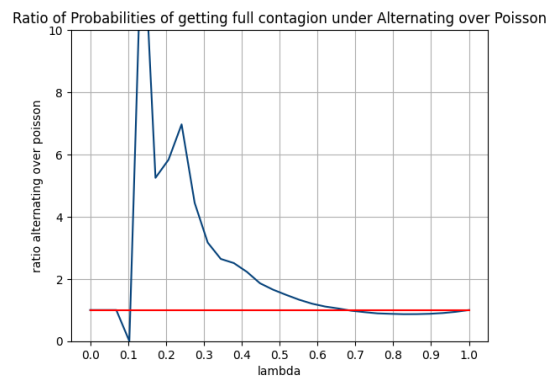


Figure 4

In the graphs for those ratios, the red line shows the indifference point between the two configurations of nodes. Above the red line are cases in which we prefer heterogeneity while below are cases in which we prefer homogeneity. We can observe that, for small values of lambda, having *poisson* and *sticky* is far better for the spread of the disease, further confirming the results of the previous statement. For larger values, the difference is quite negligible even though the ratio falls below 1, which indicates that the layout will all *poisson* performs better as predicted by the proposition. From the graph, we can see that the benefits of having always a heterogenous population far outweigh the disadvantages. Indeed, even when a homogeneous population is preferred, the ratio is barely below 1, and thus if we need to choose a population without knowing  $\lambda$ , it is always better to choose a heterogeneous one.

### 4.3. Normal contagion in star-shaped networks

Statement 3: In a star-shaped network with  $n$  leaves, with nodes that can transmit the disease for  $T = 2$  periods, the configuration that maximises the probability of a full contagion for any  $\lambda \in (0,1)$  is one such that all leaves are *poisson* and the centre is *sticky*.  $\square$

In a star-shaped network with a large number of leaves, the preferred configuration is heterogeneous, characterized by a *sticky* centre and *poisson* leaves, regardless of the activity level of the agents. Indeed, consider a star-shaped network with  $n$  leaves. In this configuration, a leaf is infected and can transmit the disease for two periods. Then, for any  $\lambda$ , the layout which maximises the probability of full contagion has a *sticky* node as the centre and *poisson* nodes as the leaves. This can be easily proved by remembering that *poisson* nodes are good receivers and *sticky* nodes are good senders. Having a *sticky* node in the centre implies that for  $n - 1$  leaves, the *sticky* node will be the sender and the *poisson* nodes the receivers, provided that the centre gets infected. Calculating the theoretical probabilities of full contagion leads to the same conclusion. Indeed, if the star has a *poisson* node as a central node then this probability will be:

$$P_P: (1 - (1 - \lambda^2)^2)(2\lambda(1 - \lambda)\lambda^{n-1} + \lambda^2(1 - (1 - \lambda)^2)^{n-1})$$

Computed by multiplying the probability of the first contagion between the leaf and the centre and, because  $T = 2$ , the sum of the event in which the centre and the leaf are active and inactive or vice versa, and the event in which they are both active at the same time. On the other hand, if the star has a *sticky* node as its centre:

$$P_S: (\lambda(1 - (1 - \lambda)^2)(1 - (1 - \lambda)^2)^{n-1} = \lambda(1 - (1 - \lambda)^n)$$

Calculated by considering the interactions *PS* for the contagion between the leaf and the core, and *SP* for the contagion between the core and the leaves. In this last case, it is possible to use the probabilities calculated in [statement 1](#) as the *sticky* node does not have a chance to change its activity. With the two probabilities, it is possible to show that the limit as  $n$  goes to infinity of the ratio between  $P_P$  and  $P_S$  is less than 1 for any  $\lambda$ . Indeed:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{P_P}{P_S} &= \lim_{n \rightarrow \infty} \frac{(1 - (1 - \lambda^2)^2)(2\lambda(1 - \lambda)\lambda^{n-1} + \lambda^2(1 - (1 - \lambda)^2)^{n-1})}{\lambda(1 - (1 - \lambda)^n)} \\ &= \frac{1 - (1 - \lambda^2)^2}{2 - \lambda} \end{aligned}$$

Which is less than 1 for every  $\lambda \in (0,1)$ . Thus, also when changing the shape of the network, having a heterogenous network is beneficial to the diffusion of a phenomenon.

#### 4.4. Normal contagion in networks with reversing nodes

Statement 4: When introducing the *reversing* node type in a line graph, with the same characteristics as those in *statement 2*, then the optimal configurations involves always *reversing* nodes and never *poisson* ones. Those optimal layouts will be either a full alternation of the form  $RSRSR\dots SR$  for low levels of  $\lambda$ , all *reversing* nodes for high levels of lambda and a mixture of *RRR*s and alternating  $SRSR\dots SR$  in between.  $\square$

When considering *reversing* nodes, if agents have a high activity level, the best configuration is to have everybody following the *reversing* behaviour. On the other hand, for middle or low activity levels, heterogeneous sequences are preferred. The whole formal proof of this statement will not be reported here, however, some hints of its proof are.

Firstly, the likelihood of transmission is increased when a *sticky* node is followed by a *reversing* node. The chance of transmission from an infected *sticky* to a *reversing* node, for instance, is 1 if  $\lambda > \frac{1}{2}$ . Indeed, if the *reversing* node is inactive during one period, it will be active during the following period and make contact during at least one of the two periods. As a result, the main disadvantage of alternating *reversing* and *sticky* nodes is using the latter as receivers, which becomes a problem for large values of  $\lambda$  thus it is preferable to use *reversing* nodes exclusively.

Then, to see why the optimal configurations do not present any *poisson* nodes notice that the probability that a *poisson* node is never active is  $(1 - \lambda)^T$ , while for the *reversing* node is  $(1 - \lambda)(1 - q)^{T-1}$  with  $1 - q < 1 - \lambda$ , since a *reversing* node has a lower probability of staying inactive than its overall probability of being in that state, as it prefers to switch. In particular,  $q = \min\left\{1, \frac{\lambda}{1-\lambda}\right\}$ , and so  $1 - q = \max\left\{0, \frac{1-2\lambda}{1-\lambda}\right\}$  with  $\frac{1-2\lambda}{1-\lambda} < 1 - \lambda$  for  $\lambda > 0$ . This implies that having  $RS$  at the start of the line always beats having  $PS$ . Similarly, having  $SR$  at the end of the line always beats having  $SP$ . Therefore, it is always best to have a *reversing* node as the beginning and ending node and this is in line with *statement 1*, where *poisson* nodes were better than *sticky* as first components. These facts imply that  $RSPS\dots PSR$  and  $RSSS\dots SSR$  dominate  $PSPS\dots PSP$

and *PSSS...SSP* respectively, when  $\lambda < \lambda^*$ . Thus, *poisson* nodes can only be found on the interior of the graph. However, by calculating the probability that a subsequent node gets infected given that the previous one was already infected, it can be shown that every interior sequence with *poisson* nodes gets dominated by one with only *reversing* and *sticky*. Those probabilities are calculated using the same reasoning showed in [statement 1](#) and [statement 2](#).

This result can also be shown empirically with simulations. For them, 50,000 iterations in which we test six combinations of five nodes in a line over the span of 30 values for  $\lambda$ . The metric to evaluate the performance of those layouts is the probability to reach a certain threshold of contagion. Also for this simulation, the contagion time is 2 and the infection starts randomly at one end. The result of this simulation is shown in [Figure 5] and [Figure 6]:

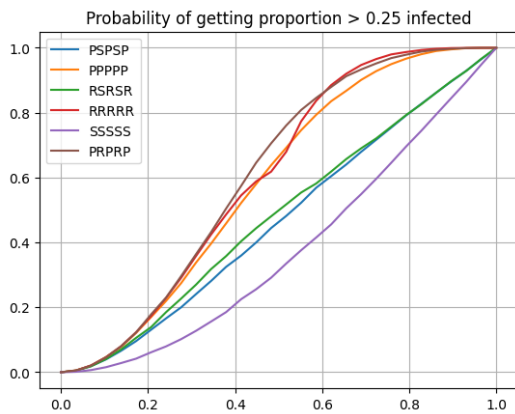


Figure 5

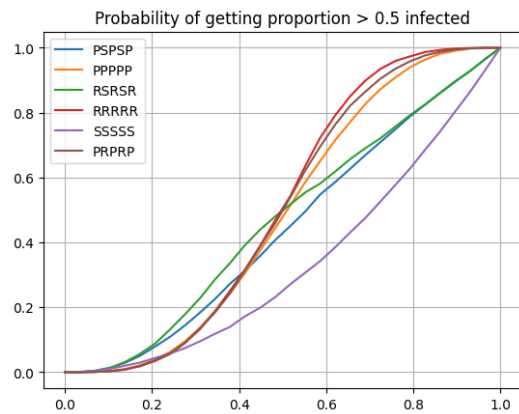


Figure 6

From the graphs, it is clear that the layouts with only *reversing* or a combination of *sticky* and *reversing* nodes are the ones which performed the best. Moreover, looking at [Figure 6] we can confirm that alternating *reversing* and *sticky* nodes is best before a certain threshold of  $\lambda$ , after which having all *reversing* become beneficial.

In conclusion, we can affirm that heterogeneity in a network helps the diffusion of a phenomenon, as it boosts the probability of infection. Thus, when the aim is to maximise the diffusion process, like when the phenomenon is a piece of information, then having a diverse network is of help. On the other hand, when the objective is to minimise the number of infected nodes, an homogeneous structure is preferred, in particular one composed of



*sticky* nodes, which by far reduces the probability of infection under all metrics and outperforms all the other combinations.

## 5. Rumours

The kind of static information seen so far is not the only one that can travel through a network. More complex kinds of information can be transmitted, for example in the form of rumours. A rumour is defined as a false circulating story, spread by some node. Understanding the dynamics of how rumours spread and how they can be debunked is crucial in real-life scenarios, particularly in the context of organizations, where misinformation can have significant consequences on brand reputation, customer trust, and even financial performance. For instance, consider a pharmaceutical company that faces a rumour regarding the safety of one of its products. The rumour spreads quickly through social media, leading to widespread fear and reduced sales. By understanding the underlying network structures and the roles of *spreaders* and *debunkers*, the company can swiftly implement a targeted communication campaign to counteract false claims.

The model used to study this new type of information features *spreaders*, which are the nodes that spread the false claim, *debunkers*, which are the nodes who know that the information is false and will try to debunk it, and simple nodes. *Debunkers* behave by spreading a counter-information in the network as soon as they hear the rumour, by following the same rules as the normal contagion process, converting neighbours to their version of the story and turning them into *debunkers*. Therefore, this counter-information is transmitted by nodes for  $T$  periods, and can spread only if both the *debunker* and the receiver are both active. Once a node has heard that the information circulating is false, it cannot get infected and will no longer be able to spread the false information, but will start spreading the true one. Moreover, everyone in the network still acts according to the Markov chains defined before and can be either *sticky*, *poisson* or *reversing*.

As those types of phenomena are present in real-life networks, it can be interesting to understand whether or not the heterogeneity of nodes helps with the contagion or not. As now information needs to travel in the network twice, the previous optimal structures for diffusion may no longer be optimal for debunking purposes. Indeed, new assumptions are possible considering that if the debunking process starts, *sticky* nodes that were active during the contagion process, are active with probability 1, and hence can play a key role in

exposing the truth. This knowledge can be invaluable in designing effective strategies for organizations to combat the spread of misinformation, promote fact-checking, and maintain their brand image.

To investigate those new systems we first present theory and expectations, then confirming the results by running simulations. By doing so, we aim to shed light on the complex interaction between various factors influencing the spread and debunking of rumours, which can have practical applications in managing misinformation and fostering a more informed and responsible digital landscape for organizations and their stakeholders.

## 6. Statements and Comments on the Debunking of Rumours

The purpose of this section is to verify that heterogeneity is beneficial to debunking a rumour by checking if the statements for the spread of information apply also in this specific case. Even if heterogeneity makes a rumour spread faster, and thus reaching a *debunker* faster, it can be detrimental to the debunking process. As *poisson* nodes act with random behaviour, if we consider the probability of a rumour to get debunked knowing that it reached a *debunker*, the network would be better off with all *sticky* nodes, as the probability of fully debunking the rumour is 1, challenging the idea that heterogeneity is beneficial also for spreading the truth. As a matter of fact, in this section, we show how, for certain thresholds of  $\lambda$ , homogeneity is preferable.

The following statements are analogous to statements 1 to 4 that we explained in the previous section, but in the presence of rumours, which have been found in this work. For the first three statements a formal proof is provided, while for the last one only simulations are shown.

### 6.1. Diffusion of rumours in line-shaped networks of 3 nodes

*Statement 1B*: In a line of three nodes, who once infected can transmit for a number of periods  $T \in \mathbb{N}$ , and where the first node is a *spreader* of a rumour, while the last node is a *debunker*, the configuration of *poisson* and *sticky* nodes that maximises the probability of fully debunking a rumour is:

- $SSS_{rumour}$  for  $\lambda < \lambda^{**}$
- $PSP_{rumour}$  for  $\lambda^{**} < \lambda < \lambda^{***}$  and

-  $PPP_{rumour}$  for  $\lambda > \lambda^{***}$

Where  $\lambda^{**}$  was defined previously and  $\lambda^{***}$  is the unique solution in (0,1) for:

$$\lambda = \left( \frac{1 - (1 - \lambda^2)^T}{1 - (1 - \lambda)^T} \right)^4.$$

Moreover, the thresholds satisfy:  $0 < \lambda^{**} < \lambda^* < \lambda^{***} < 1$ . □

Compared to normal contagion, *sticky* nodes have clearly gained an advantage over *poisson* nodes, making homogeneity beneficial, not only for high activity levels, but for low activity levels too. This can be observed by the presence of the sequence  $SSS_{rumour}$  and by the fact that the threshold between  $PSP_{rumour}$  and  $PPP_{rumour}$  has shifted to the right as  $\lambda^* < \lambda^{***}$ . In the context of rumours, *sticky* nodes gain this advantage because of their activity pattern. To understand why, consider that, to predict the outcomes of simulations, we can view this problem as a simple contagion process with lines of 5 nodes. This is because, after the rumour reaches the third node (the *debunker*), the truth has to travel through two additional nodes for the rumour to be completely debunked. Notice that, due to this mirroring, some *sticky* nodes will be correlated as the first and second nodes are the same as the fourth and fifth. For example, if the second *sticky* node of a line is active, then the fourth one will be too. Those correlated nodes are indicated with  $S_n$ , where  $n$  indicates the position of the one fixing their activity level. On the other hand, new *poisson* nodes can be considered as independent nodes because of their activity pattern. Consequently, sequences with *sticky* nodes gain an advantage from the correlation.

Because of this advantage, having a *sticky* node as a starting or ending node is no longer a dominated strategy. To see why, notice that the probabilities to debunk a rumour at the start of a line are:

- $SS_{rumour} \dots = SS \dots S_2 S_1_{contagion} : \lambda^2;$
- $SP_{rumour} \dots = SP \dots PS_1_{contagion} : \lambda(1 - (1 - \lambda)^T)^2;$
- $PS_{rumour} \dots = PS \dots S_2 P_{contagion} : \lambda(1 - (1 - \lambda)^T)^2;$
- $PP_{rumour} \dots = PP \dots PP_{contagion} : (1 - (1 - \lambda^2)^T)^2.$

Depending on lambda, a different initial sequence will be better.  $SS_{rumour} \dots$  is better than  $SP_{rumour} \dots = PS_{rumour} \dots$ , if and only if  $\lambda > (1 - (1 - \lambda)^T)^2 = g(\lambda)$ . From [statement 1](#) and [statement 2](#), we know that the function  $g(\lambda)$  is initially convex and then

concave, with a unique fixed point in  $(0,1)$ . Thus  $\lambda > g(\lambda)$  if and only if  $\lambda < \lambda^{**}$ , where  $\lambda^{**}$  is the solution to the equation  $\lambda = g(\lambda)$ . On the other hand,  $PP_{rumour} \dots$  is better than  $SP_{rumour} \dots = PS_{rumour} \dots$ , if and only if  $\lambda < \left(\frac{1-(1-\lambda^2)^T}{1-(1-\lambda)^T}\right)^2 = f(\lambda)$ . As already proven in *statement 1*, the function  $f(\lambda)$  has a unique fixed point in  $(0,1)$ , thus  $\lambda < f(\lambda)$  if and only if  $\lambda > \lambda^*$ , where  $\lambda^*$  is the solution to the equation  $\lambda = f(\lambda)$ . Since from *statement 2*,  $\lambda^{**} < \lambda^*$ , the order of optimal initial sequences is:

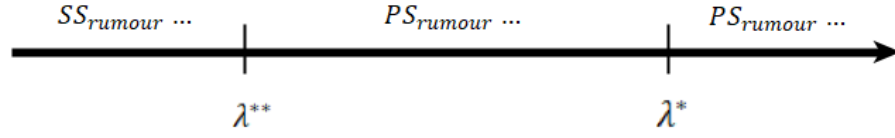


Figure 7: progression of initial sequences when the second node is sticky

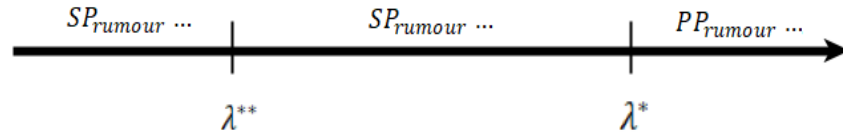


Figure 8: progression of initial sequences when the second node is poisson

Thus making *sticky* nodes a viable starter.

Next, to understand what are the optimal sequences of 3 nodes, we need to define the probabilities to debunk a rumour after the start. For this, we assume that the first node is infected for the contagion process, while the second node is counted as a *debunker* for the debunking process. In the contagion process, as the first node is assumed to be infected, any first *sticky* node is considered also active, as its activity level is set at time 0. Similarly, in the debunking process, as it happens after the contagion, any *sticky* node (regardless of the position), is considered active. On the other hand, because *poisson* nodes set their activity level each turn, they are never considered active with probability 1, and behave according to their Markov chain for contagion and debunking processes. This way of defining the probability to fully debunk a rumour works not only for ending sequences, as they have an actual debunker, but also for intermediate sequences, as when the contagion reaches the end, the truth will start spreading, thus making new debunkers in the line.

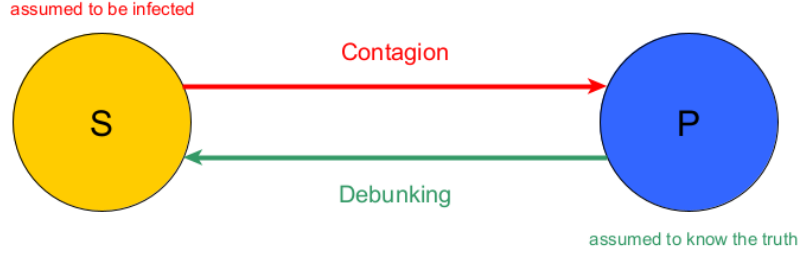


Figure 9: probability of middle and end sequences

Hence, these probabilities are:

- $SS_{rumour}: \lambda;$
- $SP_{rumour}: (1 - (1 - \lambda)^T)^2;$
- $PS_{rumour}: \lambda(1 - (1 - \lambda)^T)^2;$
- $PP_{rumour}: (1 - (1 - \lambda^2)^T)^2.$

In this way, the probability of fully debunking a rumour in a line can be calculated as the probability of fully debunking the rumour for the initial sequence, times the probability for any other -intermediate or ending- sequence. As an example, the sequence  $PSP_{rumour}$  can be divided into two subunits  $PS_{rumour} \dots$  and  $SP_{rumour}$ , so the probability of fully debunking a rumour is  $PS_{rumour} \dots \cdot SP_{rumour}$ . To define the configurations that maximise this probability for lines of three agents, we extend the intermediate sequences by adding a *sticky* or a *poisson* node at the beginning as it is no longer possible to exclude the presence of *sticky* nodes at the start. Because of this, we need to analyse every possible combination of three nodes and exclude the dominated sequences. The possible sequences that can be made with three nodes are:

- $SSS_{rumour}: \lambda^3;$
- $PSS_{rumour}: \lambda^2(1 - (1 - \lambda)^T)^2;$
- $SSP_{rumour}: \lambda^2(1 - (1 - \lambda)^T)^2;$
- $PSP_{rumour}: \lambda(1 - (1 - \lambda)^T)^4;$
- $SPS_{rumour}: \lambda^2(1 - (1 - \lambda)^T)^4;$
- $PPS_{rumour}: \lambda(1 - (1 - \lambda)^T)^2(1 - (1 - \lambda^2)^T)^2;$
- $SPP_{rumour}: \lambda(1 - (1 - \lambda)^T)^2(1 - (1 - \lambda^2)^T)^2;$
- $PPP_{rumour}: (1 - (1 - \lambda^2)^T)^4.$

Notice that,  $SSS_{rumour} > PSS_{rumour} = SSP_{rumour}$  for  $\lambda < \lambda^{**}$ , while  $PSP_{rumour} > PSS_{rumour} = SSP_{rumour}$  for  $\lambda > \lambda^{**}$ , making  $PSS_{rumour}$  and  $SSP_{rumour}$  always

dominated by other sequences. A similar argument can be used to exclude the combinations  $PSP_{rumour}$  and  $SPS_{rumour}$  as they are dominated by  $PSP_{rumour}$  for  $\lambda < \lambda^*$ , and are dominated by  $PPP_{rumour}$  for  $\lambda > \lambda^*$ . Moreover,  $SPS_{rumour}$  is always dominated by  $PSP_{rumour}$  as  $\lambda(1 - (1 - \lambda)^T)^4 > \lambda^2(1 - (1 - \lambda)^T)^4$  for any  $\lambda \in (0,1)$ . In conclusion, we expect the configuration that maximises the number of debunked nodes to be:

- $SSS_{rumour}$  for  $\lambda < \lambda^{**}$
- $PSP_{rumour}$  for  $\lambda^{**} < \lambda < \lambda^{***}$  and
- $PPP_{rumour}$  for  $\lambda > \lambda^{***}$

Where  $\lambda^{**}$  is the unique solution to:

$$\lambda = g(\lambda) = (1 - (1 - \lambda)^T)^2$$

And  $\lambda^{***}$  is the unique solution to:

$$\lambda = k(\lambda) = \left( \frac{1 - (1 - \lambda^2)^T}{1 - (1 - \lambda)^T} \right)^4$$

To obtain the equations presented, one must compare the probabilities of the three chains that are:

$$SSS_{rumour} = SS_{rumour} \dots \cdot SS_{rumour} = SSSS_2 S_{1_{contagion}}: \lambda^3$$

$$PSP_{rumour} = PS_{rumour} \dots \cdot PS_{rumour} = PPS_2 P_{contagion}: \lambda(1 - (1 - \lambda)^T)^4$$

$$PPP_{rumour} = PP_{rumour} \cdot PP_{rumour} = PPPP_{contagion}: (1 - (1 - \lambda^2)^T)^4$$

To prove that the solutions of the equations are unique, it is necessary to prove that the functions  $g(\lambda)$  and  $k(\lambda)$  have a unique fixed point. As already proven in [statement 1](#),  $g(\lambda)$  has a unique fixed point.

To prove that the other equation has a unique solution, notice that finding the fixed point of  $k(\lambda)$  is equivalent to solving  $l(\lambda) = l(\lambda^2)$ , where  $l(\lambda) = \frac{(1 - (1 - \lambda)^T)^4}{\lambda}$ . Since  $l(\lambda)$  is monotonically increasing for  $\lambda$ 's smaller than its interior inflexion point and monotonically increasing after, with  $l(1) = 0$  and  $\lim_{\lambda \rightarrow 0} l(\lambda) = 0$ , the equation must have a unique solution in the interval  $(0,1)$ .

Next, notice that when  $g(\lambda) = \lambda$ , then  $g(\lambda^2) < \lambda^2$ , since  $g(\lambda) < \lambda$  for  $\lambda < \lambda^{**}$ . This implies that  $\lambda > f(\lambda) > k(\lambda)$ , which means that  $\lambda < \lambda^{***}$  and so  $\lambda^{**} < \lambda^{***}$ . Moreover,

$\lambda^{***} > \lambda^*$  as  $k(\lambda) < f(\lambda)$  and thus  $k(\lambda^*) < \lambda^* = f(\lambda^*)$ , proving the shift of the threshold between  $PSP_{rumour}$  and  $PPP_{rumour}$ .

To confirm the theory, we performed 50,000 simulations for different levels of lambda and of every combination of three nodes and compared them to a normal contagion process. The results of them are shown below:

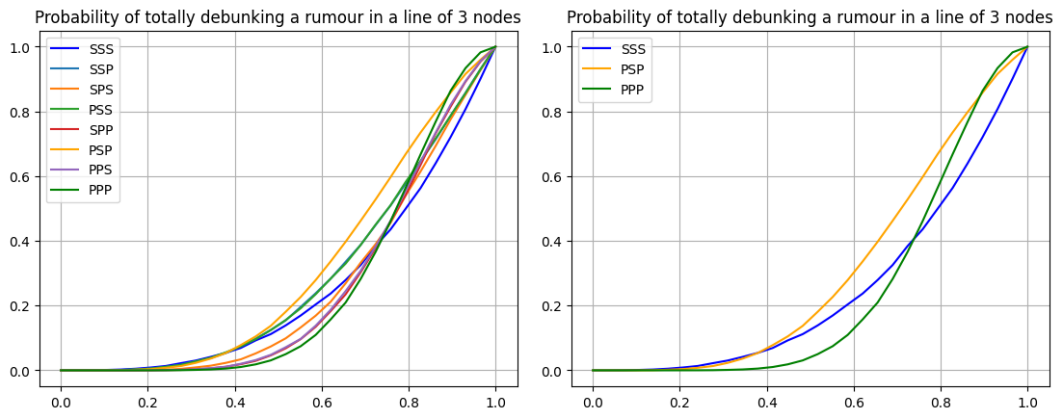


Figure 10: probabilities of total debunking

Figure 11: focus on the best combinations for debunking

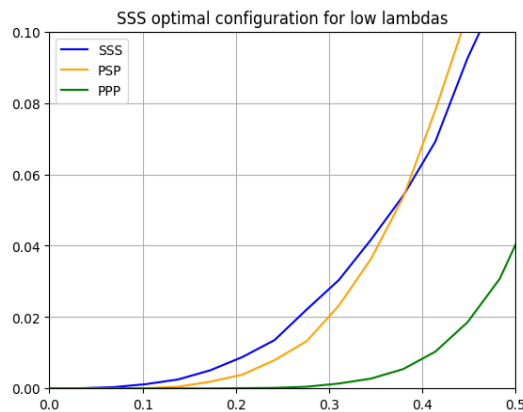


Figure 12: SSS for low lambdas

As expected, the empirical results confirm the theory. Another interesting fact about the scenario is that the optimal configurations seem to be the same as the original [statement 1](#) except for the presence of the line SSS at the very beginning. Indeed, the best configurations for the spread of a piece of information are PSP and PPP were:

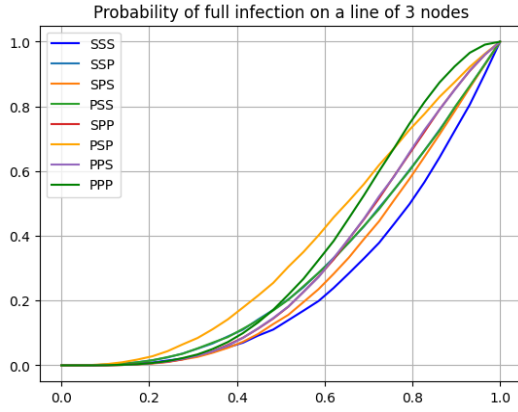


Figure 13: probabilities of full infection

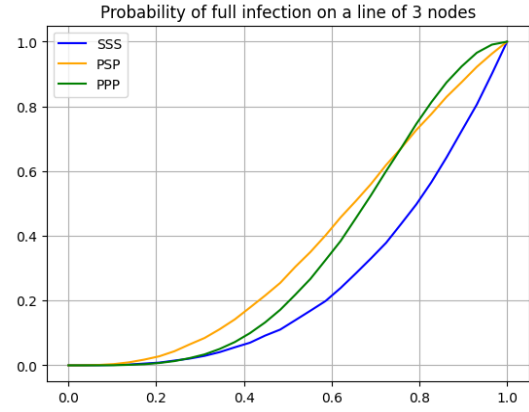


Figure 14: focus on the three combinations

This difference is because when the information needs to travel back if the network has only  $S$ , they will be active with probability 1, and thus shortening the line to an equivalent of 5 to a line of 3 nodes.

After establishing a theoretical framework and conducting simulations, it is crucial to verify that the theory holds up against the simulated data, as to ensure the validity of the model. By calculating the theoretical  $\lambda^{**}$  and  $\lambda^{***}$  for a model where  $T = 2$ , it is found that they are very close to what the simulations are showing. Indeed,  $\lambda_{simulated}^{**}$  and  $\lambda_{simulated}^{***}$ :

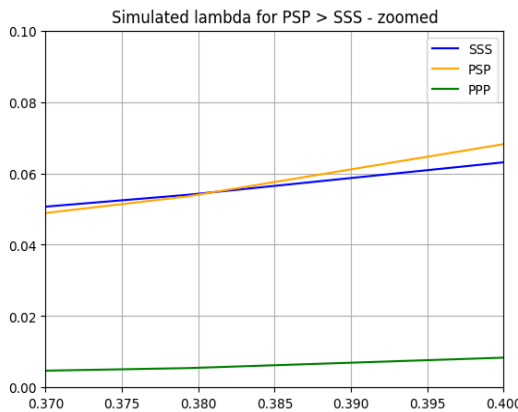


Figure 15: simulated level of lambda for PSP > SSS

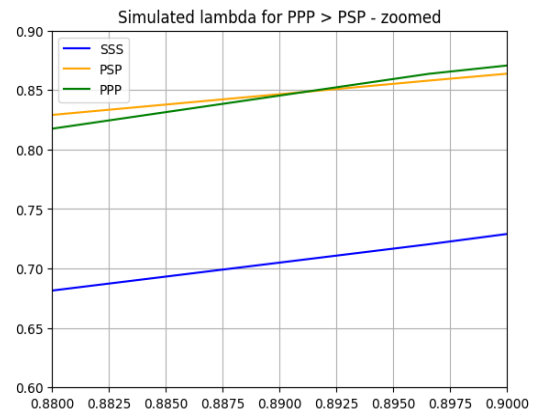


Figure 16: simulated level of lambda for PPP > PSP

are very close to  $\lambda_{theoretical}^{**} = 0.38197$  and  $\lambda_{theoretical}^{***} = 0.90166 > \lambda_{theoretical}^* = 0.75996$ .

## 6.2. Diffusion of rumours in line-shaped networks of arbitrarily large odd length

As for normal contagion processes, this statement is an extension of the previous one.



Statement 2B: In a line composed of an arbitrarily large finite odd number of nodes, where the first node is a *spreader* of a rumour, while the last node is a *debunker* and, who once infected can transmit for a number of periods  $T \in \mathbb{N}$ , the configuration of *poisson* and *sticky* nodes that maximises the probability of fully debunking a rumour is:

- $SSS \dots SSS_{rumour}$  for  $\lambda < \lambda^{**}$ ;
- $PSS \dots SSP_{rumour}$  for  $\lambda^{**} < \lambda < \lambda^{****}$ ;
- $PSP \dots PSP_{rumour}$  for  $\lambda^{****} < \lambda < \lambda^{***}$ ;
- $PPP \dots PPP_{rumour}$  for  $\lambda > \lambda^{***}$ .

Where  $\lambda^{**}$  and  $\lambda^{***}$  were defined in the previous propositions,  $\lambda^{****}$  is the interior solution of  $\lambda = (1 - (1 - \lambda)^T)^4$  and  $0 < \lambda^{**} < \lambda^{****} < \lambda^{***} < 1$ . □

The length of the line analysed in the previous statement may have a significant impact on the presence of homogeneous sequences. Therefore, it is crucial to extend the analysis to lines with an arbitrarily large, but still odd, number of nodes, in order to rule out the length of it as a factor that determines the optimal configuration.

We now proceed to the proof of this result. In order to deduce the best sequences for an arbitrarily large lines of nodes with an odd amount of nodes, it is mandatory to exclude some categories of lines as there exists an infinite number of possible combinations. First, we can demonstrate that an optimal sequence of agents must start and end with the same node type. Let us consider two sequences with the same interior nodes, but one starting and ending with the same node type, and the other with a different start and end. The comparison between the two would be equal to comparing just the beginning or the ending sequence as the rest of the sequence would be the same. Indeed, consider the sequences  $SP \dots PS_{rumour}$  and  $SP \dots PP_{rumour}$ . The latter would be better than the former if and only if  $PP_{rumour} > PS_{rumour}$ , or when  $\lambda > \lambda^*$ . Then, because  $\lambda > \lambda^*$ ,  $PP_{rumour} \dots > SP_{rumour} \dots$  as starting sequence, thus making  $SP \dots PP_{rumour}$  dominated by  $PP \dots PP_{rumour}$ , which has the same type of node at the start and end. Considering that all the other cases (*poisson* nodes at the beginning or end, *sticky* nodes at the end and *poisson* at the beginning,...) can be proven to not be optimal in the same way, we can conclude that it is impossible for an optimal sequence to start and end with different types of agents.

Next, we will mirror the proof of Akbarpour & Jackson [1] done for the normal contagion process, as their method works for this situation too. Thus, we can first

exclude all sequences containing both  $SS_{rumour}$  and  $PP_{rumour}$ . Consider a sequence starting and ending with *poisson* nodes that has some instances of  $SS_{rumour}$ . Then, in it, there must exist both  $PSS_{rumour}$  and  $SSP_{rumour}$  somewhere. If the sequence were to be optimal, it would imply that  $PSS_{rumour} > PPS_{rumour}$  and that  $SSP_{rumour} > SPP_{rumour}$ , and therefore would not be possible for  $PP_{rumour}$  to exist in the sequence. Indeed, if instead the sequence contained instances of  $PP_{rumour}$ , then, because it contains also single *sticky* nodes,  $PPS_{rumour} > PSS_{rumour}$  or  $SPP_{rumour} > SSP_{rumour}$ , thus making the presence of  $SS_{rumour}$  impossible. The same argument can be made for sequences starting and ending with *sticky* nodes. Indeed, if the sequence contained instances of  $PP_{rumour}$  and it were to be optimal, then there must exist somewhere the sequences  $SPP_{rumour}$  and  $PPS_{rumour}$ . This implies that  $SPP_{rumour} > SSP_{rumour}$  and  $PPS_{rumour} > PSS_{rumour}$ , thus making impossible for  $SS_{rumour}$  to exist in the sequence. Again, if there were a  $SS_{rumour}$ , then since the sequence has also *poisson* nodes, somewhere there is at least one  $PSS_{rumour}$  or  $SSP_{rumour}$ , making impossible for  $PP_{rumour}$  to exist.

Finally, we can prove that if there are instance of  $SS_{rumour}$ , it is impossible to have isolated *poisson* nodes on the interior. If in an optimal sequence there are two instances of  $SS_{rumour}$ , then it means that  $SSS_{rumour} > SPS_{rumour}$  (as  $SSS_{rumour}$  is  $SS_{rumour} \cdot SS_{rumour}$ ), so isolated *poisson* nodes cannot exist in the sequence. If instead those two instances had an isolated *poisson* node between them, or more than one, separated by isolated *sticky* nodes, the optimal sequence would have somewhere an  $SSPSS_{rumour}$ ,  $SSPSPSS_{rumour}$ , etc.. In the first case, to justify the presence of  $SPS_{rumour}$  in the sequence,  $SPS_{rumour} > SSS_{rumour}$ , meaning that  $SP_{rumour} \cdot PS_{rumour} > SS_{rumour}^2$ . However, if that was the case,  $SPSPS_{rumour}$  would always be better than  $SSPSS_{rumour}$  as  $SP_{rumour}^2 \cdot PS_{rumour}^2 > SS_{rumour}^2 \cdot SP_{rumour} \cdot PS_{rumour}$ , thus making impossible for one isolated *poisson* node to be between two instances of  $SS_{rumour}$ . All the other cases are extensions of the first one and can be proven with the same logic.

Thus, the only sequences left are:

- $SSS \dots SSS_{rumour}$ ;
- $SPP \dots PPS_{rumour}$ ;
- $PSS \dots SSP_{rumour}$ ;
- $SPS \dots SPS_{rumour}$ ;

- $PSP \dots PSP_{rumour}$ ;
- $PPP \dots PPP_{rumour}$ .

However,  $SPP \dots PPS_{rumour}$  is dominated by  $PPP \dots PPP_{rumour}$ , for  $\lambda > \lambda^*$ , as in that range  $PP_{rumour} \dots > SP_{rumour} \dots$ , and is dominated by  $SPS \dots SPS_{rumour}$  for  $\lambda < \lambda^*$ , as  $PSP_{rumour} > PPP_{rumour}$  for  $\lambda < \lambda^*$  as  $\lambda^* < \lambda^{***}$ , as proven in the previous statement. Furthermore,  $SPS \dots SPS_{rumour}$  is always dominated by  $PSP \dots PSP_{rumour}$  as both  $SPS_{rumour} \dots$  and  $SPS_{rumour}$  are always dominated respectively by  $PSP_{rumour} \dots$  and  $PSP_{rumour}$ . Therefore, the remaining sequences are:

- $SSS \dots SSS_{rumour}$ ;
- $PSS \dots SSP_{rumour}$ ;
- $PSP \dots PSP_{rumour}$ ;
- $PPP \dots PPP_{rumour}$ .

From the previous statement we already know that  $PPP \dots PPP_{rumour}$  will be better than  $PSP \dots PSP_{rumour}$  when  $\lambda > \lambda^{***}$ . Next, note that the comparison between  $PSS \dots SSP_{rumour}$  and  $PSP \dots PSP_{rumour}$  is the comparison between  $PS_{rumour}$  and  $SS_{rumour}$ , both following an infected *sticky* node. Thus, alternating nodes is better only if:

$$\lambda(1 - (1 - \lambda)^T)^4 > \lambda^2$$

$$\lambda < u(\lambda) = (1 - (1 - \lambda)^T)^4$$

Or when  $\lambda > \lambda^{***}$ . Finally, the comparison between  $PSS \dots SSP_{rumour}$  and  $SSS \dots SSS_{rumour}$ , is just the comparison  $PSP_{rumour}$  and  $SSS_{rumour}$ , as they are different only at the start and the end. For this we already know that  $PSP_{rumour} > SSS_{rumour}$  if and only if  $\lambda > \lambda^{**}$ . Moreover, note that when  $u(\lambda) = \lambda$ , then  $u(\lambda^2) < \lambda^2$ , since  $u(\lambda) < \lambda$  for  $\lambda < \lambda^{***}$ . This implies that  $\lambda > k(\lambda)$ , which means that  $\lambda < \lambda^{***}$  and so  $\lambda^{***} < \lambda^{**}$ . Also,  $\lambda^{***} > \lambda^{**}$  as  $(1 - (1 - \lambda)^T)^4 < (1 - (1 - \lambda)^T)^2$ .

Therefore, the order of the different sequences is going to be:

- $SSS \dots SSS_{rumour}$  for  $\lambda < \lambda^{**}$ ;
- $PSS \dots SSP_{rumour}$  for  $\lambda^{**} < \lambda < \lambda^{***}$ ;
- $PSP \dots PSP_{rumour}$  for  $\lambda^{***} < \lambda < \lambda^{**}$ ;
- $PPP \dots PPP_{rumour}$  for  $\lambda > \lambda^{**}$ .

To confirm the theory, simulations were carried out testing all the possible combinations for lines of 5 nodes, the results of which are shown below:

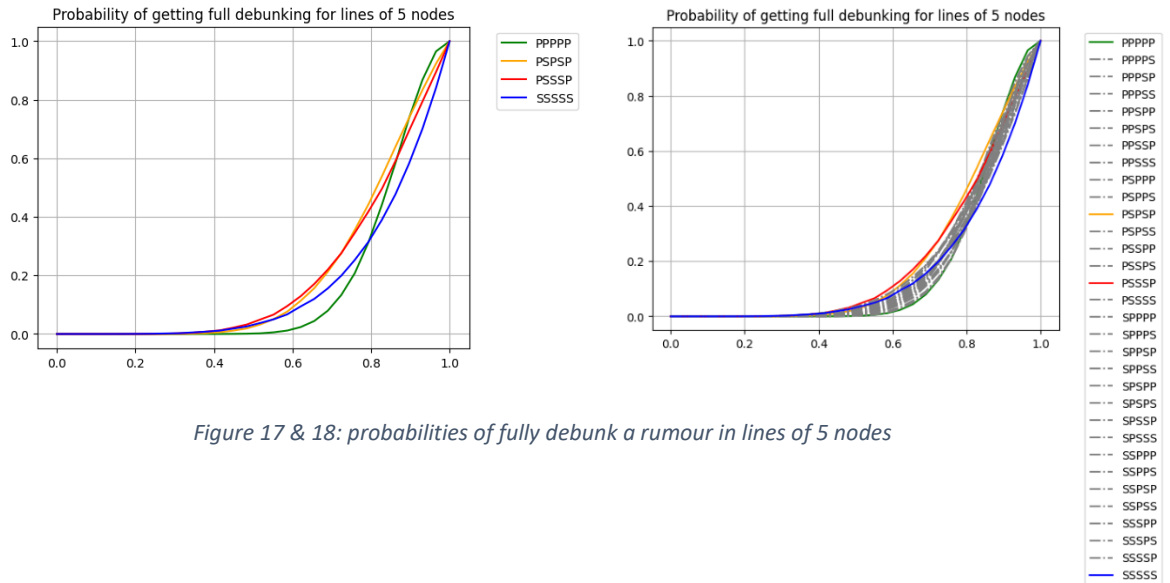


Figure 17 & 18: probabilities of fully debunk a rumour in lines of 5 nodes

As already predicted, the order and thresholds are the same as those theorised.

### 6.3. Diffusion of rumours in star-shaped networks

**Statement 3B:** In a star-shaped network with  $n$  leaves, where both *spreader* and *debunker* are both leaves, and with nodes that can transmit the disease for  $T = 2$  periods, the configuration that maximises the probability of fully debunking a rumour for any  $\lambda \in (0,1)$  is one such that all leaves are *poisson* and the centre is *sticky*.  $\square$

In line-shaped networks, for the *debunker* to learn about the rumour, all the other nodes had to hear it first. Instead, because in our model *debunkers* spread the truth also to nodes who have not heard the rumour, in star-shaped networks, as *spreader* and *debunker* are leaves, only the centre needs to hear the rumour for it to be spread to the *debunker*. For this reason, if we care to completely debunk a rumour, it does not matter if the other nodes hear the rumour in the first place. Therefore, the probability of fully debunking a rumour, is the probability that the *debunker* hears the rumour, times the probability that the truth spreads to all other nodes, accounting for correlated *sticky* nodes that we consider active during the rumour spread. The former is the probability of full infection, under normal contagion processes, in a line of three nodes -*spreader*, centre and *debunker*-. The latter is instead is the probability of full infection in a star-

shaped network, as the counter-information behaves as a normal contagion, accounting however for the *sticky* nodes counted as active for the rumour spreading process.

Moreover, in contrast to the spread of information, we cannot exclude that the optimal star-shaped configuration has *sticky* leaves, as we cannot exclude that the infection can start with a *sticky* node (proven in [statement 1B](#)). The notation used to indicate the configuration is of the form  $C_{L,type\ of\ diffusion}$ , where in the place of  $C$  there is the centre's type, and in the place of  $L$ , the leaves' type. Therefore, the possible configurations are:

- $S_{S,rumour}$ : central node is *sticky*, leaves are *sticky*;
- $P_{S,rumour}$ : central node is *poisson*, leaves are *sticky*;
- $S_{P,rumour}$ : central node is *sticky*, leaves are *poisson*;
- $P_{P,rumour}$ : central node is *poisson*, leaves are *poisson*.

Additionally, we use the amount of ' to count the number of correlated *sticky* nodes in the debunking process, as the *sticky* nodes involved in the contagion process are correlated to the ones in the debunking process. Therefore, the probability of fully debunking a rumour when the central node and leaves are *sticky*, is given by the probability of the rumour to spread in a line of three *sticky* nodes, times the probability of the truth to spread in the whole network accounting for the three *sticky* that are active:

$$S_{S,rumour}: SSS_{contagion} \cdot S'''_{S,contagion} = \lambda^3 \cdot \lambda^{n-2} = \lambda^{n+1},$$

In a similar way, the probability of fully debunking a rumour when the central node is *poisson* and the leaves are *sticky*, is:

$$\begin{aligned} P_{S,rumour}: SPS_{contagion} \cdot P''_{S,contagion} &= \lambda^2(1 - (1 - \lambda)^2) \cdot \lambda^{n-2}(1 - (1 - \lambda)^2) \\ &= \lambda^n(1 - (1 - \lambda)^2), \end{aligned}$$

the probability of fully debunking a rumour when the central node is *sticky* and the leaves are *poisson*, is:

$$\begin{aligned} S_{P,rumour}: PSP_{contagion} \cdot S'_P_{contagion} &= \lambda(1 - (1 - \lambda)^2)^2 \cdot (1 - (1 - \lambda)^2)^n \\ &= \lambda(1 - (1 - \lambda)^2)^{n+2}, \end{aligned}$$

and the probability of fully debunking a rumour when the central node and leaves are *poisson*, is:

$$\begin{aligned} P_{P,rumour}: PPP_{contagion} \cdot P_{P,contagion} \\ &= (1 - (1 - \lambda^2)^2) \cdot (1 - (1 - \lambda^2)^2)(2\lambda^n(1 - \lambda) + \lambda^{n+1}(2 - \lambda)^{n-1}) \\ &= (1 - (1 - \lambda^2)^2)^2(2\lambda^n(1 - \lambda) + \lambda^{n+1}(2 - \lambda)^{n-1}). \end{aligned}$$

To understand which structure is more convenient to our aim when the number of leaves grows to infinity, we first study the limit for  $n$  going to infinity of the ratio of the probabilities of networks with the same leaf type but different centres. The limit of  $\frac{P_{P,rumour}}{S_{P,rumour}}$  is a number between 0 and 1, thus making the configuration with a *sticky*

centre better. Indeed:

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{P_{P,rumour}}{S_{P,rumour}} &= \lim_{n \rightarrow \infty} \frac{(1 - (1 - \lambda^2)^2)^2 (2\lambda^n(1 - \lambda) + \lambda^{n+1}(2 - \lambda)^{n-1})}{\lambda(1 - (1 - \lambda)^2)^{n+1}} \\
&= \lim_{n \rightarrow \infty} (1 - (1 - \lambda^2)^2)^2 \cdot \frac{2\lambda^n(1 - \lambda) + \lambda^{n+1}(2 - \lambda)^{n-1}}{\lambda^{n+2}(2 - \lambda)^{n+1}} \\
&= \lim_{n \rightarrow \infty} (1 - (1 - \lambda^2)^2)^2 \cdot \left( \frac{2\lambda^n(1 - \lambda)}{\lambda^{n+2}(2 - \lambda)^{n+1}} + \frac{\lambda^{n+1}(2 - \lambda)^{n-1}}{\lambda^{n+2}(2 - \lambda)^{n+1}} \right) \\
&= \lim_{n \rightarrow \infty} (1 - (1 - \lambda^2)^2)^2 \cdot \left( \frac{2(1 - \lambda)}{\lambda^2(2 - \lambda)^{n+1}} + \frac{1}{\lambda(2 - \lambda)^2} \right) \\
&= \frac{(1 - (1 - \lambda^2)^2)^2}{\lambda(2 - \lambda)^2} = \lambda^3 \left( \frac{2 - \lambda^2}{2 - \lambda} \right)^2.
\end{aligned}$$

which for  $\lambda \in (0,1)$  is a number between 0 and 1. To prove this, notice that:

$$\lambda^3 \left( \frac{2 - \lambda^2}{2 - \lambda} \right)^2 < \lambda^2 \left( \frac{2 - \lambda^2}{2 - \lambda} \right)^2$$

for  $\lambda \in (0,1)$ , and thus proving that:

$$\lambda^2 \left( \frac{2 - \lambda^2}{2 - \lambda} \right)^2 < 1,$$

proves also that the result of the limit is less than 1. If we take the square root of both sides of the inequality we get:

$$-1 < \lambda \left( \frac{2 - \lambda^2}{2 - \lambda} \right) < 1 \Leftrightarrow \lambda - 2 < \lambda(2 - \lambda^2) < 2 - \lambda,$$

or

$$\lambda^3 - 3\lambda + 2 > 0$$

$$\lambda^3 - \lambda - 2 < 0.$$

Because the function  $v(\lambda) = \lambda^3 - 3\lambda + 2$  is decreasing between 0 and 1,  $v(0) = 2$  and  $v(1) = 0$ , then  $v(\lambda) > 0$  in that interval. Moreover, because the function  $w(\lambda) = \lambda^3 - \lambda - 2$  is decreasing between 0 and  $\frac{\sqrt{3}}{3}$ , is increasing between  $\frac{\sqrt{3}}{3}$  and 1, and  $w(0) = w(1) = -2$ , then  $w(\lambda) < 0$  in the interval, proving that the result of the limit is a number between 0 and 1.

This result is to be expected, as the correlation brings an advantage to the already optimal heterogenous configuration.

Similarly, the limit of  $\frac{S_{S,rumour}}{P_{S,rumour}}$  is a number between 0 and 1:

$$\lim_{n \rightarrow \infty} \frac{S_{S,rumour}}{P_{S,rumour}} = \lim_{n \rightarrow \infty} \frac{\lambda^{n+1}}{\lambda^n(1 - (1 - \lambda)^2)} = \frac{\lambda}{1 - (1 - \lambda)^2} = \frac{1}{2 - \lambda}$$

As  $0 < \frac{1}{2 - \lambda} < 1$  for  $\lambda \in (0,1)$ , making the configuration with a *poisson* centre better.

Lastly, we must compare the two heterogeneous star-shaped networks:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{S_{P,rumour}}{P_{S,rumour}} &= \lim_{n \rightarrow \infty} \frac{\lambda(1 - (1 - \lambda)^2)^{n+1}}{\lambda^n(1 - (1 - \lambda)^2)} \\ &= \lim_{n \rightarrow \infty} \lambda \left( \frac{1 - (1 - \lambda)^2}{\lambda} \right)^n = \lambda \cdot \lim_{n \rightarrow \infty} (2 - \lambda)^n = \infty \end{aligned}$$

As  $2 - \lambda \in (1,2)$  for  $\lambda \in (0,1)$ .

In conclusion, the network with a *sticky* centre and *poisson* leaves is the best star-shaped network to debunk rumours.

We can confirm this statement by simulating the configurations and looking at the different probabilities to fully debunk a rumour. The simulated star-shaped networks have 100 leaves and are evaluated 50,000 times:

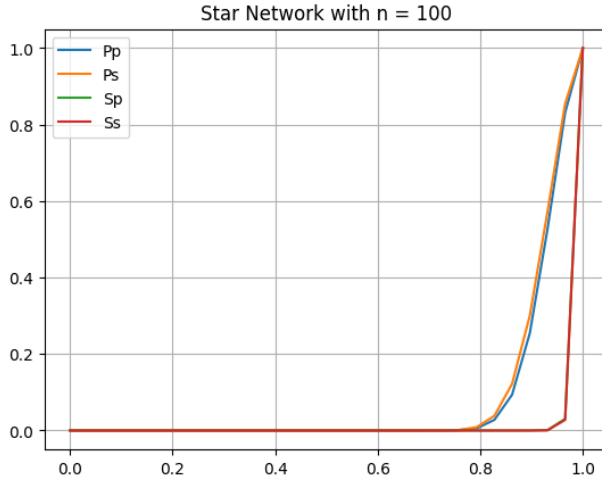


Figure 19: probabilities in star-shaped networks

As calculated before, the network with a *sticky* centre and *poisson* nodes dominates all the other configurations. Moreover, notice that it is generally unlikely to achieve the total debunking of a rumour in such networks, as the centre acts as a bottleneck for the process. Indeed, a *sticky* centre offers an advantage over a *poisson* centre, as if it is active during the contagion process, it will be active also during the debunking one. On the other hand, *poisson* leaves offer an advantage over *sticky* leaves, as in contrast to

line-shaped networks, star-shaped networks do not benefit from correlation, as leaves are important only in the debunking process. Therefore, *poisson* nodes that are inactive during the first turn in which the centre actively debunks, may turn active during the second turn, increasing the probability of fully debunking a rumour.

#### 6.4. Simulations of line-shaped networks with reversing nodes

**Statement 4B:** Consider a line-shaped network with five nodes composed of a mix of *sticky*, *poisson* or *reversing* nodes. Then, when  $T = 2$ , the configuration that maximises the probability of fully debunking a rumour is  $SSSSS_{rumour}$  for low levels of lambda,  $RSRSR_{rumour}$  for middle levels of lambda, and  $RRRRR_{rumour}$  for high levels of lambda. □

Since the results from the other statements were still partially valid even with the added complexity of the rumour, the next step is to test the hypothesis of heterogeneity in *statement 4*. Here, the lines are composed of 5 nodes and include also *reversing* nodes. Because *reversing* nodes add complexity to the model, we present only simulation without any formal proof. Before running the simulations it is possible to make some predictions about the results. Based on the results obtained for lines of 3 nodes, we can expect the line composed of only *sticky* nodes to perform better, for low levels of lambda, as they gain an advantage from correlation. Thus, at least in the beginning, we expect that to be the dominant line. However, since *reversing* nodes have a negative correlation to their previous activity status, we can also expect to pick up quite fast. For this reason, in the middle range of lambda, we can expect a combination of *reversing* and *sticky* nodes, while for higher lambda, a line composed of all *reversing* nodes. The results of the 50,000 simulations are presented in the graphs below:

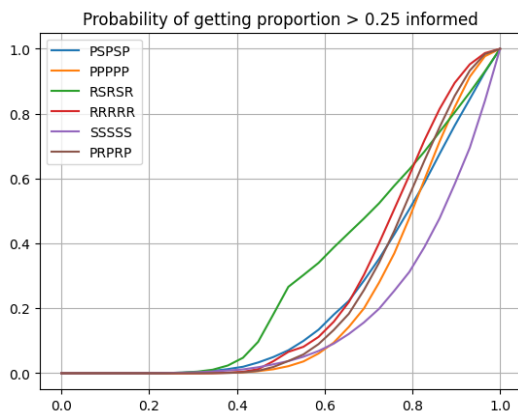


Figure 20: 25% know that the rumour is false

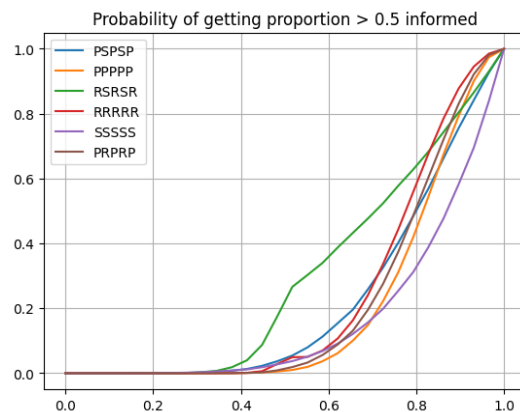


Figure 21: 50% know that the rumour is false



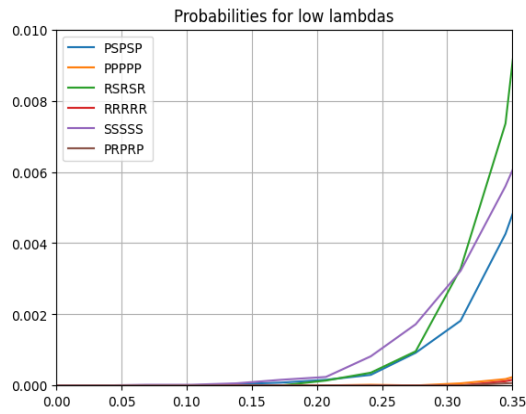


Figure 22: Restriction for low lambdas (> 0.5 prop.)

As expected, initially the line performing the best is the one composed of all *sticky* nodes, then a line of alternating *reversing* and *sticky* nodes takes the lead and finally the line with all *reversing* catches up and surpasses the others. Moreover, like for the normal diffusion, for middle level of lambdas the best combination is a mix of *RRRs* and *SRs*:

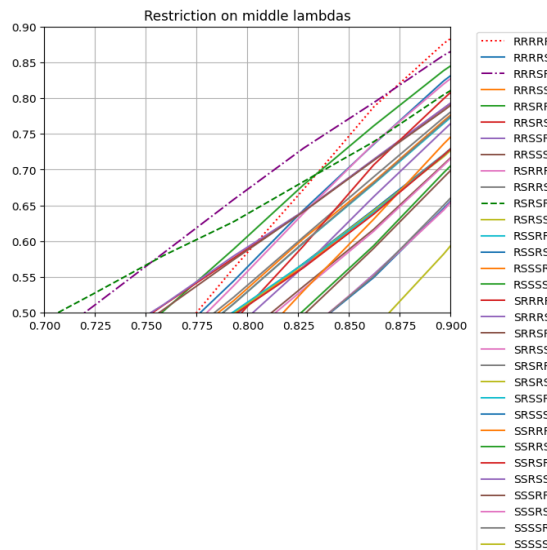


Figure 23: best configuration on middle levels of lambdas

## 7. Conclusions

The key difference between the spread of rumours and the spread of classic phenomena, like a piece of information is the debunking process. This process, absent in classic diffusion models, introduces a second piece of information that replaces the original, travelling through the network in an effort to debunk false data. Studying how the introduction of a counter-information alters the results obtained with classical contagion models is relevant,

as two messages instead of one makes a difference in the network. This dynamic in the rumour contagion model highlights the complexity of modern information exchange.

Central to our findings is the role of heterogeneity within these two models. As our study and the study of Akbarpour & Jackson [1] illustrate, heterogeneity helps the spread of true information and the debunking of false ones. Indeed, despite the inherent risk of accelerating the spread of a rumour, heterogeneity facilitates its debunking as well. As the fast spread of true and reliable information, and the correction of false data are crucial to society, heterogeneity ensures both better diffusion and debunking.

In relation to the activity levels of nodes, our study revealed interesting insights about the role of *sticky* agents. As they maintain their activity status over time, in rumour diffusion, they benefit from correlation, making them more desirable in the sequences. This result can be observed in multiple ways. First, in *statement 1B*, *statement 2B* and *statement 4B*, we observed that in contrast to *statement 1*, *statement 2* and *statement 4*, the sequences composed by all *sticky* nodes are optimal under certain thresholds of lambda (low  $\lambda$ ), showing that *sticky* agents have gained an advantage over the other types of nodes. Nonetheless, as *sticky* nodes are present in heterogeneous sequences, they too gain an advantage, making them more effective. Indeed, we observed in *statement 1B*, that the threshold between  $PSP_{rumour}$  and  $PPP_{rumour}$  has shifted to the right, meaning that  $PSP_{rumour}$  has gained an advantage over  $PPP_{rumour}$ , thanks to the presence of a *sticky* node. Moreover, the benefits of having always a heterogeneous population far outweigh the disadvantages also in the situation in which a rumour is being spread. Heterogeneous sequences proved to be much more efficient, guaranteeing a higher probability to spread for most levels of lambda. This is evident by plotting the ratios of the probabilities to fully debunk a rumour in heterogeneous sequences over the probability to fully debunk in homogeneous ones:

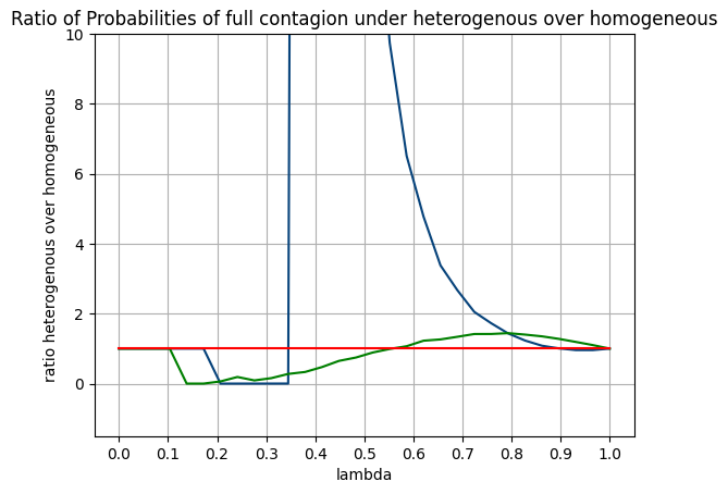


Figure 24: ratios of probabilities of full debunking in heterogeneous over homogeneous sequences

In the graph, the red line shows the indifference point, the blue line the ratio between alternating nodes and a sequence with all *poisson* nodes, and the green one is the same but the homogeneous sequence has all *sticky* nodes. Whenever the blue or green line is below the red, the homogeneous sequence is better than the heterogeneous one. Notice that, alternating nodes in a line-shaped network can be better (even more than 120 times) than having all *poisson* nodes. However, as *sticky* nodes have an advantage under our model, the advantages of heterogeneity are dampened compared to heterogeneous sequences of *sticky* nodes. Still, we can conclude that, if we need to choose a population without knowing  $\lambda$ , it is always better to choose a heterogeneous one.

To encapsulate, our study leans towards the necessity of heterogeneity in social networks as, despite its complexities, they offer unique advantages over homogeneity in the context of contagion models. As we venture deeper into the age of information, the understanding of those complex systems will significantly enhance our capacity to promote information flow, debunk misinformation and foster a healthier information ecosystem.

## 8. The Code

To support and verify the correctness of the statements made and to recreate the ones made in the paper written by Akbarpour & Jackson [1], it was developed a Python script to perform all the simulations shown. The code utilizes an object-oriented approach with classes to represent nodes, allowing for the incorporation of an update function to determine their activation status and for the storage of relevant information such as if the node is infected, how many turns still has left to infect or if it is a *debunker*. Every type of node has

its own class, that inherits the general characteristics of the node from the parent class Node. Each child class contains an update function to determine whether the node becomes active or not, based on specific criteria defined within the class. Some code snippets for the classes are shown below:

```

class Node:
    def __init__(self, edges, lambda_, c_time):
        self.edges = edges
        self.lambda_ = lambda_
        self.c_time = c_time
        self.infected_turns = 0
        self.debunking_turns = 0
        self.disease = False
        self.debunker = False # if the node initially know that the rumour circulating is false

    def __str__(self):
        return f'The edges for the node are: {self.edges} \nThe value of lambda is: {self.lambda_}'

```

Figure 25: code for general class "Node"

```

class Reversing(Node):
    def __init__(self, edges, lambda_, c_time):
        super().__init__(edges, lambda_, c_time)

        # defining the initial status of the node
        if random.random() < self.lambda_: # with probability lambda_ the node is active
            self.status = True
        else: # with probability 1-lambda_ the node is inactive
            self.status = False

        # defining p_i and q_i depending on lambda_
        if self.lambda_ <= 0.5:
            self.p = 1
            self.q = self.lambda_/(1 - self.lambda_)
        else:
            self.p = (1-self.lambda_)/self.lambda_
            self.q = 1

    def update_status(self):
        # to update the status of a Reversing node the status of the node at time t must be known
        # as if it is active it has a probability of p to become inactive and 1-p to remain active
        # while if it is inactive it has a probability of q to become active and 1-q to remain inactive
        if self.status: # if it is active
            if random.random() < self.p: # here random.random() takes values in the interval [0,1) so if
                # p is 1 it will always switch
                self.status = False
        else: # if it is inactive
            if random.random() < self.q:
                self.status = True

```

Figure 26: code for reversing nodes

Here, the rumour is represented by two Boolean values. If the node has not heard the rumour and is not a *debunker*, both the attributes `disease` and `debunker` assume the value "False". When a node hears for the first time the rumour, the attribute `disease` changes to "True" and, depending on the activity level, the node starts spreading it to its active neighbours. When the debunking process starts, the attribute `debunker` assumes the value

“True”, while the attribute `disease` the value “False”. This ensures that we can always track correctly the contagion process at every step.

The code also includes functions to set up the simulation environment, such as creating a network of nodes and spreading a disease or a rumour. In particular, the latter is a recursive function that stops when there are no more *spreaders* and *debunkers* that are able to either spread or debunk the rumour. The lists of possible *spreaders* and *debunkers* are updated at every iteration by checking if the nodes have been in the list for more than the turns they are supposed to. The spread and the debunking are done by selecting a node from one of the two lists and, if they are active, checking the activity of their neighbours. As all these characteristics are either stored as lists or Boolean values, it is quite easy to perform these operations and update the nodes with their new status.

```

def rumour(graph, spreaders, debunkers):
    """
    This recursive function simulate a rumour spreading in a network of nodes in which some nodes are able to debunk
    the information circulating

    Parameters:
    graph (list): the graph for running the simulation
    spreaders (list): list of all current spreaders
    debunkers (list): list of all current debunkers

    Returns:
    list: the graph
    """
    # stopping case
    if len(spreaders) == 0 and len(debunkers) == 0:
        return graph

    next_spreaders = []
    next_debunkers = []
    # when there are no current debunkers
    if len(debunkers) == 0:
        for spreader in spreaders:
            # active spreader
            if spreader.status:
                for receiver in spreader.edges: # check all neighbours
                    if receiver.status and not receiver.disease: # active receiver not yet infected
                        # not a debunker
                        if not receiver.debunkers:
                            receiver.disease = True
                            next_spreaders.append(receiver)
                        # if it is a debunker
                        else:
                            # becomes a new debunker only if it was not debunking before
                            if receiver.debunking_turns == 0:
                                next_debunkers.append(receiver)

                # check if the spreader should still spread next turn
                if spreader.infected_turns < spreader.c_time - 1:
                    spreader.infected_turns += 1
                    next_spreaders.append(spreader)

    # update status for every node
    for node in graph:
        node.update_status()

    # recursive call
    return rumour(graph, next_spreaders, next_debunkers)

# when there are no current spreaders
if len(spreaders) == 0:
    for debunker in debunkers:
        # active debunker
        if debunker.status:
            for receiver in debunker.edges: # check all neighbours
                if receiver.status and not receiver.debunker: # inform the node
                    receiver.debunker = True
                    receiver.disease = False # if it was not infected does not change
                    next_debunkers.append(receiver)

            # check if the debunker should still debunk next turn
            if debunker.debunking_turns < debunker.c_time - 1:
                debunker.debunking_turns += 1
                next_debunkers.append(debunker)

    # update status for every node
    for node in graph:
        node.update_status()

    # recursive call
    return rumour(graph, next_spreaders, next_debunkers)

# when the process is mixed
else:
    # carry the procedure until every node was evaluated
    randomized_action_nodes = spreaders + debunkers
    random.shuffle(randomized_action_nodes)

    for node in randomized_action_nodes:
        # if debunker
        if node.debunker:
            # active debunker
            if node.status:
                for receiver in node.edges: # check all neighbours
                    if receiver.status and not receiver.debunker: # inform the receiver
                        receiver.debunker = True
                        receiver.disease = False # if it was not infected does not change
                        next_debunkers.append(receiver)

            # check if debunker should still debunk next turn
            if node.debunking_turns < node.c_time - 1:
                node.debunking_turns += 1
                next_debunkers.append(node)

        # if spreader that was not convinced during this turn
        if node.disease:
            # active spreader
            if node.status:
                for receiver in node.edges: # check all neighbours
                    if receiver.status and not receiver.disease: # if the receiver is active and not yet infected
                        # if it is not a debunker
                        if not receiver.debunkers:
                            receiver.disease = True
                            next_spreaders.append(receiver)

                # if it is a debunker
                else:
                    # become a debunker only if was not a debunker before
                    if receiver.debunking_turns == 0:
                        next_debunkers.append(receiver)

            # check if spreader should still spread next turn
            if node.infected_turns < node.c_time - 1:
                node.infected_turns += 1
                next_spreaders.append(node)

    # update the status for every node
    for node in graph:
        node.update_status()

    # check if some spreader got convinced by a debunker (happens when the node contaminate someone and still has
    # time but then gets converted)
    convinced = []
    next_spreaders_noconv = []
    for i in range(len(next_spreaders)):
        if not next_spreaders[i].disease:
            convinced.append(i)

    if len(convinced) != 0:
        for i in range(len(convinced)):
            if i not in convinced:
                next_spreaders_noconv.append(next_spreaders[i])

    return rumour(graph, next_spreaders_noconv, next_debunkers)

# recursive call
return rumour(graph, next_spreaders, next_debunkers)

```

Figure 27: function handling the contagion and debunking processes

Also, the function update directly the network so that, after performing the simulation, is easy to do other operations in the graph such as visualising the effects of the process. Indeed, all the graphs realised were made outside the functions using the `Matplotlib` library.

The results of the simulations were obtained after running 50,000 individual graphs for 30 different values of  $\lambda$ s, to ensure that the results do not change after rerunning the code (except on simulations for *statement 3B* due to the extensive duration required for their completion). Moreover, as different and non-related simulations are performed throughout the work, the code is presented in a Jupiter Notebook, where each experiment can be run independently so as to facilitate their replicability.

In conclusion, by employing an object-oriented approach with classes to represent nodes and their update functions, the code is flexible and adaptable, allowing for easy modification to test different scenarios and assumptions. In this way, the work done can be employed for further research on the topic, as this serves as the main framework for rumour analysis.

## References

1. Akbarpour, M., & Jackson, M. O. (2018). Diffusion in networks and the virtue of burstiness. *Proceedings of the National Academy of Sciences*, *115*(30), E6996-E7004.
2. Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K. I., Kondor, I., & Barabási, A. L. (2006). Modeling bursts and heavy tails in human dynamics. *Physical Review E*, *73*(3), 036127.
3. Vazquez, A., Racz, B., Lukacs, A., & Barabasi, A. L. (2007). Impact of non-Poissonian activity patterns on spreading processes. *Physical review letters*, *98*(15), 158702.
4. Pan, R. K., & Saramäki, J. (2011). Path lengths, correlations, and centrality in temporal networks. *Physical Review E*, *84*(1), 016105.
5. Moreno, Y., Pastor-Satorras, R., & Vespignani, A. (2002). Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, *26*, 521-529.
6. Zhao, L., Wang, J., Chen, Y., Wang, Q., Cheng, J., & Cui, H. (2012). SIHR rumor spreading model in social networks. *Physica A: Statistical Mechanics and its Applications*, *391*(7), 2444-2453.
7. Isham, V., Harden, S., & Nekovee, M. (2010). Stochastic epidemics and rumours on finite random networks. *Physica A: Statistical Mechanics and its Applications*, *389*(3), 561-576.
8. Merlino, L., & Tabasso, N. (2019). Debunking Rumors in Networks. University Ca'Foscari of Venice, Dept. of Economics Research Paper Series No, 29.
9. Tabasso, Nicole. 2019. "Diffusion of Multiple Information: On Information Resilience and the Power of Segregation." *Games and Economic Behavior*, *118*: 219–40.
10. Campbell, Arthur, C. Matthew Leister, and Yves Zenou. 2019. "Social Media and Polarization." Mimeo.
11. Daley, D. J., & Kendall, D. G. (1965). Stochastic rumours. *IMA Journal of Applied Mathematics*, *1*(1), 42-55.
12. Maki, D. P., & Thompson, M. (1973). *Mathematical models and applications: with emphasis on the social life, and management sciences* (No. 511.8 M3).
13. Nekovee, M., Moreno, Y., Bianconi, G., & Marsili, M. (2007). Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, *374*(1), 457-470.
14. Zanette, D. H. (2001). Critical behavior of propagation on small-world networks. *Physical Review E*, *64*(5), 050901.
15. Zanette, D. H. (2002). Dynamics of rumor propagation on small-world networks. *Physical review E*, *65*(4), 041908.



16. Moreno, Y., Nekovee, M., & Vespignani, A. (2004). Efficiency and reliability of epidemic data dissemination in complex networks. *Physical Review E*, 69(5), 055101.
17. Moreno, Y., Nekovee, M., & Pacheco, A. F. (2004). Dynamics of rumor spreading in complex networks. *Physical review E*, 69(6), 066130.