Audun Sørheim

Misinformation...

Any undertitle is written here

Master's thesis in Physics and Mathematics
Supervisor: Astrid Silvia de Wijn
Co-supervisor: Raffaela Cabriolu
January 2026

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU

ABSTRACT

Write an abstract/summary of your thesis, and state your main findings here

A summary should be included in both English and any second language, if this is applicable, regardless if the thesis is written in English or in your preferred language. These should be on separate pages, the English version first.

PREFACE

Write the preface of your thesis here.

You may include acknowledgements and thanks as part of your preface on this page, or you may add it as a new chapter after the preface.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

List of all abbreviations in alphabetic order:

- **BA**: Barabási-Albert
- **CD**: Cognitive Dissonance
- **DHT**: Distributed Hypothesis Model
- **ER**: Erdős-Rényi
- **NTNU**: Norwegian University of Science and Technology

CHAPTER

ONE

INTRODUCTION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.1 Motivation

Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae dicta sunt explicabo. Nemo enim ipsam voluptatem quia voluptas sit aspernatur aut odit aut fugit, sed quia consequuntur magni dolores eos qui ratione voluptatem sequi nesciunt. Neque porro quisquam est, qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit, sed quia non numquam eius modi tempora incidunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim ad minima veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur? Quis autem vel eum iure reprehenderit qui in ea voluptate velit esse quam nihil molestiae consequatur, vel illum qui dolorem eum fugiat quo voluptas nulla pariatur?

At vero eos et accusamus et iusto odio dignissimos ducimus qui blanditiis praesentium voluptatum deleniti atque corrupti quos dolores et quas molestias excepturi sint occaecati cupiditate non provident, similique sunt in culpa qui officia deserunt mollitia animi, id est laborum et dolorum fuga. Et harum quidem rerum facilis est et expedita distinctio. Nam libero tempore, cum soluta nobis est eligendi optio cumque nihil impedit quo minus id quod maxime placeat facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a sapiente delectus, ut aut reiciendis voluptatibus maiores alias consequatur aut perferendis doloribus asperiores repellat.

1.2 Project description

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

1.2.1 Stakeholders

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CHAPTER

TWO

MSINFORMATION, HUMAN BEHAVIOR AND MODELLING

This chapter will contain the relevant theory of psychological effects, human behavior, and sociology which help drive the dynamics of the Distributed Hypothesis Testing (DHT) model. It is heavily inspired by, and builds upon my project thesis "Beliefs and the Propagation of Misinformation in Social Networks" at NTNU in 2025 [1], specifically sections 2.1-2.5 "Existing Literature".

This chapter is split into two main parts. The first section 2.1 presents the researched literature from sociology and psychology related to misinformation and belief dynamics. It presents psychological and sociological effects and aspect which drive misinformation propagation, on both a collective and individual level.

The second section connects the chosen behavioral insights from the first section into the DHT model used in this thesis. It explains how human behavior, such as non-Bayesian belief updating, confirmation bias, and flexibility in belief change, are included in the model. The section will reason for the choices and that they will make the model more fit to simulate misinformation propagation.

2.1 Misinformation and Human Behavior

The literature presented in this section is on the topic of misinformation through the lens of sociology and psychology. That is, what kind of human behavior affects the spread of misinformation, both on an individual and on a group level. To be able to model misinformation in a satisfactory way, it is important to understand what misinformation is, how it spreads, how it dies down, and how humans affect the propagation. The literature showcased in this section has been chosen as they cover and represent research on different aspects of misinformation.

2.1.1 Disinformation Campaigns and Polarization

- Disinformation is often collaborative; users unintentionally amplify false narratives.

- starbird2019 present three case studies:

  - The Internet Research Agency's social media campaigns before the 2016 U.S. election.
  - Discrediting of the White Helmets during the Syrian Civil War.
  - Crisis-event conspiracy theories and alternative media reinforcement.

- Polarization weakens trust in institutions and impedes constructive discourse.
- Polarized networks become resistant to factual correction (low flexibility).

Citations: starbird2019; bail2018.

2.1.2 Belief Updating and the Backfire Effect

- nyhan2010 introduced the *backfire effect*: corrections can strengthen false beliefs.
- Later studies show the effect is overstated; fact-checking works short-term but fades (nyhan2021).
- Demonstrates the need to include a parameter for belief flexibility/inertia.

Citations: nyhan2010; nyhan2021.

2.1.3 Flexibility and Belief Change

Conclusion with flexibility. It is very important that flexibility 1 makes the system deterministic. Having flexibility includes the history in beliefs for the agents. Yay! With flexibility 1 the system can oscillate between two solutions. Stiffness.

- **Psychological flexibility:** ability to adapt thoughts and actions to new evidence (kashdan2010).
- Low flexibility (rigidity) ⇒ belief inertia, persistence of misinformation (westhoff2024).
- **Cognitive flexibility / Actively Open-Minded Thinking (AOT):** willingness to evaluate and revise beliefs (stanovich2023).
- Neurological evidence: flexible individuals integrate contradictory information more effectively (romero2022).
- Flexibility corresponds to parameter ψ_i controlling belief update magnitude.

Citations: kashdan2010; westhoff2024; stanovich2023; romero2022; nyhan2010.

2.1.4 Emotional Drivers, Bias, and Echo Chambers

- **Confirmation bias:** people seek and interpret confirming information (nickerson1998).
- **Signal misinterpretation:** humans distort or underweight disconfirming evidence (defilippis2021).
- **Motivated reasoning:** evidence processing biased by identity protection (kunda1990).
- **Cognitive dissonance:** conflict between ideas leads to attitude reinforcement (festinger1957).
- **Emotions:** high arousal strengthens rigidity and polarization (altoe2022).
- These mechanisms reduce ψ_i (flexibility) and ϕ_{ij} (trust in differing neighbors) $\Rightarrow$ echo chambers.

Citations: nickerson1998; defilippis2021; kunda1990; festinger1957; altoe2022.

2.1.5 Combating Misinformation: Correction, Education, and Flexibility

- **Fact-checking:** corrects false beliefs temporarily (nyhan2021).
- **Scientific literacy & critical thinking:** correlate with lower conspiracy belief (fasce2019).
- Attitude toward critical thinking matters as much as ability.
- **Influencers and audience capture:** social incentives discourage opinion change (jurg2020).
- **Building flexibility:** open-mindedness and repeated correction yield more durable resistance.

Citations: nyhan2021; fasce2019; jurg2020; kashdan2010; westhoff2024.

2.1.6 Summary of Behavioral Insights

- Misinformation persists due to limited flexibility, emotional reinforcement, and biased signal interpretation.
- Corrective information decays unless reinforced through repetition or affective framing.
- Polarization and echo chambers arise from low ψ_i (flexibility) and low ϕ_{ij} (cross-group trust).
- These mechanisms justify extending DHT with parameters for bias, misinterpretation, and belief inertia.

2.2 The Distributed Hypothesis Model and Human Behavior

2.2.1 Classical DHT Models

- Agents exchange beliefs using Bayesian likelihoods and consensus rules.
- Assumes rational, unbiased signal interpretation.
- Convergence to the true hypothesis is exponential under these assumptions.

Citations: blum1997; nedic2017.

2.2.2 Adapting DHT to Human Agents: Bias, Misinterpretation, and Flexibility

- Humans interpret contradictory evidence non-Bayesianly (defilippis2021).
- **Signal misinterpretation:** down-weighting of disconfirming information.
- **Confirmation bias:** reduced influence of disagreeing neighbors (nickerson1998).
- **Flexibility parameter** ψ_i: degree to which beliefs update.
- **Interpersonal weight** ϕ_{ij}: trust or openness to others.
- Together, these parameters simulate realistic opinion dynamics.

Citations: defilippis2021; nickerson1998; kashdan2010; westhoff2024.

2.2.3 Implications for Misinformation Dynamics

- Low ψ_i (inflexibility) $\Rightarrow$ belief inertia and slow convergence.
- High confirmation bias (low ϕ_{ij}) $\Rightarrow$ polarization and echo chambers.
- Signal misinterpretation $\Rightarrow$ persistence of misinformation despite correction.
- Increasing ψ_i (flexibility) or ϕ_{ij} (trust) improves collective accuracy.

Citations: nyhan2010; defilippis2021; kashdan2010; westhoff2024.

2.3 Summary and Transition

- Misinformation and polarization stem from psychological limits in belief updating.
- Classical DHT assumes rational agents; human agents deviate through bias and inertia.
- The extended model includes:

- **Confirmation bias** (ϕ_{ij}) — selective trust in information sources.
- **Signal misinterpretation** — non-Bayesian weighting of evidence.
- **Belief flexibility** (ψ_i) — inertia or readiness to change beliefs.

- These components provide a realistic behavioral foundation for the simulation results in Chapter 6.

CHAPTER

THREE

GRAPHS

Here all relevant theory on graphs will be presented, only graphs used for simulation will be presented, and the methods for creating them will be described in section 5.1. This chapter is heavily inspired by my project thesis "Beliefs and the Propagation of Misinformation in Social Networks" at NTNU in 2025 [1]. Specifically, this chapter is inspired by subsection 3.1 "Graph Theory".

3.1 Graphs

CHAPTER

FOUR

DISTRIBUTED HYPOTHESIS TESTING MODEL

This chapter will contain all the relevant theory of the Distributed Hypothesis Testing (DHT) model, not including graphs or sociology. The model is applied on a network or graph of nodes, which are able to interact. First, I will introduce the developments of the model since its inception, then I will define it as it has been used for this master's thesis (with relevant constraints). Lastly I will present the algorithms which have been used and any potential differences from the project thesis preceding this master's thesis.

This chapter is inspired by my project thesis "Beliefs and the Propagation of Misinformation in Social Networks" at NTNU in 2025 [1] (specifically section 3.2), Diana Riazi's papers and PhD thesis [2, 3, 4, 5], and the paper on the DHT model by Anusha Lalitha et al. [6].

4.1 History

This section contains a brief history of developments of the DHT model from its inception by Robert Tenney and Nils Sandell Jr. in 1981[7] until its formulation by Lalitha et al. in 2014[6].

The earliest work done on the subject was motivated by the fact that in a distributed sensor network, it can be costly or inefficient to make all sensors communicate with a central hub. Thus, the sensors must be able to process signals themselves using an optimal decision rule. Tenney et al. created a model based on the formulation of decentralized hypothesis testing [7]. They found that the optimal decision rule is that the sensors perform likelihood ratio tests on the statistically independent observations they make.

Over the following decades, research further developing the model was mostly concerned on distributed sensor networks. Firooz determined the structure of the optimal decision rule given an arbitrary number of sensors and hypothesis [8]. Viswanathan et al. made a review study of the field in 1997 [9]. In 2004 Alanyali et al. applied belief propagation, and thus interaction between sensors, to make the network reach a consensus concerning an observation [10]. Halme et al. included spatial dependency in the model [11].

In 2014 (and revised in 2018), Lalitha et al. published "Social Learning and Distributed Hypothesis Testing" [6]. This paper included social learning in the

DHT model, implying that the nodes in the network interact in such a way that they can learn from each other. Lalitha et al. introduced the notion of public and private beliefs, such that interactions between nodes happened through public beliefs. These were formed by a Bayesian update of private beliefs using likelihood functions made from observations. This formulation of the model was intended to be used on directed graphs.

In 2024 and 2025, Diana Riazi published three papers with her PhD thesis using the DHT model to simulate misinformation propagation in social networks [2, 3, 4, 5]. This is the first and, so far, only study using the DHT model to simulate propagation of beliefs in a network consisting of humans. The mathematical formulation will be presented in section 4.2.

4.2 Mathematical Formulation

This section will provide the full mathematical formulation of the DHT model used in this project. The methods for coding and the setups of the simulations will be presented in chapter 5. This section entails how the model works, on what it is applied, the updating algorithm, and the dynamics applied

The DHT model is applied on a graph (or network) consisting of nodes and edges (see section for graph more on graphs). The networks utilized here are *directed networks*, which

4.3 Differences from Project Thesis

$$g(x, s) = \exp\left[-\left(\frac{x-1}{s}\right)^2\right] \tag{4.1}$$

$$smoid(x, a) = \frac{1}{1 + \exp\left[-a(x - \frac{1}{2})\right]} \tag{4.2}$$

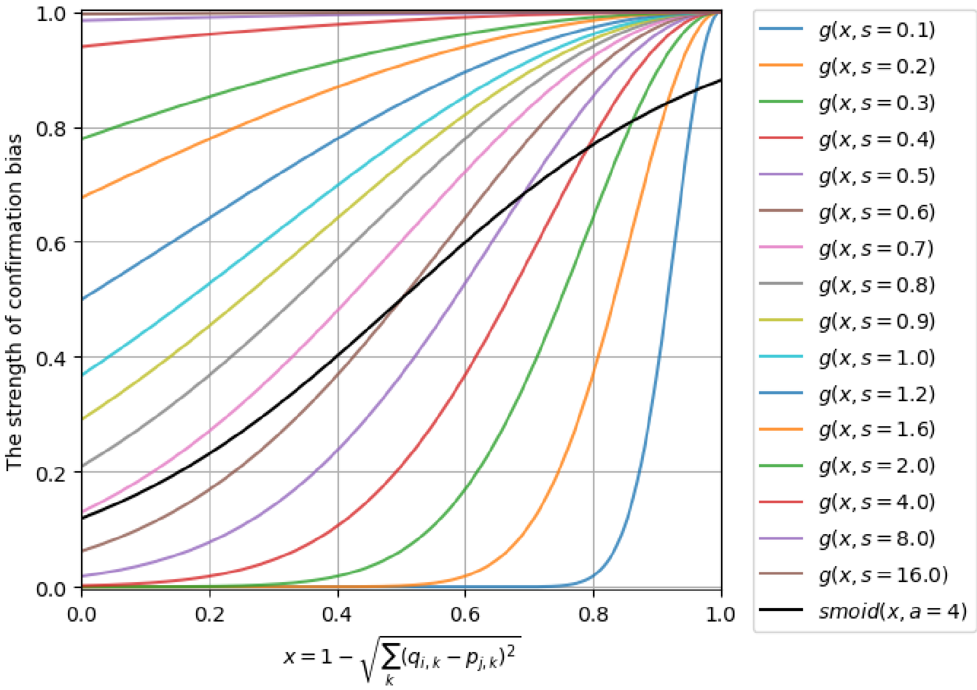

Figure 4.3.1: The strength of confirmation bias which will be multiplied with weights to neighbors. The functions $g(x, s)$ and $smoid(x, a)$ refer to equations 4.1 and 4.2 respectively. $g(x, s)$ is the new representation of confirmation bias used here and $smoid(x, a)$ is the old representation used in the project thesis.

CHAPTER

FIVE

METHODS

Include the complete description of the methods used in your research here.

Below is an example of how subsectioning works. The sections and subsections will be included in the table of contents, while subsubsections will not be in the table of contents but still have their own title in the text.

5.1 Creating graphs

5.1.1 Subsection one

5.1.2 Phase Diagram Simulations

5.1.2.1 Subsubsection Two

5.1.3 Subsection Two

5.2 Section two

Table 5.1.1: Simulation settings used for phase diagram.

Parameter	**Value**
N (agents)	4096
M (Number of hypothesis)	4
Graph type	square
k (for ER network)	10
m (for BA network)	5
Number of iterations	200
Number of simulations	200
σ_{draw}	0.75
$\sigma_{\text{likelihood}}$	0.75
Signal type	Gaussian
Confirmation bias active	True
Number of conspirators	0
Number of mega-nodes	0
Belief updating rule	Linear

Table 5.1.2: The values for the parameter F (flexibility strength) for the simulations in the phase diagram.

F							
0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400
0.450	0.500	0.550	0.600	0.650	0.700	0.750	0.800
0.850	0.900	0.950	0.980	0.985	0.990	0.995	1.000

Table 5.1.3: The values for the parameter s for the simulations in the phase diagram.

s							
0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.9	1.0	1.2	1.6	2.0	4.0	8.0	16.0

CHAPTER

SIX

RESULTS

Here results will be presented, but not discussed.

6.1 More figures

page fill page fill.

CHAPTER

SEVEN

DISCUSSION

Discuss your results here.

7.1 Future work

Include a section about what should or could be done in future research, or explain any recommended next steps based on the results you got. This should be the last section in the discussion.

CHAPTER

EIGHT

CONCLUSIONS

Give a concise summary of your research and finding here, and include a short summary of any future work as well.

Conclusion with flexibility. It is very important that flexibility 1 makes the system deterministic. Having flexibility includes the history in beliefs for the agents. Yay! With flexibility 1 the system can oscillate between two solutions. Stiffness, **inertia**.

REFERENCES

[1] Audun Sørheim. *Beliefs and the Propagation of Misinformation in Social Networks*. 2025.

[2] Diana Riazi and Giacomo Livan. "Public and private beliefs under disinformation in social networks". In: *Physica A: Statistical Mechanics and its Applications* 637 (2024), p. 129621. DOI: https://doi.org/10.1016/j.physa.2024.129621.

[3] Diana Riazi and Giacomo Livan. *Who should fight the spread of fake news?* 2024. URL: https://arxiv.org/abs/2410.08753.

[4] Diana Riazi and Giacomo Livan. "Mitigating Disinformation in Social Networks through Noise". In: *Advances in Complex Systems* (Feb. 2025). DOI: 10.1142/S0219525925500079.

[5] Diana Riazi. "Agent-based Models of Social Learning in Opinion Dynamics". Ph.D. thesis. University College London, 2025. URL: https://discovery.ucl.ac.uk/id/eprint/10205241/.

[6] Anusha Lalitha, Tara Javidi, and Anand D. Sarwate. "Social Learning and Distributed Hypothesis Testing". In: *IEEE Transactions on Information Theory* 64.9 (2018), pp. 6161–6179. DOI: 10.1109/TIT.2018.2837050.

[7] Robert R. Tenney and Nils R. Sandell. "Detection with distributed sensors". In: *IEEE Transactions on Aerospace and Electronic Systems* AES-17.4 (1981), pp. 501–510. DOI: 10.1109/TAES.1981.309178.

[8] Firooz A. Sadjadi. "Hypotheses Testing in a Distributed Environment". In: *IEEE Transactions on Aerospace and Electronic Systems* AES-22.2 (1986), pp. 134–137. DOI: 10.1109/TAES.1986.310747.

[9] R. Viswanathan and P.K. Varshney. "Distributed detection with multiple sensors Part I. Fundamentals". In: *Proceedings of the IEEE* 85.1 (1997), pp. 54–63. DOI: 10.1109/5.554208.

[10] M. Alanyali et al. "Distributed Bayesian hypothesis testing in sensor networks". In: *Proceedings of the 2004 American Control Conference*. Vol. 6. 2004, 5369–5374 vol.6. DOI: 10.23919/ACC.2004.1384706.

[11] Topi Halme, Martin Gölz, and Visa Koivunen. "Bayesian Multiple Hypothesis Testing For Distributed Detection In Sensor Networks". In: *2019 IEEE Data Science Workshop (DSW)*. 2019, pp. 105–109. DOI: 10.1109/DSW.2019.8755597.

APPENDICES

A - GITHUB REPOSITORY

All code and latex-files used in this document are included in the Github repository linked below. Further explanations are given in the readme-file.

Github repository link

- `https://github.com/ninasalvesen/thesis_latex_template`

B - SIDENOTE STATISTICS

B1 - Some random table

Remember to only include one thing per page in the appendices.

Statistic	**One**	**Two**
Count	387317	283960
Mean	130.66	134.18
Std	248.09	230.32
Q1	31.00	21.00
Median	67.00	63.00
Q3	142.00	159.00
Min	0.00	0.00
Max	14519.00	14253.00

Table B.1: Table of statistics on some sidenote data.

B2 - Some other random table

Statistic	**Three**	**Four**
Count	387317	283960
Mean	130.66	134.18
Std	248.09	230.32
Q1	31.00	21.00
Median	67.00	63.00
Q3	142.00	159.00
Min	0.00	0.00
Max	14519.00	14253.00

Table B.2: Table of statistics on some other sidenote data.