Stockholms universitet

# Stopping the Spread: A Network-Based Stochastic Model of Fake News Correction

Mathilde Coffy

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2025:8
Matematisk statistik
Juni 2025

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2025:8**
http://www.math.su.se

# Stopping the Spread: A Network-Based Stochastic Model of Fake News Correction

Mathilde Coffy*

June 2025

**Abstract**

In recent years, misinformation has become a prevalent societal challenge. Social media platforms have grown ubiquitous and have enabled the rapid and widespread dissemination of fake news. This study aims to provide practical guidance for developing more effective debunking strategies. This is a meaningful but challenging endeavor, given the fine line between beliefs catching on and being corrected. In this thesis, we develop a stochastic model to simulate the spread of misinformation and its correction in social networks. We construct the network using a configuration model and apply a modified SIR model to simulate belief spread. In our simulations, we study how misinformation propagates through the network and how the introduction of skeptical agents can help in mitigating this spread. Final outcomes depend on the timing and effectiveness of debunking interventions. We find that early introduction of skeptical individuals is crucial for limiting misinformation spread. Timely intervention curbs belief adoption effectively and delayed debunking is only successful if it is strong. This study also shows the importance of the network structure, as highly connected hubs amplify belief spread. Finally, the balance between belief persuasiveness, debunking strength, and the capacity of skeptics to influence others determines the final outcome.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: mathilde.coffy@gmail.com. Supervisor: Daniel Ahlberg, Mia Deijfen.

# Acknowledgments

This is a bachelor thesis of 15 ECTS in Mathematical Statistics at the Department of Mathematics at Stockholm University.

I would like to express my sincere gratitude to my advisors Daniel Ahlberg and Maria Deijfen for their guidance, support and inspiration. Thank you both for your insightful and timely feedback, as well as for all the thought-provoking discussions that have undoubtedly enriched this project.

I would also like to acknowledge the assistance of ChatGPT for helping with minor Python issues like correcting errors and for valuable support with LaTeX.

# Contents

# 1 Introduction

While misinformation has long existed, the rapid growth of social media over the past 15 years has amplified its spread. The algorithmic nature of these platforms and their powerful sharing capabilities allow for fake news to reach wide audiences in unprecedented ways.
This dissemination of news has proven to be both wider and deeper today than in the past and this shift has completely altered the way information is consumed. The increased exposure, coupled with the emotional appeal often present in misinformation, has rendered individuals more susceptible to believing and internalizing these fabricated stories. As a result, we are facing a higher societal vulnerability to these false narratives.

## 1.1 Background

"Fake news" refers to misleading or false information that is presented as legitimate. Its objective is often to deceive, manipulate or trick individuals, with the aim of shaping public opinion on a larger scale. It comes in many forms; ranging from distorted facts, made-up stories to click-bait articles. These stories tend to grab the reader's attention as they are dramatic, shocking or simply emotionally charged.

In their very design, these phenomena appear to be tightly connected to human psychology. Cognitive biases, such as the tendency to focus more on negative or emotionally charged content or to believe information that confirms existing views, make individuals more likely to believe and share fake news. It often resonates in a somewhat irresistible way with their worldview and values.

A recent article by Vosoughi et al. [[8]] shows that fake news tends to spread faster and more broadly than accurate information across social media networks. This trend holds across numerous news categories, ranging from politics and entertainment to business. The study also shows that despite popular belief, the spread of fake news is not so much driven by robots or influential individuals as it is by typical human behavior; it is just the novelty of this fake news, coupled with its emotional intensity or scandal value, that entices people to profusely share it. That explains the high velocity of these pieces of news spreading in a network.
These observations motivate a closer examination of how misinformation travels in a network.

## 1.2 Purpose of the study

In this study, we are seeking to develop a model to gain a better understanding of how misinformation spreads in social networks. By means of a configuration model to approximate the structure of real social networks, and an adapted version of the SIR model to depict how information is spreading, we introduce several states to show how individuals respond to a piece of misinformation. Specifically, we seek to answer the following questions:

- What is the effect of debunking, and how does its timing affect its effectiveness?

- How do changes in transmission probabilities affect the eventual spread and outcome?

As noted by Bateman and Jackson in their report for the Carnegie Endowment for International Peace [[1]], debunking interventions are an important tool to be used to limit the spread of misinformation. A thorough understanding of the spreading mechanisms is essential to design policies that will have a long-lasting effect. The authors emphasize the importance of early action, as well as the clear labeling of fake news as such. Another important aspect is that debunking should come from sources deemed trustworthy and recognized as such by the general public.

These findings lead to a more informed understanding of the link between the psychological aspects - what motivates us as social beings - and the viral nature of social media. In our model, we thus combine both the behavioral and network-based perspectives. Our goal is to develop a simulation tool for studying misinformation spread and evaluating interventions. Ultimately, we aim to understand how misinformation can most effectively be curbed in today's highly connected world.

# 2 Theory

In this section, we introduce the theoretical tools used in the construction of our model. We focus on random graph theory and epidemic modeling.

## 2.1 Configuration model

### 2.1.1 Overview of the model

To construct the social network, we use the configuration model, which is widely used for generating stochastic graphs with a specified degree sequence. This model displays important characteristics found in real social networks, such as a heavy-tailed, power-law degree distribution. In such networks, most individuals have a limited number of social connections and a small number have many, creating hubs or centers of influence.

In our setting, nodes represent individuals within their social network, and edges represent social connections through which information can spread.

An important assumption in our model is that all nodes have at least two connections, i.e., $\mathbb{P}(D \geq 2) = 1$, and that the degree is at least three with positive probability, i.e., $\mathbb{P}(D \geq 3) > 0$. Through this assumption, we make sure that the generated random graph will most likely be connected when $n$ is large and we avoid ending up with long chains or small, disconnected clusters. This is a well-established result in the study of random graphs with a given degree sequence, as shown in Molloy and Reed (1998) [6].

### 2.1.2 Graph construction

Following the theory of Deijfen and van der Hofstad (2016) [4], each node is assigned a certain number of stubs (or half-edges) based on a degree distribution that is predefined. After assigning each node $i$ a degree $D_i$, we proceed to giving it $D_i$ half-edges. These half-edges are then paired uniformly at random, in order to form edges. This ensures that the degree distribution is preserved, while also introducing randomness in the network structure. For the pairing to work, we need to make sure that the number of stubs is even. If the sum is odd, we adjust the degree of one node.

One known limitation of the configuration model, as discussed in Deijfen, Rosengren, and Trapman (2016) [3], is that it can potentially give rise to self-loops as well as multiple edges between the same pair of nodes. As these features are undesirable for our purpose, we need to refine the resulting graphs by removing these eventual self-loops and duplicate edges. As a consequence, some nodes may end up with fewer than two connections.

The network is constructed so that all connections between nodes are bidirectional. If, for example, node A is connected to node B, both can then influence each other. This would be a reflection of real-life interactions, where information and belief can flow in both directions.

#### 2.1.3 Degree distribution: Pareto sampling

In order to match real social networks, we make the assumption that each individual's number of connections follows a power-law distribution. We therefore sample degrees from the Pareto distribution, denoted Pa(k, $\alpha$). As described by Gut (2009, p. 285) [[5]], this distribution has the following density function :

$$f(x) = \begin{cases} \frac{\alpha k^\alpha}{x^{(\alpha+1)}}, & \text{for } x \geq k \\ 0 & \text{for } x < k \end{cases} \tag{1}$$

where $\alpha > 0$ is the shape parameter and $k > 0$ is the minimum degree.

The shape parameter $\alpha$ determines the properties of the distribution:

- If $\alpha > 1$, the mean of the distribution is finite:

$$E[X] = \frac{\alpha k}{\alpha - 1}$$

- If $\alpha > 2$, the variance of the distribution is finite:

$$\text{Var}(X) = \frac{\alpha k^2}{(\alpha - 1)^2(\alpha - 2)}$$

- If $\alpha \leq 1$, the mean is infinite.
- If $1 < \alpha \leq 2$, the mean is finite, but the variance is infinite.

As the Pareto distribution is continuous and we require a discrete sequence of degrees, we discretize the values by rounding to the closest integer.

This generates a degree sequence which has a power-law tail and this is consistent with many empirical social networks. In such networks, high-degree hubs are rare but important. Specific choices for $\alpha$ and their effects on network structure are discussed later in Section 4.

### 2.2 The SIR model

#### 2.2.1 Overview

The following section follows Britton's description of the SIR model (Britton, 2010) [2]. The SIR (Susceptible-Infected-Recovered) model is a commonly used framework in the field of epidemiology for modeling the spread of infectious diseases in a population.

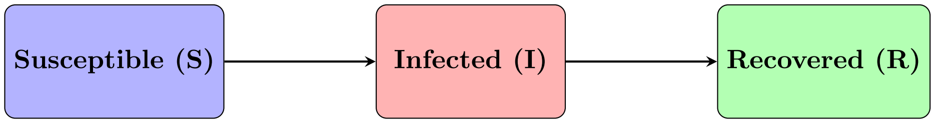


Figure 1: The SIR model showing transitions between Susceptible, Infected, and Recovered states.

As displayed in Figure 1, individuals begin in the Susceptible (S) state, and then may become Infected (I) through contact with infected individuals. They eventually transition to Recovered (R) state. Individuals progress through these different states according to specific transition rules.

This study specifically focuses on a discrete-time version of the SIR model, inspired by the Reed–Frost model. In that model, time progresses in discrete generations. Initially, one individual (the index case) is infected, and all others are susceptible. At each time-step, each infected individual independently tries to infect each susceptible neighbor with a fixed probability $p$. The infected individuals recover after that attempt and no longer play a part in the epidemic. They move to the recovered state. The process continues through successive generations until no new infections occur, and the epidemic dies out.

#### 2.2.2 Properties of the SIR model

Some important properties of the SIR model help clarify how the process unfolds:

- Each individual can only get infected once. One cannot become susceptible again after recovery.
- Recovered individuals no longer contribute to the spread. This feature creates a steady decline in the number of susceptible individuals over time.
- The process eventually stops because everyone has recovered or because no further infections are possible due to the network structure.

The SIR process can be described as Markovian as it possesses the memoryless property: future states depend solely on the current state, not on the full history. As detailed in Ross (2019, p.193) [[7]], we let $X_t$ denote the state of the network at time t, where $X_t$ is a vector that contains the status of each node in the network. The Markov property then gives us that:

$$\mathbb{P}\big(X_{t+1} \mid X_t, X_{t-1}, \ldots, X_0\big) \;=\; \mathbb{P}\big(X_{t+1} \mid X_t\big).$$

# 3 Model description

In this section, we define the model that we are using to simulate the spread of misinformation and corrective efforts in a social network. The model consists of two main components: the structure of the network and the rules for how belief and skepticism spread in this network. We base our model on the classical SIR model, but it is extended to incorporate mechanisms for debunking.

## 3.1 Network structure

We model the social network using a random graph generated by the configuration model. The degree sequence follows a power-law distribution.

The shape parameter $\alpha$ controls how heavy-tailed the distribution is. Low values of $\alpha$ give networks with more high-degree nodes (hubs), which leads to greater heterogeneity in the connections between nodes. Higher values result in more uniform degree distributions. When varying $\alpha$, we can study how differences in the structure of the network influence the spread of belief and correction.

All edges are undirected, which allows for information and influence to flow both ways between connected nodes.

## 3.2 Nodes states

Each node in the network can be in one of the four following states:

- **Susceptible (S)**: The individual has not yet formed a belief about the piece of news.
- **Believer (B)**: The individual currently believes the misinformation.
- **Non-Believer (N)**: The individual has been exposed to the true message and believes it. This state is absorbing.
- **Recovered (R)**: The individual initially believed the misinformation but has since abandoned this belief. This state is also absorbing.

## 3.3 Possible transitions between states

In the classical SIR model, individuals are either Susceptible (S), Infected (I), or Recovered (R). At each time-step, infected individuals attempt to infect their susceptible neighbors. If contacted, a susceptible individual becomes infected or remains unaffected; it is only successful infections that are counted in the model.

### 3.3.1 Overview of transitions

In our adapted model for the spread of information, we introduce additional transitions. When a Susceptible individual is exposed to a piece of news, it may:

- remain Susceptible,
- believe the fake news (transition to Believer),
- accept the true message (transition to Non-Believer).

This branching structure introduces competing narratives into the network. We can herewith model both the spread of misinformation and the emergence of skepticism. This renders the system more complex, but also more similar to how people would respond to competing versions of information in real social networks.

### 3.3.2 Visual representations of transitions

Transitions occur when one node attempts to influence a neighbor, which may lead to a change of state. Figure 2 introduces all allowed transitions in the belief spread model.

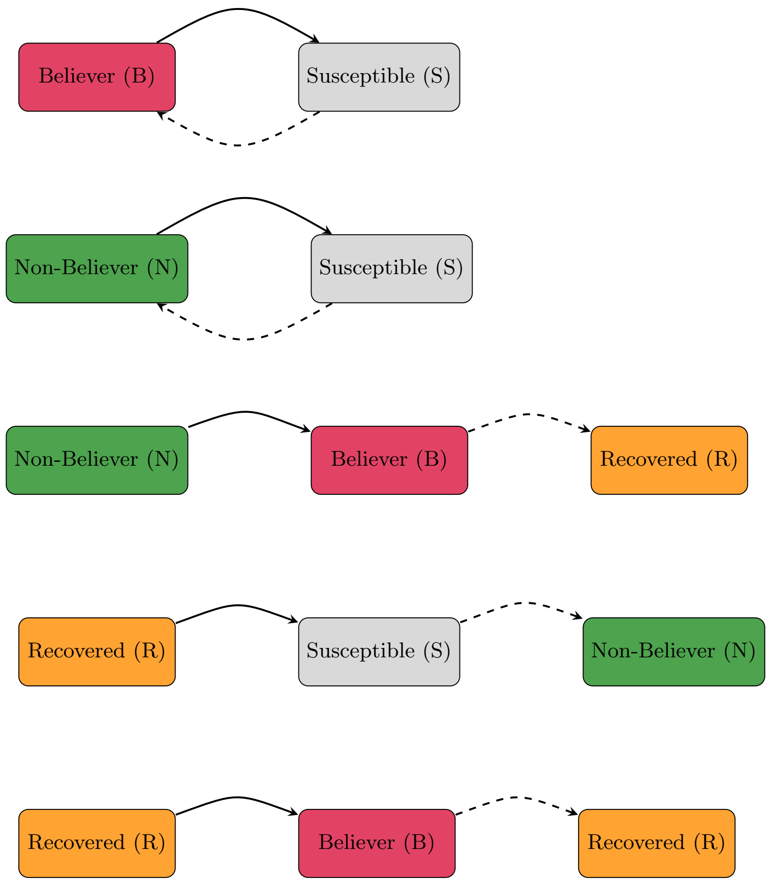


Figure 2: Transitions in the belief spread model. Solid arrows represent influence attempts; dashed arrows represent the resulting state change.

#### 3.3.3 Detailed transition descriptions

The transitions shall be interpreted as:

- Believer → Susceptible Neighbor → Believer
  A Believer can convert a Susceptible neighbor into a Believer.
- Non-Believer → Susceptible Neighbor → Non-Believer
  A Non-Believer can likewise convert a Susceptible neighbor into a Non-Believer.
- Non-Believer → Believer Neighbor → Recovered
  A Non-Believer individual can influence a Believer to recover.
- Recovered → Susceptible Neighbor → Non-Believer
  A Recovered individual can convince a Susceptible Neighbor to become skeptical.

- Recovered → Believer Neighbor → Recovered
  A recovered individual can also persuade a Believer to abandon its belief and recover.

### 3.3.4 Transition assumptions and constraints

Certain transitions are not allowed by design. For example, a Believer cannot convert a Non-Believer and a Non-Believer cannot turn a Believer into a Non-Believer, just lead the individual to recovery.

The role of Recovered individuals varies from the classical SIR model. In our model, a Recovered node (a former Believer) can directly influence a Susceptible individual to become a Non-Believer.
Additionally and unlike in the classical SIR model where recovery can happen spontaneously, Believers in our model can only abandon their belief through external influence from Recovered or Non-Believer nodes.
Finally, each node gets exactly one chance to attempt to spread its belief or skepticism after a state change. After this attempt, it becomes inactive. This would reflect the idea that individuals typically share a piece of information only once.

In our model, an individual who is in the Recovered state behaves identically to an individual in the Non-Believer state.

## 3.4 Transition probabilities

In order to model the patterns of influence in the network, we define the following parameters. Each corresponds to a type of social interaction:

- $\beta$: **Belief persuasiveness**
  The probability that a Believer convinces a Susceptible individual to adopt the belief and become a Believer. This parameter gives a measure of the scandal value of the information.

- $\gamma$: **Debunking strength**
  The probability that a Non-Believer or Recovered individual convinces a Believer to abandon the belief and transition to the Recovered state. This parameter measures how successful corrective messages are in changing beliefs.

- $\delta$: **Skeptic spread strength**
  The probability that a Non-Believer or Recovered individual persuades a Susceptible to become a Non-Believer. This represents the spread of skepticism.

Understanding the relationship between the model parameters $\beta$ and $\delta$ helps us build a clearer understanding of the processes at stake. These parameters reveal how easily misinformation spreads and how effectively it can be corrected. They reflect various real-world scenarios. We can for instance consider the two following relationships:

- High scandal value scenario ($\beta > \delta$): This illustrates situation where the piece of misinformation is highly sensational or emotionally charged. That renders it more likely to spread than the true version. Examples of this could include viral conspiracy theories or celebrity scandals. In these cases, the initial false claims would typically spread widely before any corrective information can be introduced.

- Correction scenario ($\delta > \beta$): In this case, the corrective information ends up being more convincing and hereby spreads faster than misinformation. This may occur when trustworthy institutions swiftly counter false claims. For example, when health organizations debunk misinformation about vaccines.

These two scenarios illustrate the extremes of misinformation spread. The adjustment of the parameters allows us to study the conditions under which misinformation thrives and propagates or can be mitigated.

## 3.5 Time dynamics

### 3.5.1 Spread of information

Our simulations are operated in discrete time steps (t = 0, 1, 2, ...). The spread of information in the network occurs via a discrete-time stochastic process, following the classic Reed-Frost model. At each time-step:

- Nodes that have transitioned into a new state at the current time-step become active. These active nodes try to influence each of their direct neighbors.

- Each attempt to influence someone succeeds independently and with a fixed probability, depending on the states of the individual trying to exert influence and its target.

- Once an active node has attempted to influence its neighbors, it becomes inactive. This reflects the idea that individuals would only typically share information once.

- All state transitions happen simultaneously at each time-step. This contributes to a consistent evolution of the network.

Whether a node transitions to another state is determined by an independent stochastic Bernoulli trial. For every interaction, a uniformly random number between 0 and 1 is drawn; if it lies below the corresponding transition probability, the node changes state in the next time-step. It remains unchanged otherwise. Over multiple time-steps, the number of successful transitions for the nodes follows a binomial distribution. This provides us with the overall outcome of the spread.

Given the stochastic nature of the model, small differences in early interactions can have a strong impact on the final outcome and this leads to wide variability between simulations. As documented in Britton (2010) [[2]], such models often display high variability in the early phases, where chance can ultimately lead to a wide spread or a fading-out process. To mitigate this effect, we run a large number of trials which we average over and always ensure that a steady state is reached before analyzing outcomes.

### 3.5.2 Introduction and spread of skepticism

Debunking plays a defining role in the containment of misinformation. Debunking and skepticism are modeled through the introduction of a Non-Believer node at a deterministic time point. This node can be seen as the embodiment of a fact-checking or debunking entity entering the network.

Once introduced, the Non-Believer node attempts to influence its neighbors by:

- either convincing a Susceptible neighbor to become a Non-Believer,
- or persuading a Believer to abandon its belief and transition to the Recovered state.

The Non-Believer and Recovered states are absorbing: once a node enters one of these states, it stays there permanently. It will attempt to influence its neighbors once, upon arrival in that absorbing state.

Compared to the classical SIR model, our setup introduces an extension through the addition of the Non-Believer state. This addition allows us to differentiate between:

- Non-Believers: individuals that never believed the misinformation,
- Recovered: Individuals who once believed but have since rejected the misinformation,
- Susceptible: Individuals who have not yet formed a belief.

This structure portrays a more accurate picture of how misinformation and debunking spread within social networks.

### 3.5.3 Step-by-step spread example

At the start of each simulation, all nodes start in the Susceptible state except for one randomly selected Believer who initiates the spread of a piece of misinformation. We choose to start with only one random Believer so as to better isolate the effect of the structure of the network, and the timing of countermeasures like debunking.

To best explain the model, we start with an example. Figure 3 below shows how influence spreads from a single source during the early simulation stages, before skepticism is introduced. The layout supplies us with an illustration of how the network structure shapes the initial spread.

The node colors represent different states:

- **Gray**: Susceptible (S)
- **Red**: Believer (B)
- **Green**: Non-Believer (N)
- **Orange**: Recovered (R)

The following steps illustrates how the belief spreads through the network, during the first few steps:

- **Step 0:** The simulation begins. All nodes are Susceptible except for a single random Believer node (#15), which initiates the spread.
- **Step 1:** The initial Believer tries to influence its neighbors. Some nodes get converted into Believers (#7, #14 and #19). Each node only gets one chance to influence others.
- **Step 2:** The newly converted Believers from step 1 now influence their own neighbors, which leads to an extension of the red region. We randomly also turn a single Susceptible into a Non-Believer (#0).
- **Step 3:** The spread slows down. No new nodes are converted into Believers. The single Non-Believer converts two Susceptibles into Non-Believers (#5 and #10) and three Believers into Recovered (#9, #14 and #16).

It is important to note that in this example we only uses 20 nodes, but the full simulations are conducted using a network of 10,000 nodes.

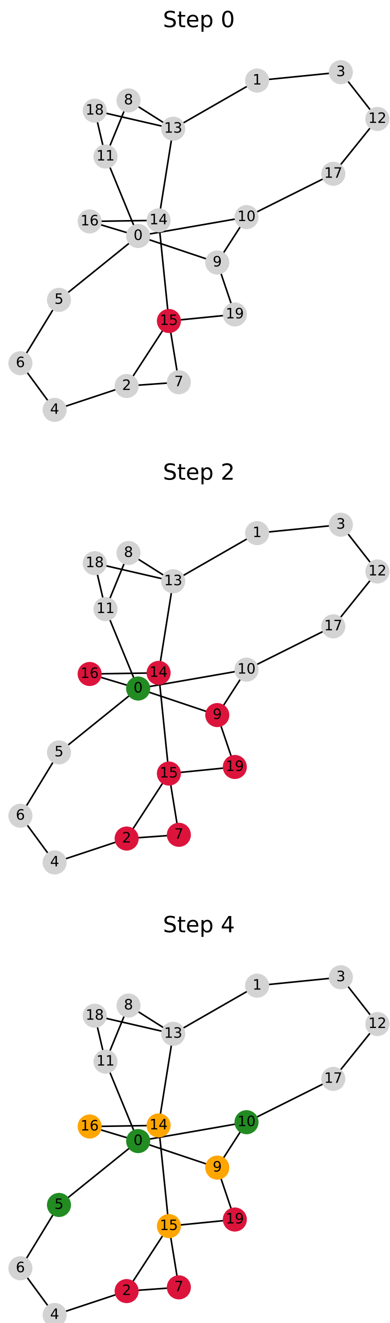


Figure 3: Visualizing the spread: network evolution over the first 6 steps of belief spread. After step five, the news is contained and will no longer spread. Node colors represent different states: Gray indicates Susceptible (S), Red indicates Believer (B), Green indicates Non-Believer (N), and Orange indicates Recovered (R)

### 3.6 Model's assumptions

Before presenting the simulation results, we summarize here our assumptions for the structure of the model:

- **Initial condition**: The spread starts with one randomly selected Believer node in the network.
- **Simultaneous updates**: Time is discrete. All nodes update their states simultaneously at each time-step, based on the state of the network at the previous time-step.
- **Single attempt**: Each node gets exactly one chance to influence its neighbors after entering a new state.
- **Bidirectional interactions**: All edges are undirected; influence can flow in both directions between connected nodes.
- **Absorbing states**: Once a node becomes a Non-Believer or Recovered, it will remain in that state permanently.
- **No spontaneous recovery**: A Believer can only recover through external influence, via a Non-Believer or a Recovered individual.
- **Debunking introduction**: A single Non-Believer is introduced deterministically at a specified time-step. Until then, no skepticism is present in the system.
- **Fixed transition probabilities**: Three probabilities rule the spreading: belief spread ($\beta$), skepticism spread ($\delta$), and debunking strength ($\gamma$). They are kept constant within each simulation run.
- **Static network**: The network is generated once (via the configuration model with a power-law degree distribution) and remains fixed throughout the simulation.

# 4 Results

In this part, we present the results of our simulations. We study how belief spreads and how intervention strategies can affect the final distribution of node states.

Our objective is to evaluate the effectiveness of debunking under different conditions. We focus on three questions:

- When should we intervene in order to maximize the effect of debunking?
- Can late interventions still mitigate belief spread?
- How do different spreading strengths change the system's long-run outcome?

To answer these questions, we perform simulations that vary the timing of non-believer introduction and the transition probabilities: the persuasiveness of belief, the effectiveness of skepticism, and the strength of debunking. Each simulation runs for 30 discrete time-steps and is repeated 1,000 times, to take into account stochastic variability. In all the simulations presenting the average evolution of node states presented in this study, we include only simulations where both Believer and Non-Believer populations independently reach at least 100 nodes. This ensures that both belief and skepticism have a chance to spread and that allows us to convincingly evaluate the impact of debunking.

## 4.1 Simulation setup and execution

The simulations for this study were carried out using Python. A few standard libraries were used:

- `networkx` to build and handle the networks,
- `numpy` for random number generation (including Pareto distribution) and calculations,
- `matplotlib` for creating plots as well as some built-in modules like random and time.

All parts of the simulations (how networks were built, how belief spread was simulated, and how node states changed over time) were coded manually without relying on simulation packages.

Each simulation involved constructing a network of 10,000 nodes and simulating the spread over 30 time-steps, and repeating this process 1000 times for each setting. This was repeated across 15 different introduction times for the non-believer intervention. Each such set of simulations generally required 2 hours to complete on an AMD $7800X3D$ processor with 8 cores. The final data used in the thesis is based on 10 such simulations. In total 48 simulations with varying parameters were performed, but all data has not been included in the final report as the method has evolved over time.

## 4.2 Timing of introduction

An important parameter in this model is the timing of introducing a Non-Believer individual who acts as a debunker. By varying this timing, we can analyze the impact on the overall spread of the belief and its potential suppression.

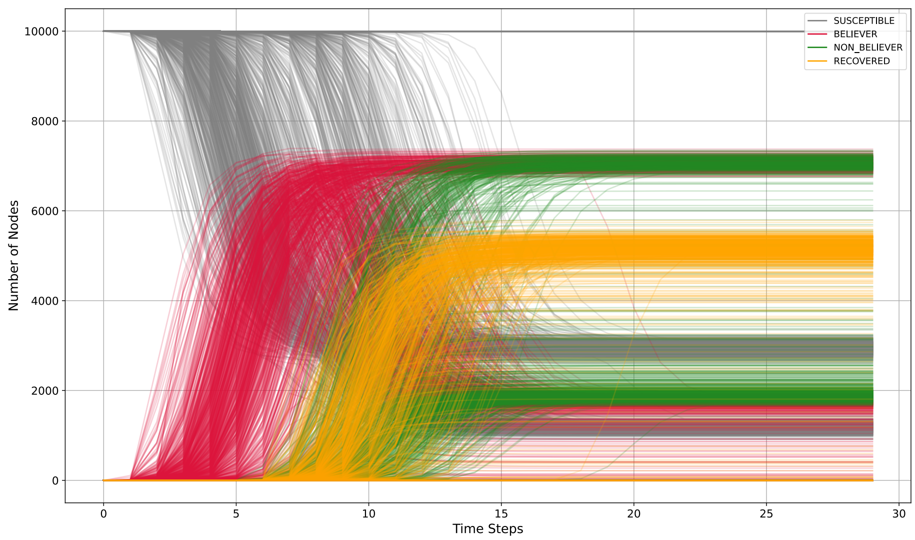


Figure 4: Evolution of node states over time for 10,000 nodes and across 1,000 simulations. A Non-Believer is introduced at time-step 5. Parameters: $\alpha = 1.5$, $\beta = 0.5$, $\delta = 0.5$, and $\gamma = 0.5$

Figure 4 shows the evolution of node states across 30 time-steps for 1,000 simulations on a network of 10,000 nodes. Each line depicts the trajectory of one simulation, for a given state. In the above figure, the Non-Believer is introduced at time-step 5 and the transition probabilities are all set at 0.5. This entails that all types of influence are equally likely.

In the early phase, it is solely Believers that are active and they convert Susceptible individuals, which leads to an unchecked spread of misinformation and high virality. In the vast majority of simulations, the Believers reach their peak at around time-step 8, with about 75% of the population convinced.

From time-step 5 and onward, the number of Non-Believers begins to rise. Those skeptical nodes curb belief propagation and lead to a wide range of final outcomes. This can be observed by the variation in the green lines. In some simulations, skepticism spreads widely. In others, debunking fails to gain ground.

As Non-Believers interact with Believers, an important number of nodes transitions to the Recovered state. This contributes to the decline of the number of Believers after the peak.

The combined presence of Non-Believers and Recovered nodes provides a measure of overall debunking success. However, despite constant values for $\beta, \delta$ and $\gamma$, we note a great diversity of outcomes across run. That shows how unpredictable the belief spread can be.

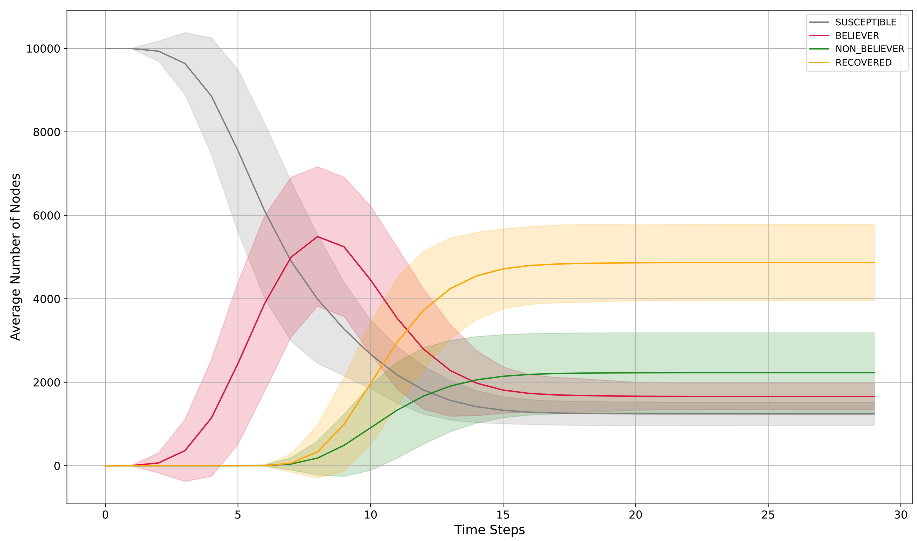


Figure 5: Evolution of node states over time for 10,000 nodes, averaged across 489 simulations (We ran 1000 simulations, but 511 were removed as neither Believers nor Non-Believers reached at least 100 nodes). A Non-Believer is introduced at time-step 5. Parameters: $\alpha = 1.5$, $\beta = 0.5$, $\delta = 0.5$, and $\gamma = 0.5$. The lines show the average number of Susceptible (gray), Believer (red), Non-Believer (green), and Recovered (orange) nodes. Shaded areas represent one standard deviation.

Figure 5 complements Figure 4 as it provides a summary of the average number of nodes in each state over time, across all simulations where the spread of both Believers and Non-Believers have reached at least 100 nodes.

This aggregate view gives a clearer idea of the typical trajectory of belief spread and debunking by smoothing out the fluctuations that occur in the individuals runs. The width of the colored bands around each curve depicts a standard deviation, showing how much simulation outcomes vary at each time-

step. The variability is the strongest for Believers shortly after skepticism is introduced into the system. We recognize this feature from the divergent outcomes seen in individual simulations. This is an expected outcome of the model, and it demonstrate that it behaves as intended.
As expected, the Believers curve reaches its peak just before the Non-Believers begin to rise. It then steadily declines as both skepticism and recovery take effect. This further confirms a correct implementation of the model.

Over time, the system reaches a steady state, but the final distribution of states still displays a high level of spread, especially for Recovered and Non-Believers. This is due to the model's noted randomness.

We proceed now to further analysis with figure 6 below. This graph illustrates the average state evolution when skepticism is introduced at time-step 14.

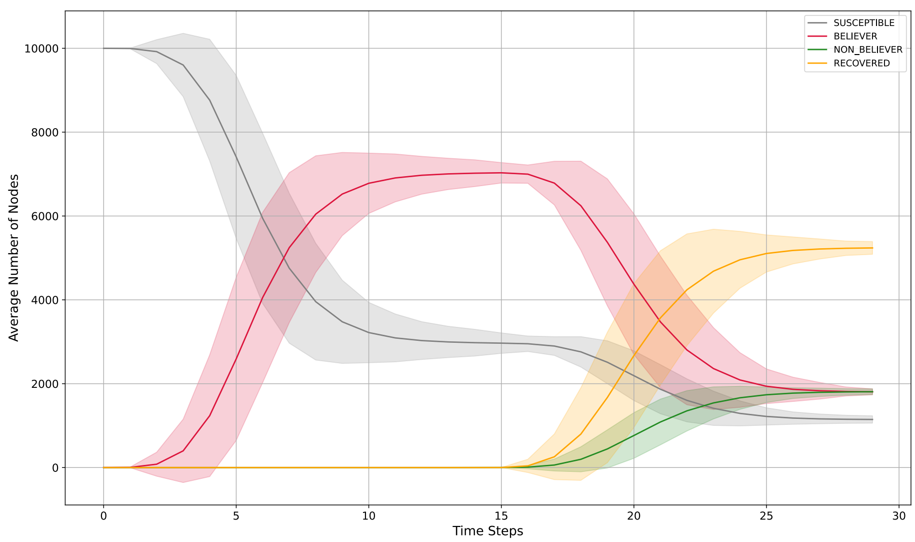


Figure 6: Evolution of node states over time for 10,000 nodes, averaged across 428 simulations (We ran 1000 simulations, but 572 were removed as neither Believers nor Non-Believers reached at least 100 nodes). A Non-Believer is introduced at time-step 14. Parameters: $\alpha = 1.5$, $\beta = 0.5$, $\delta = 0.5$, and $\gamma = 0.5$. The lines show the average number of Susceptible (gray), Believer (red), Non-Believer (green), and Recovered (orange) nodes. Shaded areas represent one standard deviation.

Compared to the earlier case, where a Non-Believer was introduced at time-step 5, we now observe that by the time skepticism is introduced, the number of Believers has already reached a local maximum and remains stable for several time-steps. This indicates that a steady state has been achieved among Believers.

As a result, the corrective influence of the Non-Believer takes place in a population where belief spread has already peaked. Introducing a Non-Believer at this or any later time step would therefore be unlikely to change the final outcome.

Figures 5 and 6 illustrate two distinct phases in the introduction of Non-Believers: one where the spread of Believers has not yet stabilized and another where it has reached a steady state. In the next section, we present a more detailed analysis of these phases.

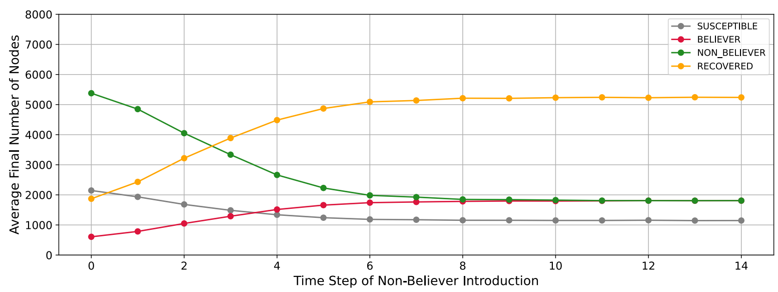


Figure 7: Average final state distributions of nodes in the network across 1,000 simulation runs, for varying debunking timing. Parameters: $\alpha = 1.5$, $\beta = 0.5$, $\delta = 0.5$, and $\gamma = 0.5$.

To understand how the timing of debunking intervention affects the spread of beliefs, we now examine figure 7. It depicts the average final state distributions of nodes in the network across 1,000 simulation runs. Simulations require established spread ($\geq$ 100 nodes) by both Believers and Non-Believers to be included. The x-axis shows the time-step at which the first Non-Believer (debunker) is introduced, ranging from time-step 0 to 14. On the y-axis, we find the average number of nodes in each of the four possible states when steady state has been reached (i.e., at $t = 30$).

As this is an aggregate view, it smoothes out individual variability between runs and helps study general trends. It allows us to observe how the effectiveness of debunking is influenced by the timing of intervention and the persuasiveness of the initial belief.

In the next section, this network set-up is used as a baseline which we compare subsequent findings with. We keep the same method with regards to averaging, finding the steady state and requiring established spread of both Believers and Non-Believers.

## 4.3 Belief strength and convergence

We now proceed with simulations that examine varying transition probabilities, thereby testing different levels of persuasiveness and resistance within the network. These parameters indicate how likely individuals are to influence others, or to be influenced by either misinformation or corrective messages. By adjusting these probabilities, we try to reflect real-life situations and how people actually behave. In reality, individuals may encounter highly convincing misinformation or unconvincing skepticism.

As the simulation progresses, the system converges to a steady state in which no further transitions occur. This happens when all nodes enter an absorbing state or no longer have active neighbors who can influence them. In our model, each node has only one chance at influencing its neighbor, which does accelerate the stabilization. By observing the system at steady state, we analyze how belief and skepticism interact and stabilize.

### 4.3.1 Persuasiveness of belief

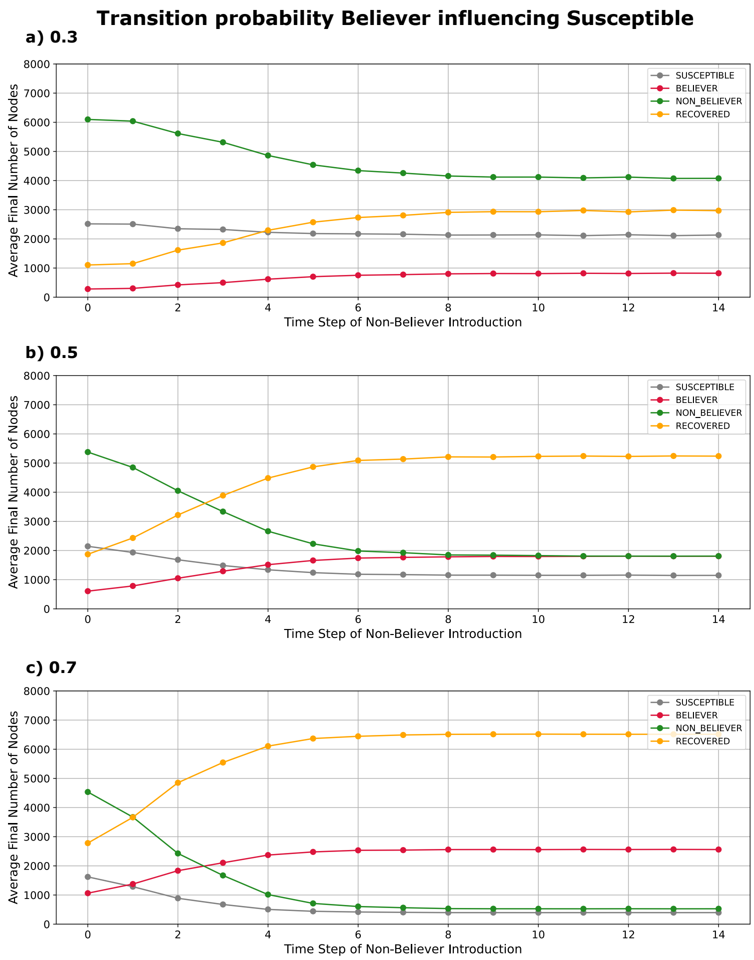


Figure 8: Final state distributions for varying belief persuasiveness ($\beta$) and debunking timing. Subplots show results for $\beta$ values of 0.3 (a), 0.5 (b), and 0.7 (c). Other parameters: $\alpha = 1.5$, $\delta = 0.5$, and $\gamma = 0.5$.

Figure 8 presents three graphs depicting the average final state distributions of nodes in the network across 1,000 simulation runs and subplot b is indentical to figure 7. The three subplots correspond to different values of the belief persuasiveness parameter $\beta$, which determines how easily belief can spread through the network.

All three subplots in figure 8 show that when debunking intervention is delayed, the system eventually reaches a steady state. This indicates that, in these late-intervention scenarios, the primary spread of belief has largely concluded before the introduction of the Non-Believer. Comparing the steady state across the three subplots, we see that the number of Recovered nodes increases as $\beta$ increases. This can be explained by the fact that there are more Believers to convert for larger $\beta$. However, not all Believers are converted, which explains why the final number of Believers also rises with increasing $\beta$. The number of Non-Believers is much lower for larger $\beta$; the reason for this is again that we have more Believers for larger $\beta$ and thus there are few Susceptible nodes available for Non-Believers to convert. Finally, and as expected, the number of Susceptible nodes consistently declines as $\beta$ increases, as individuals are more likely to get influenced.

We now focus on the part of figure 8 where intervention is introduced before the system reaches a steady state. This corresponds to situations where both Believer and Non-Believer/Recovered actively spread simultaneously. When the Believer and Non-Believer nodes are introduced at the same time, we see that the spread of Non-Believers consistently outcompetes the spread of Believers. The reason for this outcome is that both the Non-Believer and Recovered states are absorbing, while the Believer and Susceptible states are not. As the Non-Believer is introduced later and later, we eventually observe a crossover point, where the number of Non-Believer nodes is surpassed by the sum of Believer and Recovered nodes (the number that did believe, or do believe). At higher $\beta$ values, we reach this crossover sooner. For $\beta = 0.3$ we hardly reach this crossover prior to reaching steady state.

### 4.3.2 Debunking strength

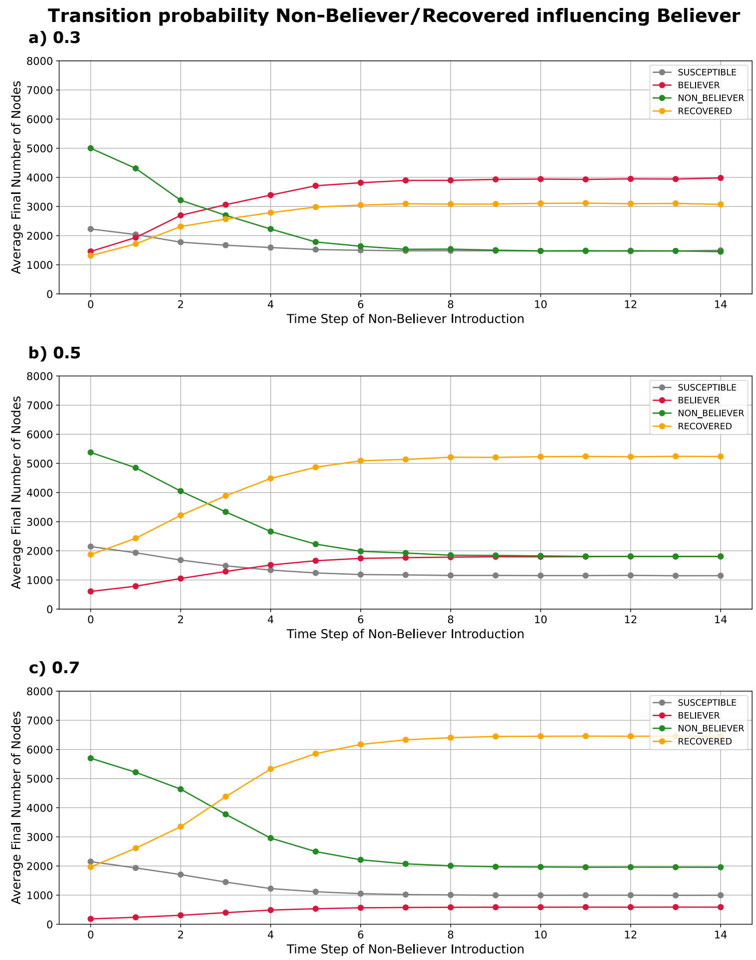


Figure 9: Final state distributions for varying debunking strength ($\gamma$) and debunking timing. Subplots show results for $\gamma$ values of 0.3 (a), 0.5 (b), and 0.7 (c). Other parameters: $\alpha = 1.5$, $\beta = 0.5$, and $\delta = 0.5$.

In figure 9, we now examine how the system responds to varying levels of debunking strengths ($\gamma$). As in figure 8, we see that when debunking intervention is delayed, a steady state is reached. When debunking occurs earlier, we here again can observe how both Believers and Non-Believers/Recovered actively spread simultaneously.

Within the steady state region, we notice that the combined total of Believer and Recovered nodes remains constant across different values of $\gamma$. This is because $\gamma$ directly affects the probability of a Believer transitioning to the Recovered state, but ultimately does not change the total number of nodes that have been influenced (either as Believers or Recovered). As a consequence of the sum of Believer and Recovered being constant, the sum of Susceptible and Non-Believer must also be constant.

However, we also observe that the number of Susceptible nodes decreases as $\gamma$ increases. This reduction is explained by the fact that Recovered nodes can influence Susceptible nodes to become Non-Believers. As $\gamma$ increases, the probability of a Believer becoming Recovered rises, which in turn increases the likelihood that Susceptible nodes will be converted to Non-Believers.

### 4.3.3 Skeptic spread strength

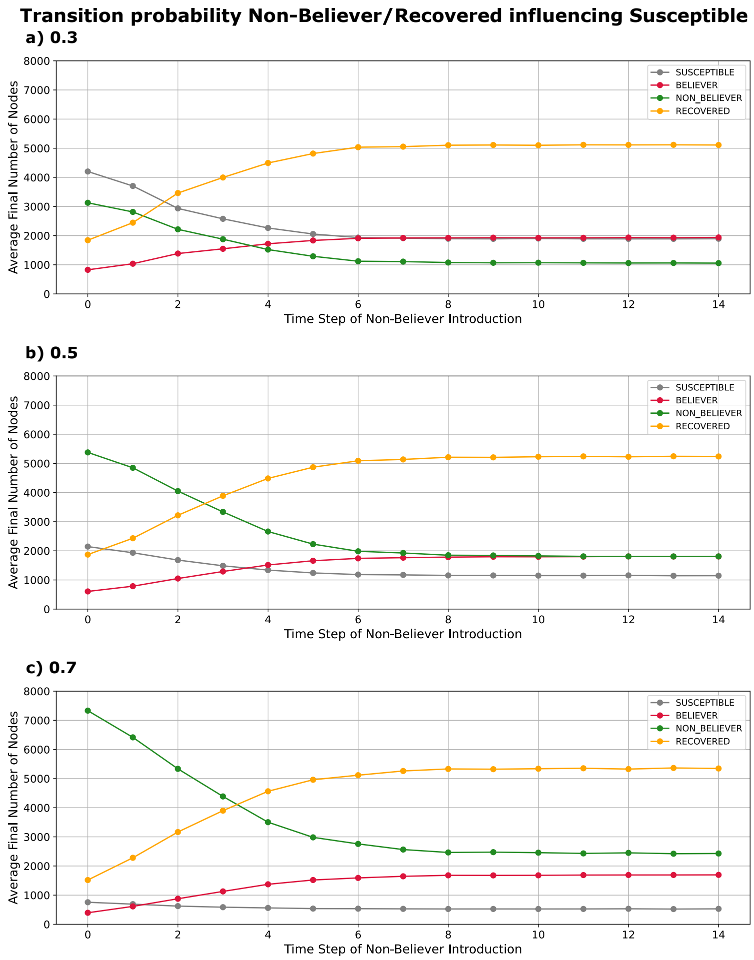


Figure 10: Final state distributions for varying skeptic spread strength ($\delta$) and debunking timing. Subplots show results for $\delta$ values of 0.3 (a), 0.5 (b), and 0.7 (c). Other parameters: $\alpha = 1.5$, $\beta = 0.5$, and $\gamma = 0.5$.

In the graphs shown in figure 10, we vary the transition probabilities for Non-Believer or Recovered individuals to influence Susceptible individuals ($\delta$), and observe the final distribution of states. Just as in figures 9, and 8, we observe that a steady state is consistently reached when debunking intervention is delayed.

Within the steady state region, the number of Recovered nodes slightly increases as $\delta$ increases. This effect occurs because higher values of $\delta$ lead to more Susceptible nodes becoming Non-Believers, who can then convert Believers into Recovered nodes. This is however a secondary effect, so its impact on the system remains limited. The same mechanism explains the small decrease in the number of Believers with increasing $\delta$.

The strongest effect $\delta$ has on the system is, as expected, that:

1. the number of Susceptible nodes decreases as $\delta$ increases,

2. the number of Non-Believers increases with increasing $\delta$. This effect is especially strong when debunking is introduced early.

These observations are in line with the expected outcomes, as a higher $\delta$ strengthens the ability of Non-Believers and Recovered individuals to convert Susceptible nodes.

## 4.4 Network structure

In order to study how network structure can influence belief propagation and correction, we vary the Pareto parameter $\alpha$, which sets the degree distribution of the network. All previous simulations were carried out with $\alpha = 1.5$, but we now introduce variations in $\alpha$ as an additional angle of analysis. The aim is to study how different network structures can affect the spread of misinformation and the effectiveness of debunking.

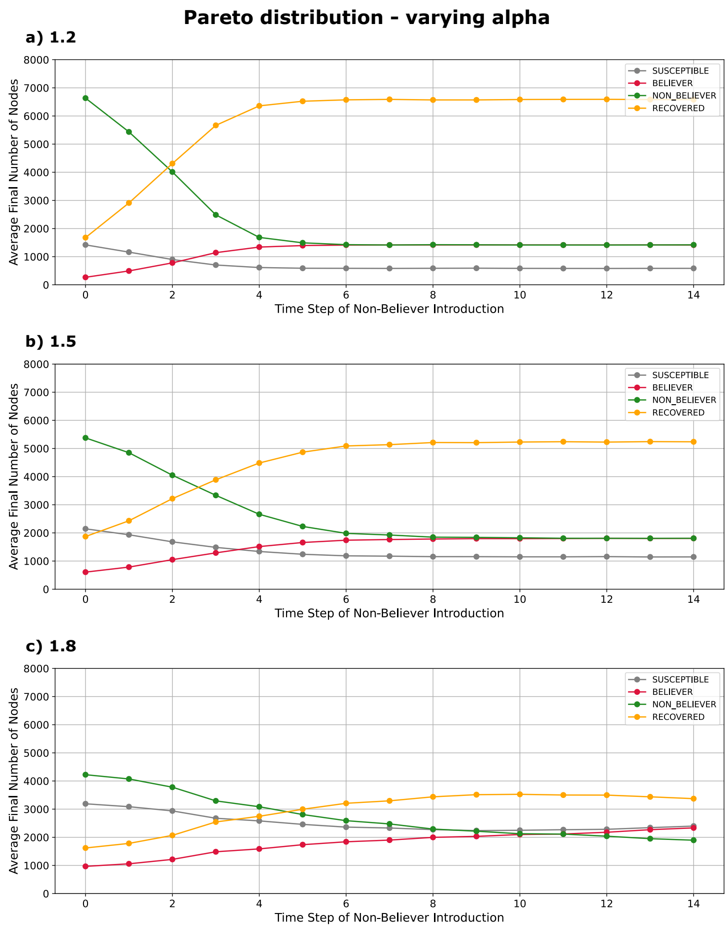


Figure 11: Impact of Pareto degree distribution for varying $\alpha$. Subplots show results for $\alpha$ values of 1.2 (a), 1.5 (b), and 1.8 (c). Other parameters: $\beta = 0.5$, $\delta = 0.5$, and $\gamma = 0.5$.

The three subplots in figure 11 show how the average final state distribution of the network varies depending on when the first Non-Believer is introduced. The difference between these subplots lies in the value of the Pareto parameter $\alpha$, which shapes the network's degree distribution..

The first plot, corresponding to $\alpha = 1.2$, represents a highly heavy-tailed network. This entails that a small number of nodes have very high degrees and most have very few. This structure enables a quick spread of misinformation and early debunking is therefore critical. If Non-Believers are not introduced quickly, misinformation can quickly propagate to a large portion of the network. When debunking is initiated early, however, it successfully converts many believers into recovered nodes.

When $\alpha = 1.8$, the network becomes more balanced, with most nodes having a similar number of connections. All state curves flatten; the influence of hubs is reduced and the system is less sensitive to the timing of debunking introduction. Both misinformation spread and correction take place more gradually, and the final state distribution of nodes is quite stable, regardless of when debunking is introduced.

At $\alpha = 1.5$, the network represents an intermediate case. Debunking remains effective, but its timing is less critical. The final state is more balanced between Believers and Non-Believers. For early debunking, the Non-Believer group is smaller than at $\alpha = 1.2$. The Recovered individuals still dominate when the Non-believer introduction is delayed. The number of Believers slightly increases, and Susceptible nodes are quickly eliminated.

These observations show how the network's degree distribution, controlled by $\alpha$, has a strong influence on the effectiveness and timing of debunking efforts.

# 5 Discussion

## 5.1 Time is of the essence

The findings of this thesis highlight the critical importance of timing in the efforts to counter the spread of misinformation. We show that the efficiency of introducing skeptical agents (Non-Believers) is linked to how early this intervention occurs. Early introduction of Non-Believers, before misinformation becomes widespread, enhances the impact of corrective effort. This effect is analogous to a vaccine, where early administration prevents large infection. Additionally, as Non-Believers actively try to correct Believers, we see a secondary effect where the early spread of Non-Believers effectively corrects individuals as soon as they become Believers.

However, even when interventions are delayed and occur after the belief in the misinformation has already reached a steady state, debunking can still be effective. In these cases, the timing of intervention is less critical, and the focus should shift to the quality of debunking (debunking effectiveness rate $\gamma$). Our simulations clearly illustrate this point. For instance, with weak debunking (figure 9, $\gamma = 0.3$) approximately 40 % of the population remains Believers, while a better-implemented debunking (figure 9, $\gamma = 0.7$) results in only 5% of the population remaining Believers. Before debunking was introduced, 70% of the population were Believers in both cases.

## 5.2 Importance of the network's structure

The network structure plays an important role in helping determine the optimal timing for introducing debunking efforts. In networks with a few highly connected individuals (hubs), it is paramount to start the debunking effort as soon as possible.
If debunking is delayed in such networks, the misinformation quickly spreads throughout the network, and the focus has to shift from preventing misinformation to converting existing Believers into Recovered individuals. With an early intervention, the design of the network can be used to one's own advantage and quickly spread skepticism throughout the population.

## 5.3 Debunking effectiveness and spreading strengths

In the case of early interventions, the scandal value ($\beta$) of the fake news plays a minor role in determining the initial spread, as misinformation is quickly countered by the actions of Non-Believers and Recovered individuals. However, a high scandal value still affects the final population distribution, as it leads to a larger proportion of Recovered individuals. The proportion of Believers remains low, as many nodes are briefly influenced but in the end reject the misinformation and transition to the Recovered state.

For instance, when debunking is introduced immediately, we see that a low scandal value (Figure 8, $\beta = 0.3$) results in less than 5% of the population remaining as Believers and around 10% becoming Recovered. In contrast, a large scandal value ((Figure 8, $\beta = 0.7$), leads to approximately 10% of the population remaining Believers, and almost 30% becoming Recovered.

Debunking can also be designed to specifically target Susceptible individuals (Skepticism spread strength, $\delta$). We have shown that such debunking is especially effective when introduced as soon as possible. By enlightening the population prior to being exposed to misinformation, this approach effectively acts as a ”vaccine”, preventing the spread of disinformation.

## 5.4 When and how to intervene?

Based on the findings in this thesis, we recommend that debunking interventions be initiated as soon as possible, even if that means the debunking might not be efficiently implemented. Early intervention prevents misinformation from spreading widely and can act as a preemptive measure.

If misinformation has already reached a critical mass, the focus should shift to designing a well-planned debunking campaign. The temporary spread cannot be stopped, but a well-executed and curated debunking effort could still lead to a substantial proportion of the population recovering from misinformation.

It is important to note, however, that these recommendations are based solely on the scenarios analyzed in this thesis. More advanced debunking strategies may very well yield better results.

## 5.5 Possible improvements ahead

In this project, we made the simplifying assumption of modeling the spread of a single piece of news. However, individuals in real-world settings are rarely exposed to just one message. They are instead encountering a multitude of news items through different channels and media and they are all shared simultaneously. Each source competes with others for attention. To reflect this complexity, future work could could explore a multi-type information spread model. Such extensions would allow us to simulate the spread of competing pieces of information and study how individuals form or reinforce beliefs when exposed to conflicting messages.

Another potential development would involve redefining the timing of intervention. We could introduce a debunking strategy where the first Non-Believer enters the network when a certain proportion of it has become Believers. This scale-based setup could reveal tipping points beyond which debunking becomes ineffective. It would also align more closely with how real-world interventions often respond to observed prevalence, such as media fact-checking efforts coming into play once misinformation reaches a visible threshold.

Finally, we could consider introducing a variable time delay before individuals attempt to influence their neighbors. In the current model, spreading occurs immediately in the next time-step after a node becomes active. This feature often leads to very swift propagation. As a result, the timing of the introduction of Non-Believers may have limited effect, given that a large portion of the spread has already unfolded. A more realistic alternative could be to assign each individual a waiting time drawn from, for example, a geometric distribution, before they attempt to influence others. This modification would slow the spread and thereby render the timing of debunking more critical and relevant. Different delay distributions could also be assigned to Believers and Non-Believers to reflect differences in how actively they spread their views. This extension was not implemented in the current model, but it represents a valuable direction for future work.

# References


[1] Bateman, J. and Jackson, D. (2021). Countering disinformation effectively: An evidence-based policy guide. Technical report, Carnegie Endowment for International Peace.

[2] Britton, T. (2010). Stochastic epidemic models: A survey. *Math. Biosci, 225, 24-35.*

[3] Deijfen, M., Rosengren, S., and Trapman, P. (2018). The tail does not determine the size of the giant. *J Stat Phys 173, 736-745.*

[4] Deijfen, M. and van der Hofstad, R. (2016). The winner takes it all. *The Annals of Applied Probability, 26(4), 2419–2453.*

[5] Gut, A. (2009). *An Intermediate Course in Probability.* Springer Texts in Statistics. Springer, 2nd edition.

[6] Molloy, M. and Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7(3):295–305.

[7] Ross, S. M. (2019). *Introduction to Probability Models.* Academic Press, London, United Kingdom, 12th edition.

[8] Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science, 359(6380), pp. 1146–1151.*