UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Competitive influence maximisation in social networks

by

Sukankana Chakraborty

ORCiD: 0000-0002-2115-8531

A thesis for the degree of
Doctor of Philosophy

4th January 2023

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Electronics and Computer Science

Doctor of Philosophy

Competitive influence maximisation in social networks

by Sukankana Chakraborty

Network-based interventions have shown immense potential in prompting behaviour changes in populations. Their implementation in the real world however, is often difficult and prone to failure as they are typically delivered on limited budgets and in many instances can be met with resistance in populations. Therefore, utilising available and limited resources optimally through careful and efficient planning is key for the successful implementation of any intervention. An important development in this aspect, is the *influence maximisation* framework —which lies at the interface of *network science* and *computer science* —and is commonly used to study network-based interventions in a theoretical setup with the aim of determining best practices that can optimise intervention outcomes in the real world.

In this thesis, we explore the *influence maximisation* problem in a competitive setting (inspired by real-world conditions) where *two* contenders compete to maximise the spread of their intervention (or influence) in a social network. In its traditional form, the *influence maximisation* process identifies the k most influential nodes in a network —where k is given by a fixed budget. In this thesis, we propose the *influence maximisation* model with continuous distribution of influence where individuals are targeted heterogeneously based on their role in the influence spread process. This approach allows policymakers to obtain a detailed plan of the optimal distribution of budgets which is otherwise abstracted in traditional methods. In the rest of the thesis we use this approach to study multiple real-world settings.

We first propose the competitive *influence maximisation* model with continuous allocation of resources. We then determine optimal intervention strategies against known competitor allocations in a network and show that continuous distribution of resources consistently outperform traditional approaches where influence is concentrated on a few nodes in the network (i.e. *k* optimal nodes). We further extend the model to a game-theoretic framework which helps us examine settings with no prior information about competitor strategies. We find that the equilibrium solution in this setting is to uniformly target the network —implying that all nodes, irrespective of their topological positions, contribute equally to the *influence maximisation* process.

We extend this model further in *two* directions.
First, we introduce the notion of adoption barriers to the competitive *influence maximisation* model, where an additional cost is paid every time an individual is approached for intervention. We find that this cost-of-access parameter ties our model to traditional methods, where only *k* individuals are discretely targeted. We further generalise the model to study other real-world settings where the strength of influence changes nonlinearly with allocations. Here we identify *two* distinct regimes —one where optimal strategies offer significant gains, and the other where they do not yield any gains. The *two* regimes also vary in their sensitivity to budget availability, and we find that in some cases, even a *tenfold* increase in the budget only marginally improves the outcome of the intervention.
Second, we extend the continuous allocation model to analyse network-based interventions in the presence of negative ties. Individuals sharing a negative tie typically influence each other to adopt opposing views, and hence they can be detrimental to the influence spread process if not considered in the dynamics. We show that in general it is important to consider negative ties when planning an intervention, and at the same time we identify settings where the knowledge of negative ties yields no gains, or leads to less favourable outcomes.

Contents

List of Figures

Declaration of Authorship

I, Sukankana Chakraborty, declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as:
   – Sukankana Chakraborty, Sebastian Stein, Markus Brede, Ananthram Swami, Geeth de Mel, and Valerio Restocchi. Competitive influence maximisation using voting dynamics. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 978–985, 2019
   – Guillermo Romero Moreno, Sukankana Chakraborty, and Markus Brede. Shadowing and shielding: Effective heuristics for continuous influence maximisation in the voting dynamics. *Plos one*, 16(6):e0252515, 2021

Signed:

Date:

Acknowledgements

Where the mind is without fear and the head is held high;
Where knowledge is free;
Where the world has not been broken up into fragments by narrow
domestic walls;
Where words come out from the depth of truth;
Where tireless striving stretches its arms towards perfection;
Where the clear stream of reason has not lost its way into the dreary
desert sand of dead habit;
Where the mind is led forward by thee into ever-widening thought
and action—
Into that heaven of freedom, my Father, let my country awake.

- Rabindranath Tagore (**Where the mind is without fear** from
Gitanjali).

To Dadai and Dida, I miss you every day.

Definitions and Abbreviations

ER   Erdos-Renyi model

SW   Strogatz-Watts model

BA   Barabasi-Albert model

SIR   Susceptible-infected-recovered model

SIS   Susceptible-infected-susceptible model

Chapter 1

Introduction

1.1 Motivation

Interactions with our peers impact decisions and behaviours in our daily lives
(Moreno (1934); Maxwell (2002); Ritzer et al. (2007)) —often shaping our health
practices, voting choices and addiction patterns (Christakis and Fowler (2007, 2008);
Rosenquist et al. (2010); Rogers (2010)). These behaviours that spread through social
contact, collectively lead to global outcomes (or mass behaviours) in populations
(Solovei and van den Putte (2020); Centola (2010)). Because of this, social networks
have attracted considerable attention from policy-makers as a tool for interventions,
used to promote behaviour change at large scales (Valente (2012); Allcott and
Gentzkow (2017); Steinert-Threlkeld et al. (2017)).

Indeed, network-based interventions have shown immense potential as instigators of
social change (Valente (2012); Snyder (2007)). One of the earliest examples of a
successful intervention was the Taichung study of 1964, where the spread of
information through interpersonal ties increased the uptake of family-planning
methods in the population (Rogers (2010)). Several other studies have since been
conducted to show how social connectivity can be exploited to promote interventions
that contribute to, for instance, better health and improved financial planning in
low-income populations with limited access to technology and information (Banerjee
et al. (2013); Oliver-Williams et al. (2017); Abaluck et al. (2021)).

In addition to promoting social reforms, network-based interventions have also
drawn attention as a means to mitigate the harmful spread of misinformation in social
networks (Massey et al. (2020); Young et al. (2021)). Indeed while on one hand, the
internet has improved our quality of life in several aspects (Catarinucci et al. (2015);
Stavropoulos et al. (2020)), in a world where information is predominantly consumed
through social media (Bergström and Jervelycke Belfrage (2018); Boczkowski et al.

(2017)), it has also precipitated the easy manipulation of masses through the spread of distorted information (Leopold and Bell (2017); Hillstrom (2018)) —often leading to catastrophic impacts that have long-lasting effects on societies (Ricard and Medeiros (2020); Coutts et al. (2013); Flintham et al. (2018)). Consider the many negative effects that have perpetuated through the unchecked use of social media in the recent years (Tench and Jones (2015); Moreno and Whitehill (2014); Vishwanath (2015)) —including protests and uprisings (Guadagno et al. (2010); Scherman et al. (2015); Kharroub and Bas (2016); Maireder and Schwarzenegger (2012)), threats to public health (Dubé et al. (2015); Hotez (2016); Roozenbeek et al. (2020); Siani (2019); Tasnim et al. (2020)) and obstruction of democracy (Allcott and Gentzkow (2017); Pennycook et al. (2018)). Moreover, given the current challenges of imposing regulations in this sector (Bromell (2022); Weiser (2009); Leopold and Bell (2017); Hillstrom (2018)), policy-makers are gradually turning to network-based interventions to immunise populations against misinformation spread (Zarocostas (2020); van der Linden (2022)).

The implementation of network-based interventions in the real world entails many challenges (Valente (2012) —some of which can be handled through the meticulous planning of intervention methods. One way to design efficient and effective interventions, is to first explore it in a theoretical setup that simulates real-world conditions (Yadav et al. (2015, 2016, 2018b,a); Wilder et al. (2018b)). This forms the underlying theme of this thesis where we focus on a theoretical framework, known as *influence maximisation*, to study network-based interventions as a means to regulate collective behaviours in populations. Our goal is to propose efficient algorithms that optimise intervention outcomes in the real world.

While significant strides have been made in this area, a general limitation of theoretical models is that they are often too simplistic in their approach (Newman et al. (2011)). In this thesis, we highlight some of the limitations of existing methods, and propose extensions that aim to alleviate them —ultimately providing preliminary results that can help design empirical experiments to progress research in this area.

1.2   Research outline

The tactical spread of influence in social networks is a multi-faceted problem with attractive applications across many areas (Easley et al. (2010); Jackson (2010)). The *influence maximisation* problem, formalised in the seminal paper by Kempe et al. (2003), aims to exploit interpersonal ties to maximise the adoption of an innovation or a behaviour in a population (Ferguson (2008); Domingos and Richardson (2001); De Bruyn and Lilien (2008)) It is typically framed as an optimisation problem, and is used to identify the most influential individuals in a network —who can maximise the spread of a desired behaviour in the rest of the population. The concept reflects the

two-step flow hypothesis where influence in a population spreads by means of a two-step process (Katz (1957)). In the first step, influential individuals in a population —also termed as *opinion leaders* —adopt a new idea or innovation. Following this, the adoption behaviour propagates from these *opinion leaders* to the rest of the population.

Influence maximisation has been extensively studied in the past, producing a rich body of literature that spans across multiple disciplines (Leskovec et al. (2007); Budak et al. (2011); He et al. (2012)). One way to atomise this exhaustive body of work is to consider how influence flows in the model. As such, we identify two independent branches of models - (i) diffusion models championed by Kempe et al. (2003), and (ii) dynamic models from the field of sociophysics (Sen and Chakrabarti (2014); Galam (2004)).

Diffusion models have been categorically used in *computer science* in the past, to study the spread of influence (or information) in social networks. Some stereotypical examples of diffusion models are the *Independent Cascade* model and the *Linear Threshold* model (Kempe et al. (2003)). Some epidemiological models that study the spread of infection in populations consider a similar approach, with the exception that, individuals may recover from the infection and be either removed from the network, or become susceptible to being infected again (Kermack and McKendrick (1927); Hethcote (2000)).

While such models aptly describe how decisions and behaviours virally spread in social systems, their representation of individual states reflect long-term commitments such as buying a car. The one-off, immutable nature of decisions in these models make them unsuitable for studying settings where individual choices (or opinions) are transient and free of abiding commitments. For example, consider instances where people switch between mobile networks for affordability reasons (Wu and Wang (2005)), or discontinue and resume weight-loss programmes based on how motivated they feel (James (2009)), and in some cases "unfollow" users on Twitter they once chose to "follow", due to lack of reciprocity and informativeness (Moon (2011)).

In contrast, individuals in dynamic models are at liberty to frequently change their decisions —based on repeated interactions with their social neighbourhood (Castellano et al. (2009a); Sen and Chakrabarti (2014)). In this thesis, we consider the dynamic model known as the voter model to capture influence flow in a population (Clifford and Sudbury (1973); Holley and Liggett (1975); Sood and Redner (2005)). This paradigmatic model is characterised by its simple but effective approach to study reality-based social dynamics (Redner (2019); Braha and de Aguiar (2017)). At large, it has been used to study the evolution of opinions in populations, where individuals switch between binary opinion states at a rate proportional to the fraction of opinion shares in their neighbourhood. In the classical voter model, individuals do not display any self-beliefs or biases, and form their opinions strictly through social interactions.

Thus implying an implicit lack of information (or understanding) about the propagating opinions in the network, which also aptly captures our settings of interest described in Section 1.1.

Moreover, a noted advantage of the voter model is that the equations of the flow of influence can be solved analytically for any simplified network structure —however large —eliminating the need to strictly rely on tedious computational methods. Although, such analytical solutions can only be obtained in simple network structures, they are important in their own right as they provide benchmarks for computational tools, and can help interpret numerical results in more complex settings.

1.3   Research contributions

We study the *influence maximisation* problem in the voter model under competitive settings, where two contenders (or controllers) external to the network compete to maximise their influence in the population. Our work is inspired by the settings presented in Section 1.1, where network-based interventions are often used to eradicate undesirable behaviours (e.g. poor health choices, misinformation) in populations. We apply the problem to many real-world settings and determine approaches that work well to contend with competing influences in networks, and in doing so, we make the following contributions to the existing field of research.

1.3.1   Continuous allocation of resources

The voter model has attracted considerable attention within the *influence maximisation* research (Masuda (2015); Kuhlman et al. (2013); Yildiz et al. (2013); Brede et al. (2018, 2019)). Most of this work however, mimics the traditional setting where a limited budget is used to convert (flip or activate) a small number of individuals in the population —who subsequently influence the rest of the population.

This approach focuses on identifying the most influential individuals in the network and abstracts all information about the way in which the budget should be used (or distributed over the network) to maximise influence spread. In the real world, *influence maximisation* efforts are predominantly led with resources such as time and money, and thus a strategy detailing how these resources (and budget) should be utilised and distributed, would be immensely helpful. With this in mind, we propose a novel approach to study *influence maximisation* using continuous allocation of resources where individuals are targeted with varying intensities (or strengths of influence) —based on their role in the influence spread process. Given that allocations are now continuous, it opens up the opportunity for all nodes to be targeted, instead

of a limited few. Traditional models also assume that selected nodes are activated with complete certainty at the start of the influence spread process. However, this is not a straight-forward process in the real world and may involve a lot of uncertainty, which we capture in our model by assuming that influence is applied in an uninterrupted manner (as is observed in digital marketing today), and individuals occasionally interact with this influence to make their decisions.

Furthermore, it has been well-established that dynamics in the voter model either converge to an ordered consensus, or reach a fragmented state as the population reaches an equilibrium (Castellano et al. (2009a)). As consensus is rarely ever achieved in the real world (Garimella et al. (2018); Tucker et al. (2018)), it has motivated significant efforts to study conditions and properties of fragmentation in networks, with special attention to competitive settings. Competitive *influence maximisation* has been fairly well-studied in the voter model (Masuda (2015); Yildiz et al. (2013); Kuhlman et al. (2010)), however most of this work notably assumes prior knowledge of competitor strategies which is not always easy to retrieve in the real world. In our work, we consider both known and unknown competitor allocations, and systematically optimise *influence maximisation* strategies for both instances. Where no explicit information about the competitor strategy is available, we employ a game-theoretic framework (Chasparis and Shamma (2010); Fazeli and Jadbabaie (2012); Masucci and Silva (2014); Fazeli et al. (2016)), to optimise counter-strategies.

1.3.2   Barriers to adoption and nonlinear cost of allocations

Despite our best efforts, network-based interventions may not always yield the outcome we expect in the real world. In some cases, low adoption of interventions result from, what are known as, barriers to adoption (Butler and Sellbom (2002); Vanclay (1992)). Such barriers often result from the lack of financial means to acquire the product, the lack of access to the product or the lack of information about the product, which need to be remedied in order for the external interventions to succeed (Mobarak and Saldanha (2022); Gates (2019)).

With this in mind, we introduce a fixed cost of access parameter in our model. This is the cost paid by an external controller each time they want to access an individual with the intent of influencing them. It is analogous to paying a cost to remove existing barriers of adoption, and consequently priming the network for interventions. For instance, Gates (2019) showed that distributing free contraceptives at children's vaccination centres (that came at an additional cost), increased the uptake of vaccinations among children. This was because the policy attracted more women to visit vaccination centres to obtain safe contraception, where they could be educated about the health benefits of vaccines.

Assuming that the cost of the intervention is the subsidy provided on a product and the effect is the probability of adoption, we also observe a nonlinear relation between the cost and the effect of interventions in many real-world instances. For example, people are often naturally reluctant to adopt innovations or new technologies, unless reasonably subsidised to lower the stakes of adoption (Sardianou and Genoudi (2013); Nicolini and Tavoni (2017)). In this case, smaller subsidies will first have marginal effects on the probability of adoption. However once the subsidy breaches a certain threshold (of affordability and risk), the probability of adoption will increase exponentially. Similarly, when considering interventions that people readily adopt (such as subscription trials), we find that a small initial subsidy (i.e. offering a trial) has a significant effect on the probability of adoption, but any subsequent increases in influence (e.g. increasing the length of the trial period) only marginally increases the probability of adoption, yielding a nonlinear relation between the cost and the effect of influence. With the aim of including these real world settings in our work, we modify our existing model to consider nonlinear costs of allocation —which to the best of our knowledge has not been considered in *influence maximisation* research before.

1.3.3   Effect of negative ties on *influence maximisation* efforts

Finally, it should be noted that a majority of the *influence maximisation* work is focused on networks of friendship, where people influence one another strictly positively, i.e. an adoption of a product or an idea only increases the probability of a subsequent adoption in a social neighbour. However real-world networks often contain negative relationships in addition to positive ties (Leskovec et al. (2010)), which if discounted from the dynamics can erroneously perpetuate undesirable influence in the network (Guha et al. (2004)). With this in mind, we explore the competitive *influence maximisation* problem in the presence of negative edges, to study how negative ties impact *influence maximisation* efforts in the competitive setting against both known and unknown competitor allocations.

1.4   Summary of Results

We now briefly highlight some important results obtained in this thesis.

1.4.1   Continuous allocations of influence

We modify the traditional voter model to study competitive *influence maximisation* under continuous allocation of external influence, and subsequently employ it to characterise optimal strategies against known and unknown competitor allocations.

We first validate our proposed model by using it to study influence dynamics in a prototypical leader-follower-based star topology. We do this primarily to gather intuitions about the research problem from analytical results, that can later be used to design numerical methods in larger networks. In addition, we illustrate that it is always favourable to target the network continuously, as opposed to discretely activating a few nodes (Kuhlman et al. (2010); Masuda (2015)). Finally, we study the problem in a game-theoretic setting, where we show that the optimal strategy is to always target everyone in the network uniformly, with the same intensity.

1.4.2   Nonlinear effect of allocations

To generalise our model further, we introduce a new parameter that represents a fixed cost incurred by the controller to access nodes in a network. We first obtain analytical results in star networks, and show how analytically derived results can help design efficient heuristics to optimally maximise influence in larger networks. Moreover, we find that the cost of access parameter bridges the gap between the traditional *influence maximisation* models that allocate resources discretely[1], and our proposed continuous model of allocations. We further study the problem under other, more general, nonlinear cost functions, where we determine optimal patterns of allocations and further demonstrate the importance of considering these nonlinear constraints when maximising influence in networks.

1.4.3   Effect of negative ties

Finally, we propose a novel method to optimally maximise influence spread in networks with negative edges, and compare our approach to a traditional setup where all edges are assumed to be positive. We find that the knowledge of negative ties enhances influence spread in instances where the competitor deliberately avoids these edges. Conversely, our method yields no gains in certain network topologies, where the competitor targets all nodes uniformly. Finally, we explore the problem in a game-theoretic setting and show that knowledge of negative ties can, in some cases, adversely affect the outcomes of an intervention.

1.5   Publications

Some of the work in this thesis has led to published proceedings. Results discussed in Chapter 3 have been published in the following papers:

---

[1]The top k individuals with the highest number of social connections are targeted.

1. Chakraborty, S., Stein, S., Brede, M., Swami, A., DeMel, G., Restocchi, V. Competitive Influence Maximisation using Voting Dynamics (2019). In: The Social Influence Workshop at the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, Canada.

2. Romero Moreno, G., Chakraborty, S., Brede, M. Shadowing and shielding: Effective heuristics for continuous influence maximisation in the voting dynamics (2021). In: Plos one journal.

Below are some of our papers that summarise the results discussed in Chapter 4 and Chapter 5 respectively:

1. Chakraborty, S., Stein, S. Competitive influence maximisation with nonlinear costs of allocation. *(Accepted at the 11th International Conference on Computational Data and Social Networks 2022).*

2. Chakraborty, S., Stein, S., Brede, M., Swami, A. Competitive influence maximisation in the presence of negative ties. *(Working paper).*

1.6   Outline of the thesis

This chapter motivates the work presented in the rest of the thesis by providing a glimpse of existing gaps in the literature and outlining the research problem and its real-world impact.

In Chapter 2 we lay out the important theoretical concepts required to study the current research problem. The chapter provides a detailed account of many network science concepts and is aimed at readers who have limited expertise in the area. We first discuss network representations of social systems and describe methods used to model influence processes in social networks. We then present a review of the available literature on the competitive *influence maximisation* problem and highlight gaps that this thesis intends to fill.

In Chapter 3 we introduce the competitive *influence maximisation* problem with continuous allocations of external influence. Optimal counter-strategies against both known and unknown competitor allocations are derived using algorithmic approaches as well as analytical methods.

We then extend this model in two directions (as shown in Fig. 1.1). First, we change the constraints of the optimisation problem presented in Chapter 3 to include more realistic and complex settings. This work is presented in Chapter 4 where we study the competitive *influence maximisation* problem assuming a fixed cost of access and other
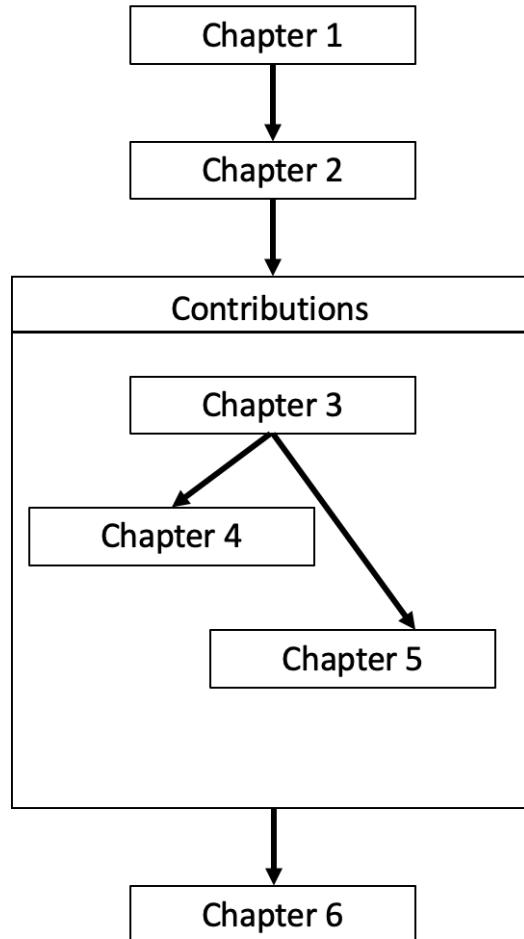
FIGURE 1.1: Figure showing the structure followed in this thesis.

nonlinear costs of allocation. Numerical approaches, and where possible, analytical methods are proposed to determine optimal allocations under these settings.

Second, we modify the model presented in Chapter 3 to study the competitive *influence maximisation* problem in networks with signed edges. More specifically, in Chapter 5 we study the impact of negative ties on *influence maximisation* efforts in competitive settings. We optimise strategies against known and unknown competitor allocations and highlight settings where knowledge of negative ties result in significant gain in influence spread, and similarly draw attention to instances where it yields no gains or results in a loss.

Chapter 6 summarises some important results obtained in Chapters 3 to 5, and highlights the outcome of the research problem studied in this thesis. Finally, it briefly outlines future directions in which the current work can be extended.

Chapter 2

Background

The study of network-based interventions requires careful analysis of real-world social dynamics. This however can be challenging due to the complex nature of human interactions, and the scale at which they occur (Barabási (2003)). Devising a tractable way to study real-world social systems thus requires an approach that can tame this complexity without losing any important properties of the system. A key contribution in this aspect was the notion of using networks to capture the structure and connectivity of complex social settings. This concept —that eventually evolved into the *network science* discipline —provides a level of abstraction to deal with the complexity of social systems, and yields a mathematical framework that can be ubiquitously applied to study any real-world setting (Barabasi (2019); Vespignani (2018)).

In the rest of the chapter, we first familiarise the reader with terminologies and concepts from the field of *network science* that are fundamental to the study of complex social networks. This is followed by a discussion on the building blocks of the research model. Studying the effect of network-based interventions on collective behaviours in a theoretical setting relies on capturing two important aspects of real-world social systems: (i) its connectivity, and (ii) the dynamics of social interactions. Consequently, we provide a comprehensive review of approaches taken to represent the structure and dynamics of real-world social systems.

We then proceed to describe the *influence maximisation* problem which is extensively used to investigate the optimisation of network-based interventions in populations and is the core theme of this research. The existing literature on *influence maximisation* is extensively broad, and we streamline our review of this field to the aspects of the research problem considered in this work. For a more comprehensive background of this field, see Li et al. (2018); Banerjee et al. (2020); Azaouzi et al. (2021). Finally, we highlight gaps in the existing literature that motivate our research aims presented in Chapter 1, and which have been addressed in the rest of the thesis.
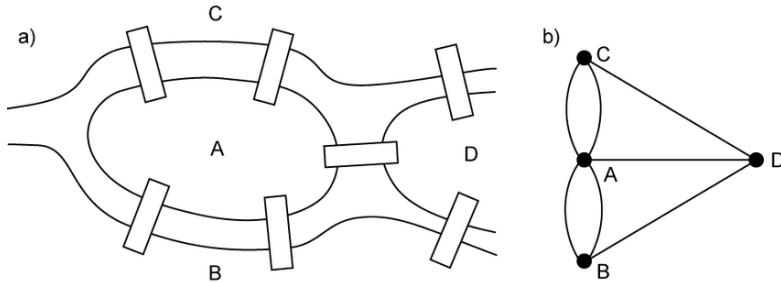
FIGURE 2.1: Figure (a) shows a schematic of the Könisberg bridge problem, and Figure (b) shows the graphical representation of the same problem (Figure taken from Boguslawski (2011)). The city of Könisberg was built on the islands A and D, surrounded by the Pregel river, and connected to the adjacent landmasses B and C by seven bridges. The aim of the Könisberg bridge problem was to determine if there exists a single path that traversed each bridge exactly once.

2.1   Network theory

Network science has its origin in *graph theory*. Graphs typically depict pairwise associations between objects in a set. In principal, a graph can be used to represent any collection of objects with connections between them, such as people and the relationships between them, websites and the links between them, or places and the transport links connecting them (Newman (2003)).

Using graphs to solve problems is a well-established method in *discrete mathematics*, and was initially proposed by the Swiss mathematician Leonhard Euler in 1736, while attempting to solve the landmark Könisberg bridge problem. The Könisberg problem shown in Fig. 2.1, is given by the following. The city of Könisberg was built on *two* islands situated on the Pregel river, connected to each other as well as to the adjacent river banks via *seven* bridges. The challenge of the Könisberg bridge problem was to find a single path that would traverse all *seven* bridges exactly once.

Using a graphical representation of the problem (see Fig. 2.1), Euler showed that for such a path to exist, there can be at most *two* landmasses with an odd number of bridges. These points in the graph would serve as entry or exit points for the path, and since all *four* landmasses in this instance had odd numbers of bridges, he conclusively showed that there was no such path that traversed all bridges exactly once. By representing the problem graphically, Euler devised a method of abstraction that could effectively remove all irrelevant details from a problem setting to simply focus on core of the problem —which in this case was the connectivity between the landmasses. This property of abstraction is what makes graphs such an attractive tool in the study complex systems —thus conceptualising the field of *network theory*.

2.1.1   Some *network theory* concepts

Before proceeding further, we introduce some common *network theory* concepts to ease the reader into terminologies and concepts that have been liberally used in the rest of the thesis.

1. **Vertex:** Each item in the set of objects (to be represented as a graph) forms a node or a vertex.

2. **Edge:** Relationships between items are depicted as connections or edges (between nodes).

3. **Sub-graph**: A smaller graph formed by the subset of the vertices (and the edges between them).

4. **Types of connections:** Often connections between nodes are of different types.

   (a) **Directed:** A directed edge represents a one-sided relationship, where an edge starts from a source node and ends at a target node. Hence, these edges can be traversed only in one direction (source $\longrightarrow$ target).

   (b) **Undirected:** These edges represent two-sided relationships such as friendships, and can be traversed in both directions.

   (c) **Weighted:** Weighted edges in a network indicate that not all connections are the same, and some connections are stronger than others (i.e. friends *versus* acquaintances).

   (d) **Unweighted:** Unweighted edges on the other hand assume all connections in the network have the same strength.

5. **Degree:** The number of connecting links (or edges) a node has (with other nodes in the network). In a directed graph there are *two* types of degrees:

   (a) **in-degree:** The number of incoming links of a node (i.e. number of connections where the node is the target).

   (b) **out-degree:** The number of outgoing links of a node (i.e. number of connections where the node is the source).

6. **Path:** The sequence of edges from one node to another.

7. **Clique:** A clique or a complete graph is one where all nodes are connected to each other.

8. **Centrality:** Centrality measure is a topological property of a node that indicates its importance in a network flow process[1]. There are different types of centrality

---

[1] For example, nodes that are important in the spread of information in a network.

measures e.g. degree, closeness[2] and so on. More such centrality measures can
be found in Barabási and Pósfai (2016).

9. **Clustering coefficient:** It is the ratio of existing connections between the
   neighbours of a node to all possible connections between them.

10. **Assortativity:** The tendency of nodes to connect to other nodes with similar
    properties. Similarly, disassortativity is when nodes connect to dissimilar nodes.

2.2   Representing real-world networks

Drawing meaningful and relevant insights about real-world social dynamics require
analysing real-world data, which may not always be easy to obtain. One way to
generalise theoretical results to real-world settings is to study the problem in synthetic
networks that aptly capture properties of real-world networks. In the remainder of
this section, we provide a comprehensive overview of broadly used network models
that generate synthetic networks used for the theoretical investigation of real-world
dynamics.

2.2.1   The small-world effect

As of 2017, the world is populated by over 7.6 billion people[3], and yet, we often claim
that it is a "small world". The phrase in general, hints at the overall connectedness in
the world, where often random strangers find themselves separated by a surprisingly
low number of intermediary relationships.

The intuition of short average path length in social networks was first proposed by
Frigyes Karinthy in 1929, and later illustrated through an experiment nearly 40 years
later by Stanley Milgram (Travers and Milgram (1967)). At the time, Milgram showed
that on average there are a maximum of six intermediary relationships between any
two people in the world. A phenomenon commonly known as the "Six degrees of
separation"[4], which given the advancement of technology in the last two decades, and
the internet has been revised to "Three-and-a-half degrees of separation" as per a
recent study on the Facebook network (Edunov et al. (2016)).

One of the first prominent network models to efficiently capture this small-world
property was the random graph model proposed by mathematicians Erdös and Rényi

---

[2]Indicates the average length of the shortest path between the node and all other nodes in the network
[3]Data obtained from the 2017 UN Report.
[4]implying that any *two* disconnected nodes in a social network have at most a path-length of *six* edges
between them —where path-length is a relational concept (i.e. represents social relationships) and not a
spatial one.

([Erdos and Rényi](1960)). To this day, the model is extensively used in *network science* research given its simplicity and relevance ([Newman](2000)). The random graph considers a population of N people, connected through $(N \cdot p/2)$ connections, where p[5] is the probability any two randomly chosen vertices have an edge between them. The degrees of separation in the resulting network is given by $\log N$, which therefore preserves the small-world property even when N is very large.

However, despite its advantages, the random graph model falls short in capturing two other commonly observed properties of real-world social networks. One of them is the clustering property[6]. The other is the heavy-tailed power-law degree distribution of real-world networks which is distinctly different from the Poisson-like degree distribution observed in random graphs ([Caldarelli](2007)).

2.2.2 The clustering effect

In 1998, nearly 40 years later, Watts and Strogatz proposed their own small-world model ([Watts and Strogatz](1998)). In the beginning, a lattice is considered where every node is connected to k of their nearest neighbours. Edges in the network are then rewired with a probability p[7]. The relevance of this model was justified in [Watts and Strogatz](1998) by highlighting that people usually form connections with those in their immediate surroundings (e.g. neighbours, colleagues and so on), while also occasionally forging ties with people outside of their social neighbourhoods (for instance, acquaintances formed through chance encounters). The lattice like structure of the network where people are connected to those in their immediate vicinity, results in an overlap in social neighbourhoods ('friends of friends') between neighbouring nodes, thus capturing the clustering effect observed in real-world social networks (that is missing in random graphs). The rewiring step on the other hand, randomly connects distant nodes that reduces the average path length in the network, thus also capturing the "small-world" effect.

Although the Watts-Strogatz small-world model is a useful model to study real-world networks, it has been criticised for its failure to capture the shape of degree distributions observed in the real world. A widely accepted property of real-world social networks as determined from empirical investigations is that they display a heavy-tailed power-law degree distribution ([Caldarelli](2007)) where a small number of nodes (called 'hubs') are densely connected (with many edges) within the network while the remaining nodes have sparse connections.

---

[5]Also termed as the coordination number of the network

[6]The tendency of forging friendships with the friends of a friend.

[7]The connection between two nodes is broken and a new connection is established with a randomly chosen node in the network.
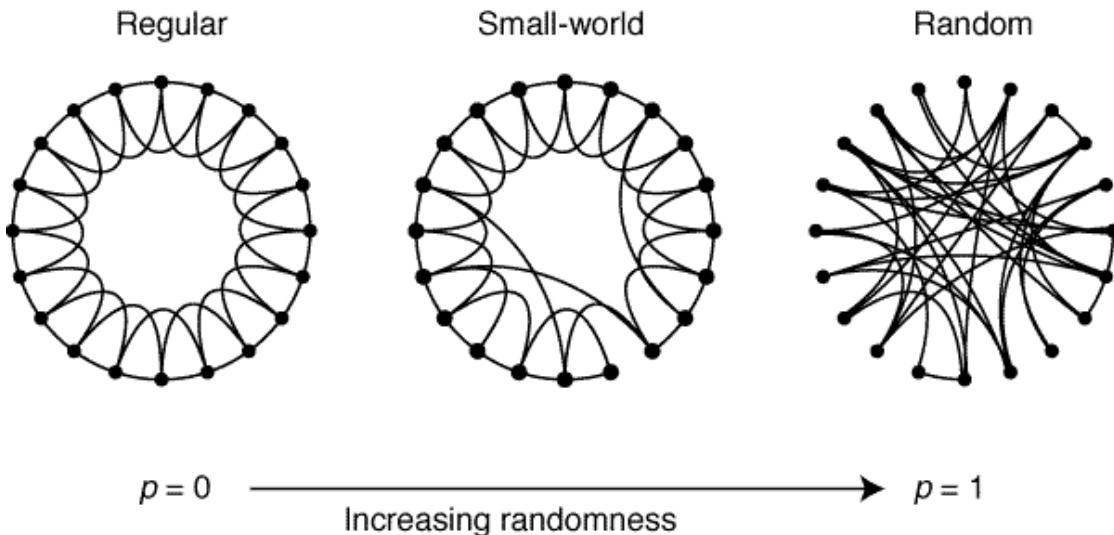
FIGURE 2.2: Figure from De Stefano et al. (2010), showing the transition from a regular
lattice, to the small-world network and finally to the random graph as the probability
of randomly rewiring links in the network p is increased.

2.2.3   Preferential attachment

In 1999, Barabási and Albert (1999) proposed a network model that effectively
captured the power-law degree distribution. Known as the preferential attachment
model, it was based on, what is known as the "rich gets richer" phenomenon. To
explain further, the preferential attachment model is a growth model that describes
how a network expands over time. As a consequence of growth in the network, older
nodes acquire more connections at the expense of newly added nodes, thus growing
"richer" over time.

A regular connected component of size m_0 nodes is considered. Newly added nodes
form connections with m existing nodes with a probability p that is directly
proportional to the degree of the existing nodes (see Fig. 2.3). Implying that new
nodes have a higher chance of connecting to older nodes that have formed more
connections over time. Thus enabling the "rich" nodes to get even "richer" as the
network expands.

The resulting network is known as a scale-free network that is widely observed in
several biological, financial and social settings (Barabási and Pósfai (2016)). The
degree-distribution in a scale-free network resembles a power-law distribution given
by $P(k) \propto k^{-\gamma}$, where $P(k)$ is the probability distribution of degrees in the network, k
is the degree of nodes in the network and γ is a constant. Empirical studies reveal that
γ is usually between $2 < \gamma < 3$ (Barabási et al. (1999); Caldarelli (2007)).

While the above models provide suitable templates to represent connectivity in the
real world, the structure of the network (known as the network topology) also
influences the dynamical processes in the network. For instance, existence of social
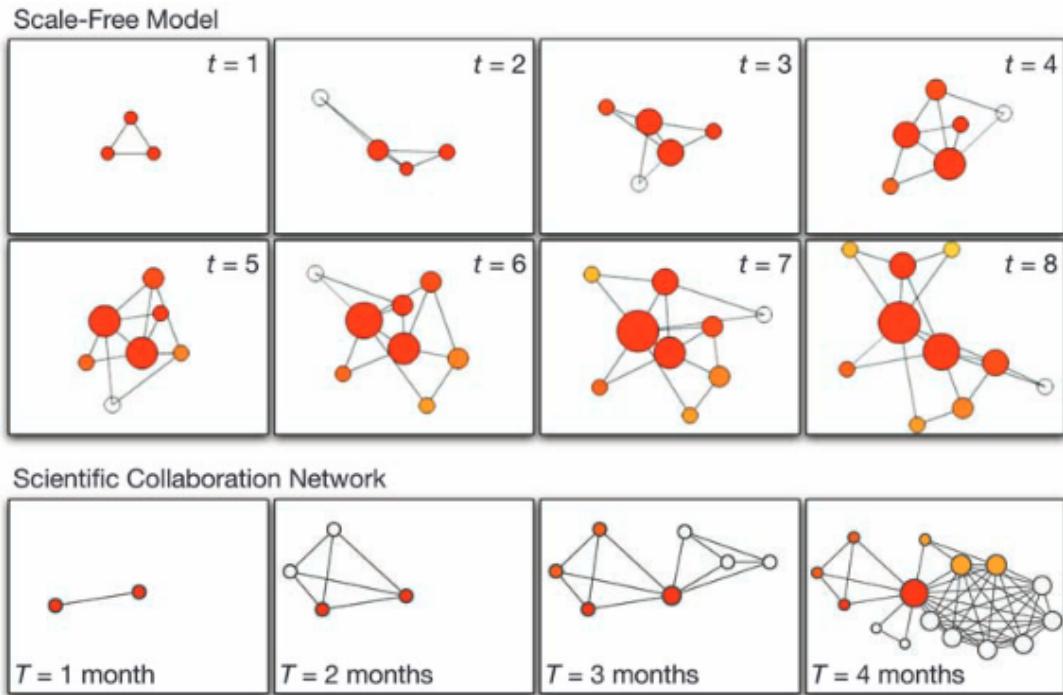
FIGURE 2.3: Figure from Barabási (2009) showing how the preferential attachment model works. The top panel illustrates the workings of the preferential attachment (or the scale-free) model over several time steps. Observe that the model begins with a connected component of $m_0 = 3$ nodes. As new nodes enter the system, they connect to existing nodes with a probability proportional to their degree. Thus, older nodes that were selected for new connections early on tend, to have more connections (or get richer) over time. The bottom panel shows how a collaboration network among physicists grows over time.

communities in highly clustered networks can often be detrimental to the spread process. As a next step, we determine models that effectively capture social processes observed in the real world, and use the following section to discuss approaches taken to effectively capture social interactions in a theoretical setting. Due to the nature of the research problem at hand, we focus our review strictly on methods that capture the propagation of influence in social networks.

2.3 Modelling influence propagation in social networks

Individual decisions and behaviours are often rooted in peer influence. Thus the key to understanding social behaviours —their formation and evolution —lies in the study of how influence flows in a population. Unfortunately, there is no one true model that can be universally applied to study the spread of influence in social settings. Instead, models vary based on the research setting being considered. For instance, infectious diseases propagate in a manner different to how culture spreads in a population. Hence models that study the spread of diseases starkly vary from

models that study the formation of languages and culture (Castellano et al. (2009a)). The challenge is then to choose an appropriate model that closely reflects the real-world setting being studied.

2.3.1   Social influence models

One of the earliest investigations of contagion[8] in social networks was presented by Moreno (1934), where he studied a running away epidemic among teenage girls at a reformation school in New York. Moreno captured the relationships between pupils in the form of sociograms[9], and used them to show that the decision to run away, typically reflected the position of the individual in the network and their relationships with others, as opposed to individual personalities or personal life events (as one would otherwise expect).

This study laid the cornerstone for research on social influence —specifically, its propagation and effects on individual and collective behaviours in populations. Some other formative works that pioneered this field include Latanè's social impact theory that described how individuals influence each other within a peer group (Latané (1981)). Schelling and Axelrod's models of cultural dynamics, and Granovetter's threshold model to study collective behaviours in society are prominently used to this day in sociological studies (Schelling (1971); Axelrod (1997); Granovetter (1978)). It is worth mentioning that none of the above works (in exception of Moreno (1934)) are based on network representations of populations. Instead, they use agent-based models (which is a field of research on its own) that simulate individual behaviours in a multi-agent setting to study macroscopic behaviours and outcomes (see Crooks and Heppenstall (2012); Janssen (2005) for more details). In this thesis however, we strictly focus on network-based social processes.

2.3.2   Diffusion models

One of the most established methods to study influence propagation in networks uses diffusion models. These models rely on word-of-mouth recommendations where only a few individuals initially adopt an innovation (or intervention), and subsequently propagate this adoption behaviour to the rest of the network. As information spreads from one person to another, it creates cascades of influence throughout the network (and are hence also known as cascade models), increasing the global share of adoption in the population.

Key examples of cascade models that largely dominate the field are the *Independent Cascade* model and the *Linear Threshold* model, both of which were introduced in the

---

[8]The infectious nature of peer influence.
[9]Graphical representation of relationships between members of a social group.

paper by Kempe et al. (2003). The *two* models differ in the way influence propagates in the network. Influence in the *Independent Cascade* model spreads by means of simple contagion where an infected node i activates its neighbouring node $j \in \{1 \ldots \mathcal{N}_i\}$[10] with a probability p_{ij} (through a single contact event). Whereas, nodes in the *Linear Threshold* model update their states only when the total influence in their neighbourhood breaches a certain threshold. This is known as complex contagion and is reflective of the way in which social norms evolve in a population.

In addition to the *Independent Cascade* model, simple contagion processes are also studied in epidemiological models such as the *SIR* (susceptible-infected-removed) model and the *SIS* (susceptible-infected-susceptible) model (Anderson and May (1992); Prakash et al. (2012); Watkins et al. (2016)). Although these models were primarily designed to capture processes of disease spread in populations, the mechanism followed in these models can also be applied to the study diffusion of influence in social settings.

Infection (or influence) spreads in the *SIR* model as infected individuals (I) transmit the disease to their susceptible neighbours (S) at a rate determined by the transmissibility of the disease (analogous to the probability of activation p_{ij} in the *Independent Cascade* model). These infected nodes (I) remain in the network for a duration T, also known as the infection period, during which they infect other susceptible individuals in their social neighbourhood. At the end of the infection period, infected nodes (I) are removed (R) from the system assuming death (or recovery[11]). The *SIS* follows a similar model of disease transmission, but applies to diseases where the infected individuals can become susceptible to re-infections at the end of the infection period (e.g. flu or common cold).

Determining the optimal set of nodes that maximises influence in a network, when considering diffusion models, is an NP-hard problem (Kempe et al. (2003)). However, the collective influence in a diffusion model is a monotonically increasing function with diminishing returns[12] —implying submodularity[13], which be leveraged to design polynomial-time, near-optimal greedy approaches that solve the otherwise NP-hard problem.

Over the years, many algorithmic approaches have been employed to determine the optimal solution to the *influence maximisation* problem in diffusion models, such as: (i) the *simulation-based* approach, (ii) the *proxy-based* approach and (iii) the *sketch-based* approach. The *simulation-based* approach employs *Monte-Carlo (MC)* simulations to determine the optimal seed set, where every possible combination of seeds are

---

[10]Here $\mathcal{N}_i$ is the neighbourhood of node i.

[11]Once the individual is recovered they can never contract the disease again, e.g. measles

[12]This is because nodes in the network can only be activated and never deactivated.

[13]Submodularity states that the marginal value gained from adding an element e to a set S is less than the marginal gain obtained when e is added to a subset of S (Krause and Golovin (2014); Lovász (1983); Fujishige (2005)).

evaluated to determine the set that maximises influence in the network. Although, it provides theoretical convergence under specific settings, and its simple approach can be easily incorporated into any diffusion model —this method can rapidly escalate into an expensive computational task in large networks, even for a reasonable tolerance of error. The *proxy-based* approach on the other hand, addresses the computational inefficiency in the *simulation-based* approach by reducing complex influence models into standard proxy models such as Page-Rank. However, despite its benefits, this method fails to offer any theoretical guarantees for the optimal solution, and the optimal seed set changes significantly in some cases even for a minor change in the underlying network. The *sketch-based* approach is widely known for its theoretical efficiency as well as low time-complexity. The approach uses *Monte-Carlo* simulations to determine the optimal seed set, but significantly reduces the computation time by using an oracle-based system (that pre-computes the influence path of each node). Examples of this algorithm include TIM/TIM+ (Tang et al. (2014)), IMM (Tang et al. (2015)), StaticGreedy (Cheng et al. (2013)) etc.

However, despite their relevance and popularity, the static nature of individual states in diffusion models (as discussed in Chapter 1) is a limitation when studying dynamic social behaviours. Although, epidemiological models such as the *SIR* model and the *SIS* model allow node states to be flexible to a certain extent, and are occasionally used to study dynamical social systems (Romero et al. (2009); Stauffer and Sahimi (2006); Woo et al. (2011)), they are not suited for exploring instances where individual states change repeatedly and stochastically.

2.3.3   Dynamic models

Dynamical approaches are effective when studying emerging macroscopic behaviours (e.g. majority opinion) in populations —where individuals repeatedly switch between different opinion states (Barrat et al. (2008)). Here nodes update their states each time they interact with their neighbours. The process of updating the state of a node can be synchronous, where all nodes update their states simultaneously, or can be asynchronous, where nodes in the network update their states in a particular order. Although the stochasticity of the system can appear intimidating to study, a way forward is often to focus on macroscopic behaviours instead of microscopic interactions in the population (Castellano et al. (2009a); Sen and Chakrabarti (2014)), and thus the system is studied when it reaches a steady-state, i.e. global consensus, or the fraction of opinion states in the network remain static, and no longer change with time.

A conventional approach to determine collective behaviour in such dynamic systems uses an assembly of models from sociophysics (Castellano et al. (2009a); Sen and Chakrabarti (2014)), where individual states are expressed using either discrete

variables (Clifford and Sudbury (1973); Galam (1999); Holley and Liggett (1975); Krapivsky and Redner (2003)) or continuous (Deffuant et al. (2000); Hegselmann et al. (2002)) variables. One of the most prominent dynamic models, widely used to study opinion dynamics, is the voter model, that assumes discrete opinion states for individuals in the population(Clifford and Sudbury (1973); Holley and Liggett (1975)). Despite the abundance of dynamic models used to study social processes, the voter model stands out for its simplicity and fair representation of real-world social dynamics (Castellano et al. (2009b); Redner (2019)). Individuals in this model update their states through imitation behaviour where a node is first picked uniformly at random from the population to update its state. The selected node then randomly picks one of its neighbours, and copies their state. This process continues till the entire population reaches an equilibrium state (or steady-state).

Other variations of the voter model have been also proposed. In the reverse voter model a node is randomly chosen from the network and its state is copied to one of its randomly selected neighbours (Castellano (2005)). Other variants include the constrained voter model, where nodes are assumed to be either centrists or extremists (leftists, or rightists). Influence propagates in the system as centrists communicate with extremists, and other centrists. The extremists (leftists and rightists) on the other hand do not engage with one another as they are on the opposite ends of the opinion spectrum (Vazquez et al. (2003)). In the noisy voter model, individuals are considered to spontaneously flip from one state to another at a rate given by q such that when q is set to *zero*, we retain the classical voter model (Granovsky and Madras (1995)).

Other dynamic models include the majority rule model where individual states are given by discrete variables. This model was proposed to illustrate influence propagation through group discussions and public debates, where individuals —at the time of updating their states —interact with a fraction of their neighbourhood, and then adopt the most prevalent opinion state within the group (or randomly in case of ties Galam (2002)). Other models that assume discrete individual states include the Naming Game which is primarily used to study language dynamics (Steels (1995)), and the Sznajd model, where influence propagates through social validation (Sznajd-Weron (2005)). Alternatively, individual states (or opinions) can be continuous in nature where individuals rate their preference towards an opinion over a defined range (e.g. 0 to 1), which starkly contrasts discrete opinions (e.g. binary opinions given by ± 1) Dynamic models which consider continuous opinions are known as bounded confidence models, where people strictly interact with those who have opinions similar to them. Examples include the Hegselmann-Krause model (Hegselmann et al. (2002)) and the Deffuant model (Deffuant et al. (2000)). For a more comprehensive review on dynamic models see Castellano et al. (2009a).

2.4   *Influence maximisation* **in competitive settings**

The *influence maximisation* problem has been extensively studied under competitive settings, especially given its commercial appeal. It has been extensively explored in diffusion models, as extensions to the *Independent Cascade* model (Bharathi et al. (2007); Carnes et al. (2007)), the *Linear Threshold* model (Borodin et al. (2010); He et al. (2012)) and other *two*-step diffusion models as well (Goyal et al. (2014)). It has also been well-studied in dynamic models, particularly in the voter model using zealots (Mobilia (2003)). Generally, zealots are biased individuals (Masuda et al. (2010); Mobilia (2003)), or radical agents in the network who are impervious to any neighbourly influence (Barrat et al. (2008); Kuhlman et al. (2013); Mobilia et al. (2007); Mobilia (2015)), and are responsible for preventing the population from reaching a consensus (Acemoğlu et al. (2013); Mobilia et al. (2007); Yildiz et al. (2013)). Thus in a competitive setting, zealots are used as "influence blockers" that limit the spread of any opposing influence in network (Kuhlman et al. (2010); Yildiz et al. (2013)).

In a more recent approach, rooted in *control theory*, zealots were used as agents external to the network, that unidirectionally targeted selected individuals in the network (Masuda (2015)). By controlling an optimal set of nodes in the network, the external controller maximised their control (or influence) over the rest of the network. Masuda (2015) showed that under competitive settings, the optimal strategy to achieve maximum influence spread in network is to preferentially target hub nodes. However, this was a generic result was further honed by studying the problem in noisy and dynamic settings (Brede et al. (2018, 2019)). Brede et al. (2018) showed that while targeting hub nodes is optimal for maintaining long-term control over the network, targeting low-degree peripheral nodes is more effective for short-term *influence maximisation*. In addition, Brede et al. (2019) studied the problem in a noisy voter model and showed that optimal control shifts from high-degree hub nodes to low-degree peripheral nodes as noise in the model increases.

We expand this line of investigation in Chapter 3 and extensively compare the roles of high-degree and low-degree nodes in the *influence maximisation* process —by continuously allocating resources (or external influence) to nodes in the network. As nodes are targeted with varying intensities in this method, it is used to determine settings where high-degree nodes or low-degree nodes should be preferentially targeted to maximise influence spread.

Additionally, the *influence maximisation* problem is typically a constrained optimisation problem where resources allocated to the network are constrained by a budget[14]. In traditional (discrete) methods used to allocate resources to a network, the budget constrains the number of nodes converted (or "seeded") at the start of the dynamics.

---

[14]In comparison, the unbudgeted case is a trivial problem.

Thus in the continuous approach, as resources are heterogeneously spread over the network and are constrained by an overall budget, we need to define a relationship between the amount of resources allocated to a node and the influence experienced by them. As a preliminary step, we consider this relation to be linear in Chapter 3. In Chapter 4 however, we take a step further in this direction by considering other nonlinear relations between allocations and influence experienced by nodes. Influence dynamics in the classical voter model is uniquely linear, and nonlinearity has been studied in voter dynamics typically in terms of spread dynamics —specifically when considering majority dynamics where individuals consult and respond to the majority view in their neighbourhood (Peralta et al. (2018); Schweitzer and Behera (2015)). For example, we obtain nonlinear spread dynamics in the *q*-voter model where individuals either adopt the unanimous view of *q* neighbours or spontaneously switch their opinions with a probability ϵ (Castellano et al. (2009b)), or in models with contrarians where some individuals (or contrarians) consistently oppose the majority view in their neighbourhood (Tanabe and Masuda (2013)). Nonlinearity however, to the best of our knowledge has never been considered in the voter model in terms of allocations to the network and our work is the first to do so.

Finally in Chapter 5 we study competitive *influence maximisation* in networks with negative ties —which has received limited attention in the past (see Girdhar and Bharadwaj (2016) for a detailed review). Many of the works in this area explore the problem in diffusion models —such as the *Independent Cascade* model (Ju et al. (2020); Li et al. (2014); Liu et al. (2019)) and the *Linear Threshold* model (He et al. (2019); Liang et al. (2019); Shen et al. (2015)) —where traditional greedy heuristics are principally employed to determine optimal solutions to the problem. Some other contributions to this area include using a modified integrated page-rank algorithm (Chen and He (2015)), or simulated annealing (Li et al. (2017)) to determine the optimal seed set in a diffusion model, while Srivastava et al. (2015) discusses a comparative analytical approach to the problem. The problem is studied in the voter model in Li et al. (2013), and possibly bears the closest resemblance to our work. However, our research extends this work in several aspects. First as foremost, we consider continuous allocation of resources, which distinctly contrasts the traditional discrete method employed in Li et al. (2013) and in general, in this research area. Additionally, we study the problem in a competitive setting which is starkly different from the single-controller setting studied in Li et al. (2013).

Chapter 3

Continuous allocation of resources

3.1   Introduction

The research problem central to this thesis (as described in Chapter 1), is to determine ways in which the spread of desirable influence can be maximised in populations through the optimal distribution of limited resources in the presence of competition. To do this, we employ the paradigmatic voter model (Clifford and Sudbury (1973); Holley and Liggett (1975); Sood and Redner (2005)), for which our motivations are highlighted in Section 1.2. The voter model also closely represents real-world dynamics (Braha and de Aguiar (2017); Redner (2019)), and the simplicity and tractability of its approach enables analytical examination of influence dynamics in complex networks, which is rare and useful when drawing insights about aggregate behaviours in populations (Castellano et al. (2009a)). As discussed in Section 2.4, despite its rich background most of the historic work using the voter model is characterised by discrete allocation of external influence on the network —where nodes are either influenced or not. In contrast, here we consider a continuous approach where a broad spectrum of nodes are targeted with varying amounts of influence. Naturally, the resources allocated to the network must also be continuous in nature, like time and money, which are also commonly used to influence networks in the real world, thus making our model more realistic in comparison to traditional models.

As stated earlier, we study the problem in a competitive setting, where *two* controllers compete to maximise their influence in a population. Over the years, *competitive influence maximisation* has attracted considerable attention, given its commercial value in many real-world settings (Bharathi et al. (2007); Wilder and Vorobeychik (2018)). Here we consider two scenarios: (i) where we have prior knowledge of competitor allocations in the network, and (ii) where we have incomplete knowledge of competitor allocations. Where competitor allocations are not fixed, game-theory

provides a viable approach to determine optimal allocations for both controllers. While cascade models have been abundantly used to study *influence maximisation* in game-theoretic settings (Clark and Poovendran (2011)), dynamical models have received limited attention in this aspect. Some exceptions include Chasparis and Shamma (2010); Fazeli et al. (2016); Masucci and Silva (2014), of which we find Masucci and Silva (2014) to bear the closest resemblance to our work. In Masucci and Silva (2014) however, networks are targeted using the single-injection approach of cascade models where a set of individuals are flipped (or infected) at the start of the dynamics. This is starkly different from our approach where instead of fixing the state of certain nodes, influence is applied externally, and continually to the network.

In this chapter, we first analyse the problem in star networks. Our reasons for this are three-fold. First, it is an archetypal example of leader-follower structures in social networks. The bimodal degree distribution in star networks creates a *trade-off* between the roles of *low-degree follower* nodes and *high degree leader* nodes in the influence spread process which we can exploit, to compare the traditional discrete method (focused on the hub (Masuda (2015); Kuhlman et al. (2013))) and the proposed continuous method of allocating influence. Second, centralised structures such as the star graph are noted for their efficiency in information propagation and provides a suitable template to test our model (Leavitt (1951)). Finally, star graphs are largely prevalent in online social networks (e.g., Twitter (Rathnayake and Suthers (2016))) and organisational networks (Tichy et al. (1979)), which makes our results relevant in the real world.

3.2   Outline

The chapter is structured as follows:

In Section 3.3 we introduce the competitive influence maximisation model with continuous allocations for voter dynamics.

We then examine patterns of optimal allocations against a passive competitor in Section 3.4. We consider several instances of fixed competitor allocations in a star network, and present closed-form analytical solutions in each case. We also compare the optimal strategy to other heuristics that are commonly used to maximise influence social networks. We then show how much vote-share a controller can gain by optimally allocating resources.

Finally, in Section 3.5, we study the instance where both controllers simultaneously target the network. We obtain optimal allocations for both controllers first in the star network using analytical methods, and then in a real-world network using numerical methods.

In Section 3.6 we summarise and revisit some of our results. We discuss some limitations of our current setting and discuss potential future directions to extend the work.

3.3   The opinion dynamics model

We consider a population of N individuals, connected through a social network. The structure of the network is given by a graph $G(V, E)$, where each vertex $i \in V = \{1, 2, \ldots, N\}$ represents an individual in the population, connected to their immediate social neighbourhood $\{j \in V; j \neq i\}$ through a subset of E[1]. Edges have weights associated with them. These weights w_{ij} quantify the strength of a connection between any individual i and their neighbour j, and in turn ascertains the intensity with which i influences j and vice-versa. We assume the weight of every edge is pre-determined, and is captured in a non-negative weighted-adjacency matrix $W \in \mathbb{R}_+^{N \times N}$.

We explore a setting where two controllers (A and B) compete to maximise their influence (or opinions) in the population. At any given point in time, individuals in the network strictly adhere to one of two opinions (A or B), corresponding to each controller. Opinions are characterised using binary state variables $\sigma_{A,i}(t) \in \{0, 1\}$, where $\sigma_{A,i}(t) = 1$ implies node i is in state A, or in state B ($\sigma_{A,i}(t) = 0$) at time t. From here, it is easy to follow that $\sigma_{B,i}(t) = 1 - \sigma_{A,i}(t)$.

Controllers maximise opinion shares in the population by influencing the network externally (shown in Fig. 3.1). For any node i, external influence from controllers A and B is quantified as $p_{A,i}$ and $p_{B,i}$ respectively, and controller allocations over the whole network is described using vectors p_A and p_B. Allocations are non-negative, $p_A \in \mathbb{R}_+^N$ and $p_B \in \mathbb{R}_+^N$ and linearly constrained by the budget B_A and B_B available to each controller, as $\sum_i p_{A,i} = B_A$ and $\sum_i p_{A,i} = B_B$.

Nodes update their opinions using voter dynamics (Holley and Liggett (1975)), at every time step, a node is selected uniformly at random to update their opinion state where they copy the state of a neighbouring node j with the probability $w_{ji} / (\sum_{j \in \mathcal{N}_i} w_{ji} + p_{A,i} + p_{B,i})$ where $\mathcal{N}_i$ is the immediate social neighbourhood of i, or copy the state of an external controller (say A) with the probability $p_{A,i} / (\sum_{j \in \mathcal{N}} w_{ji} + p_{A,i} + p_{B,i})$.

As opinions are stochastic, we approximate the global behaviour in the system by assuming that $x_{A,i}$ is the probability a node i is in state $\sigma_{A,i} = 1$, which gives us the

---

[1]We ignore self-loops in our model. This is because we assume that individuals have limited knowledge about the propagated opinions and rely purely on social interactions to choose their state of opinion. However, it is worth noting that self-loops can be easily added in the model without any fundamental changes to the underlying mathematics, as shown in Masuda (2015).
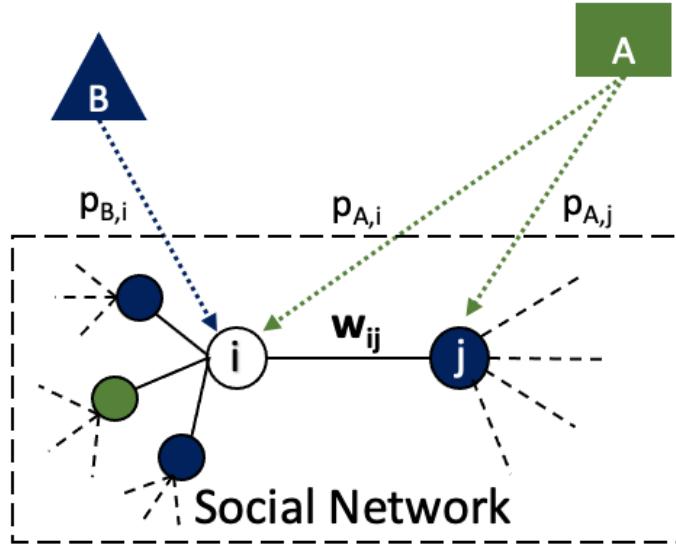
FIGURE 3.1: Schematic of the voter model with external controllers. Nodes (circles) represent individuals in the network. The network here is unweighted and undirected ($w_{ij} = 1; \forall \{i,j\} \in V; i \neq j$), i.e. all nodes influence each other equally. External controllers A (square) and B (triangle) unidirectionally influence the network with resources $p_{A,i}$ and $p_{B,i}$ ($\forall i \in V$). Note that i experiences external influence from both controllers, whereas j is targeted only by controller A. Filled green circles correspond to nodes in state A while filled blue circles represent those in state B. Say node i is randomly picked to update its state (shown by the void circle), and controllers influence the network with unit resources $p_{A,i} = p_{B,i} = 1$, ($\forall i \in V$). Observe that 1/4 of its social neighbourhood is in state A, while the rest is in state B. Therefore, node i picks state A or B with probabilities 1/3 and 2/3 respectively.

rate at which it chooses to remain in opinion state A as,

$$\frac{dx_{A,i}}{dt} = (1 - x_{A,i})\frac{\sum_j w_{ji} x_{A,j} + p_{A,i}}{\sum_j w_{ji} + p_{A,i} + p_{B,i}} - x_{A,i}\frac{\sum_j w_{ji}(1 - x_{A,j}) + p_{B,i}}{\sum_j w_{ji} + p_{A,i} + p_{B,i}}. \tag{3.1}$$

Here the terms $\frac{\sum_j w_{ji} x_{A,j} + p_{A,i}}{\sum_j w_{ji} + p_{A,i} + p_{B,i}}$ and $\frac{\sum_j w_{ji}(1 - x_{A,j}) + p_{B,i}}{\sum_j w_{ji} + p_{A,i} + p_{B,i}}$ quantify the total influence a node i experiences from their immediate neighbourhood and from external controllers in favour of opinions A and B respectively.

We estimate the global behaviour of the population by estimating the total share of opinions[2] obtained by each controller at equilibrium. We determine steady-state conditions by setting $\frac{dx_{A,i}}{dt} = 0$ in Eq. (3.1), which for an arbitrary network of size N

---

[2]Often referred to as vote-shares in the voter-model.

yields

$$\left[L + diag(p_A + p_B) \right] x_A = p_A,$$

where L is the Laplacian of the network given by a $N \times N$ matrix with diagonal elements representing the total strength of all edges on a node[3] ($L_{ii} = \sum_j w_{ji}$) and off-diagonal elements are $L_{ij} = -w_{ij}$. The *diag* function implies element-wise addition of allocation vectors p_A and p_B to the diagonal of the Laplacian matrix L, such that the i-th diagonal element of $[L + diag(p_A + p_B)]$ is given by $L_{ii} + p_{A,i} + p_{B,i}$.

The total vote-share obtained by controller A at equilibrium is then given by

$$\implies X_A = \frac{1}{N} \vec{1}^T x_A = \frac{1}{N} \vec{1}^T [L + diag(p_A + p_B)]^{-1} p_A. \tag{3.2}$$

Here $\vec{1}^T$ is a column-vector with ones in all its positions. The optimisation problem therefore can be stated as,

$$p_A^* = \arg\max_{p_A \in \mathcal{P}} X_A^*(L, p_B), \tag{3.3}$$

where $\mathcal{P}$ is a set of all possible allocations p_A such that $0 \leq p_{A,i} \leq B_A$ and $\sum_{i=1}^N p_{A,i} = B_A$.

Note that, here we have chosen A as our focal controller. Similar expressions can also be derived for controller B.

3.4 Maximising influence against fixed controller allocations

We begin by examining the optimisation problem in a star graph. Star-like structures are prevalent in many real-world social networks (Rathnayake and Suthers (2016); Tichy et al. (1979)), and are an archetypal example of leader-follower networks, where a leader (hub node) is connected to several followers (peripheral nodes), who do not have any connections between themselves (as shown in Fig. 3.2). The bimodal degree distribution facilitates a comparative analysis between the role of the hub and the periphery in the *influence maximisation* process, and in doing so, allows us to compare between discrete allocations (that focus exclusively on the hub), and more flexible continuous allocations (over several nodes).

In this section, we first explore the problem in a traditional setting where the competitor is passive (Bharathi et al. (2007); Masuda (2015)), and their allocations to

---

[3]Given by the node degree in unweighted graphs.
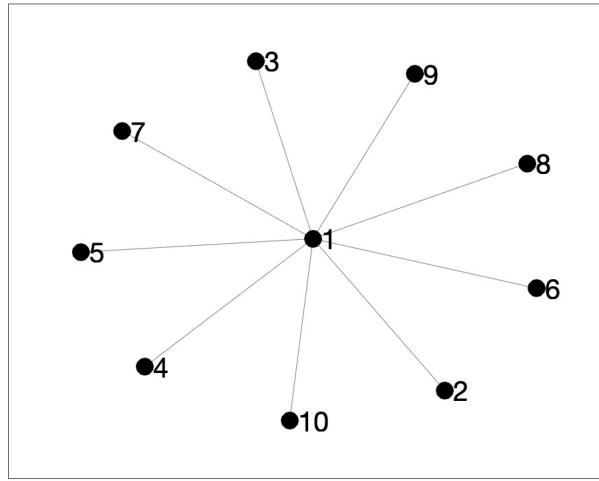
FIGURE 3.2: Figure showing a star graph of size $N = 10$. Node $i = 1$ is the hub node with degree 9. All other nodes ($2 \leq i \leq N$) with degree 1 are termed peripheral nodes.

the network are fixed. Later in the section, we extend the model to cope with more realistic settings, where competitor allocations are not disclosed in advance, and both controllers simultaneously target the network.

In the rest of the chapter, unless otherwise stated, we obtain our results in star networks of size $N = 1000$ and $n = (N - 1)$ peripheral nodes. As we are motivated to understand how topological properties (i.e. degree centralities) drive *influence maximisation* decisions, we restrict our analysis to unweighted and undirected graphs, where binary weights are used to capture the structure of the network, such that $w_{ij} = 1$, if i and j have an edge between them, else $w_{ij} = 0$. Additionally, *W* here is symmetric as we consider undirected graphs.

We assume controllers A and B have budgets B_A and B_B assigned to them which they use to influence the network. For ease of exposition, we use $B_A = a$ and $B_B = b$ in our closed-form analytical expressions. Competitor B here is passive, i.e. their allocations to the network are fixed and known. Inspired by the unique structure of the star graph, here we examine three natural instances of competitor allocations: (i) competitor targets the network uniformly, (ii) competitor targets the hub and (iii) competitor targets the periphery.

3.4.1   Competitor targets the network uniformly

We first consider a setting where the competitor targets the network uniformly. This is analogous to a generic form of marketing, where companies offer indiscriminate and uniform discounts on products. A different interpretation of this setting would be to

think of it as existing resistance in the network to a new idea or opinion, often caused by factors that uniformly apply to a population, such as socio-economic status (Devasenapathy et al. (2016); Mobarak and Saldanha (2022); Rogers (2010)).

We now proceed to define the setting mathematically. The allocation vector for the competitor is given by $p_{B,i} = b/(n+1)$, $\forall i \in \{1, 2, \ldots, n+1\}$. We parameterise the allocation vector for controller A as $p_A = (\alpha, k_A)$, where α is used to regulate the fraction of the total budget allocated to the peripheral nodes, and k_A controls the number of targeted peripheral nodes. Thus each peripheral node receives (α/k_A) resources and the hub is targeted with the residual budget $(a - \alpha)$. Note that as all peripheral nodes are identical in nature in an unweighted, undirected star graph, there is no need to differentiate between them.

We use Eq. (3.2) to obtain an expression for the total vote-share $X_A(\alpha, k_A)$. Here parameterising the allocation vector significantly reduces the degrees of freedom in the system and makes the analytical set-up easier to handle. We can now obtain closed-form analytical solutions for optimal allocations (α^*, k_A^*) by solving the partial derivatives $\nabla_{p_A} X_A = \left(\frac{\partial X_A}{\partial \alpha}, \frac{\partial X_A}{\partial k_A}\right) = 0$.

First, we observe[4] that $\nabla_{k_A} X_A(k_A) \geq 0$. This suggests that vote-share monotonically increases with k_A. The optimal k_A^* therefore lies on the boundary $k_A^* = n$, implying that the best response, in the current setting, is to always target all peripheral nodes. We can now replace $k_A = n$ in the earlier vote-share expression and obtain the partial derivative $\nabla_\alpha X_A(\alpha)$. Solving $\nabla_\alpha X_A(\alpha) = 0$ gives us two stationary solutions,

$$\alpha^* = \left\{ \frac{an}{n+1}, \frac{n(a(n+1) + 2b + 2(n+1)^2)}{(n+1)(n-1)} \right\}.$$

By rewriting the second expression as,

$$\alpha^* = \frac{n(a(n+1) + 2b + 2(n+1)^2)}{(n+1)(n-1)} = \frac{an}{n-1} + \frac{2bn}{(n+1)(n-1)} + \frac{2n(n+1)}{(n-1)},$$

we can easily show that it violates the budget constraint a. Therefore, our optimal solution for α in this instance is given by

$$\alpha^* = \frac{an}{n+1}. \tag{3.4}$$

Observe that the optimal strategy mirrors competitor allocations, and distributes resources uniformly over the network independent of budget availability (a/b). This also implies that continuous allocation of resources are optimal when the competitor employs a continuous approach to allocate resources over a network.

---

[4]Details of the derivations can be found in Appendix A.1.3

We now explore if optimal allocations vary when the competitor preferentially targets the hub or the periphery.

3.4.2   Competitor targets the hub

We consider an alternative allocation strategy for competitor B. In this setting, B focuses solely on the hub node and allocates the entire budget B_B to the hub. This is a well-established strategy in marketing, where firms commonly rely on "opinion leaders" (or hub nodes) to maximise the sale of their products in a population (Haenlein and Libai (2013)). Opinion leaders in social networks resemble hub nodes with high measures of centrality (e.g. degree, betweenness), such as public figures or celebrities who are often appointed as brand ambassadors for many companies (Pringle and Binet (2005)).

Under these settings, the hub node $i = 1$ receives the entire budget available to the competitor $p_{B,1} = B_B = b$, and none of the peripheral nodes receive any influence $p_{B,i} = 0$, $(2 \leq i \leq N)$. Taking the same approach as before[5], we determine the partial derivatives $(\nabla_\alpha X_A, \nabla_{k_A} X_A)$ and find that $\nabla_{k_A} X_A \geq 0$. Once again we replace $k_A = n$ in X_A and proceed to evaluate α^* from $\nabla_\alpha X_A = 0$. The optimal allocation is obtained as $\alpha^* = \{(\pm\sqrt{(n+1)^2 + (a+b)} - (n+1))n\}$. Here the solution with the negative root violates the non-negativity constraint on p_A, and therefore we only accept

$$\alpha^* = n(\sqrt{(n+1)^2 + (a+b)} - (n+1)), \tag{3.5}$$

as the solution for optimal allocations, subject to $0 < \alpha^* \leq a$.

We make the following observations from this result. First, we find that continuous allocation of resources is preferred over discrete allocations even when the competitor follows a discrete approach. Second, in contrast to the setting where the competitor targets the network uniformly, here we observe a correlation between optimal allocation α^* and controller budgets a and b.

As a first step, we examine optimal allocations in the limit of large budgets. For large controller budget $(a >> b)$, allocations to the hub $(1 - (\alpha^*/a)) \to 1$, whereas for large competitor budget $b >> a$, we find that $(1 - (\alpha^*/a)) \to 0$. This implies that the optimal strategy will regulate between discrete and continuous allocations as the amount of resources available to the controller varies in comparison to the competitor budget.

To obtain a richer understanding of how optimal allocations vary with controller budgets, we plot allocations to the hub as a fraction of the total budget $(1 - (\alpha^*/a))$ in

---

[5]Details of the derivations can be found in Appendix A.1.1
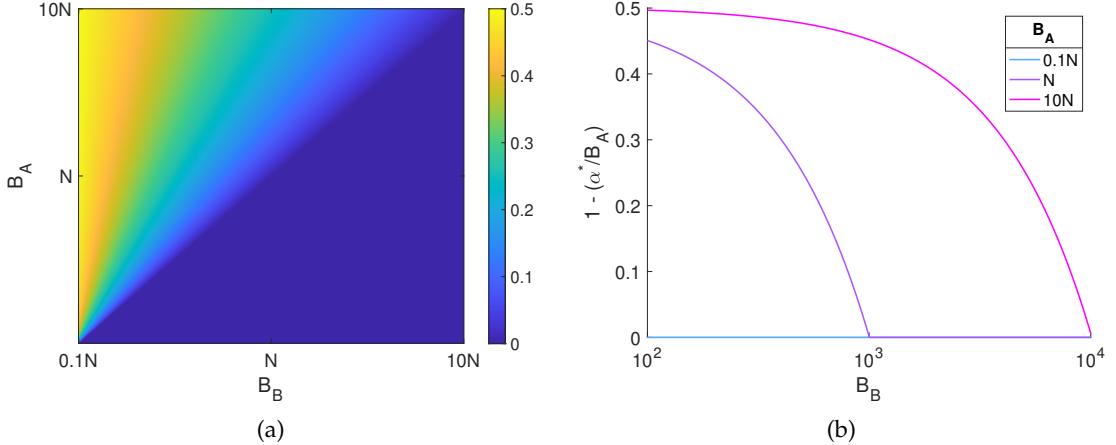
FIGURE 3.3: Figures showing optimal allocations to the hub node in a star network of size $N = 1000$. Competitor B discretely targets the hub with $B_B = b$. Figure (a) shows allocations to the hub node $(1 - \alpha^*/B_A)$ as a fraction of the budget, when controller budgets B_A and B_B are varied. Figure (b) shows allocations to the hub node for specific instances of controller budget $B_A \in \{0.1N, N, 10N\}$ in a semi-log plot as competitor budget is varied.

Fig. 3.3a, for budgets $B_A = a$ and $B_B = b$ where $0.1N \leq B_A \leq 10N$ and $0.1N \leq B_B \leq 10N$. Observe that allocations to the hub node are positive only when the budget available to the controller is equal to, or more than the competitor budget, $B_A \geq B_B$. This implies that the controller competes over the targeted node (here, the hub node) only when they have sufficient resources. For all other cases where $B_A < B_B$, the controller avoids the hub node completely. We show this more clearly in Fig. 3.3b, where we plot optimal allocations for specific values of a against varying competitor budget b. Observe how allocations to the hub node reduces to 0, when the competitor budget reaches a critical point b_c. We obtain an analytical expression for b_c by setting $\alpha^* = a$ in Eq. (3.5) and solving for b,

$$b_c = \frac{a(n(n+2) + a)}{n^2}. \tag{3.6}$$

Similarly, to determine conditions under which the hub receives the entire budget a, we set $\alpha^* = 0$ in Eq. (3.5), and obtain $b = -a$, which is not feasible. We can therefore conclude that irrespective of the budget ratios, the periphery is always targeted in this setting.

We now proceed to explore the instance where the competitor favourably targets the periphery.

3.4.3   Competitor targets the periphery

The *influence maximisation* literature commonly advocates the importance of hub nodes, or *leaders* in the *influence maximisation* process (Masuda (2015)). However, more recently, it has been argued that in organisational settings, *followers* play an equally important role in the influence spread process, and can often be responsible for influencing the behaviour of the *leaders* in the network (Lapierre and Bremner (2010)).

With this in mind, we define the setting where the competitor exclusively targets the periphery. The allocation to the hub node here is $p_{B,1} = 0$, while all other peripheral nodes receive $p_{B,i} = b/n, 2 \leq i \leq N$. Using the same approach as before[6], we obtain an expression for the total vote-share $X_A(\alpha, k_A)$ and find that $\nabla_{k_A} X_A \geq 0$. Thus we set $k_A = n$ and solve $\nabla_\alpha X_A = 0$ to obtain the solution for optimal allocations as $\alpha^* = \{(a + n + 1) \pm \sqrt{(n+1)^2 + (a+b)}\}$. Here we discard the solution with the positive root as it violates the budget constraint. The final solution for optimal allocation is therefore given by

$$\alpha^* = (a + n + 1) - \sqrt{(n+1)^2 + (a+b)}, \tag{3.7}$$

subject to $0 \leq \alpha^* \leq a$.

Once again, we find that optimal allocations are determined by the controller budgets a and b. Observe that in the limit of large competitor budget $b >> a$, we get $\alpha^* \to 0$. Implying that resources should be diverted away from the periphery, under low budget conditions. Similarly, for $a >> b$, we get $\alpha^* \to 1$. Thus the controller targets the periphery when more budget is available to them.

Setting $\alpha^* = 0$ in Eq. (3.7) yields the critical point when controller A withdraws all its resources from the periphery as,

$$b_c = a(a + 2n + 1). \tag{3.8}$$

Setting $\alpha^* = a$ on the other hand, yields $b = -a$, implying that the hub is always targeted when competitor allocations are focused on the periphery.

3.4.4   Comparing optimal allocations to common heuristics

So far we have presented analytical solutions for optimal allocations in a star network. Next, we examine how effective the optimal strategy is against common heuristics often employed to target networks. More specifically, we determine how much

---

[6]Details of the derivations can be found in Appendix A.1.2

vote-share a controller can gain from employing the optimal strategy in contrast to naïve heuristic approaches.

Here we generalise our results by considering a flexible allocation strategy for competitor B, where they allocate a fraction ϵ_B of their budget to the peripheral nodes and the rest $(1 - \epsilon_B)$ to the hub. Therefore by varying ϵ_B, controller B regulates the fractions of resources allocated to the hub (or the periphery). Observe that, here $\epsilon_B = 0$ presents the setting where allocations to the network are discrete, while $\epsilon_B > 0$ indicates continuous allocations to the network. Given that vote-shares are shown to increase monotonically with the number of targeted peripheral nodes in a star graph (against both discrete and continuous competitor allocations), here we assume that the ϵ_B fraction of the budget is uniformly distributed over all n peripheral nodes.

Assuming $p_A(\alpha, k_A)$, we find[7] that $\nabla_{k_A} X_A \geq 0$. Thus setting $k_A^* = n$ yields,

$$X_A = \frac{(((1 - \epsilon_B)n - \epsilon_B)b + a(n+1))\alpha + a(b\epsilon_B + n(n+1)) - (n+1)\alpha^2}{(((1 - 2\epsilon_B)b + a - \alpha^2)\alpha + (1 - \epsilon_B)b^2\epsilon_B + (a\epsilon_B + n)b + an)(n+1)}. \tag{3.9}$$

Solving the partial derivative $\nabla_\alpha X_A = 0$ then gives us solutions for optimal allocations,

$$\alpha^* = \frac{1}{1 - \epsilon_B(n-1)}\Big(b\epsilon_B(\epsilon_B(n+1) - 1) - n(\epsilon_B(a+b) + n)$$
$$\pm \sqrt{n(a+b)((n+1)^2 + (a+b))(n + \epsilon_B(1 - \epsilon_B)) - n}\Big).$$

We ignore the expression with the negative root as it violates the non-negativity constraint ($\alpha^* < 0$), leaving us with one feasible solution for optimal allocations in this setting,

$$\alpha^* = \frac{1}{1 - \epsilon_B(n-1)}\Big(b\epsilon_B(\epsilon_B(n+1) - 1) - n(\epsilon_B(a+b) + n)$$
$$+ \sqrt{n(a+b)((n+1)^2 + (a+b))(n + \epsilon_B(1 - \epsilon_B)) - n}\Big). \tag{3.10}$$

Observe that setting $\epsilon_B \in \{\frac{n}{n+1}, 0, 1\}$ in Eq. (3.10) yield the respective expressions for optimal allocations obtained in Sections 3.4.1 to 3.4.3.

We now determine the effectiveness of the optimal strategies against heuristics that are commonly employed to maximise influence in social networks. Inspired by the structure of the star network we consider four such approaches for controller A: (a) targeting the hub, (b) targeting the periphery uniformly, (c) targeting all nodes

---

[7]We refrain from showing the derivations for this instance as the resulting equations are unwieldy. However, please note that this case is easily captured in the game-theoretic setting (the derivations for which are shown in Appendix A.2), where we show that the result $\partial X_A / \partial k_A \geq 0$ is true for any arbitrary competitor strategy.

uniformly and (d) degree-based targeting. The first three approaches have been discussed in detail in earlier sections. In degree-based targeting, we assign allocations to nodes in proportion to their degrees. In the case of star networks with n peripheral nodes, $\alpha = 0.5B_A$.
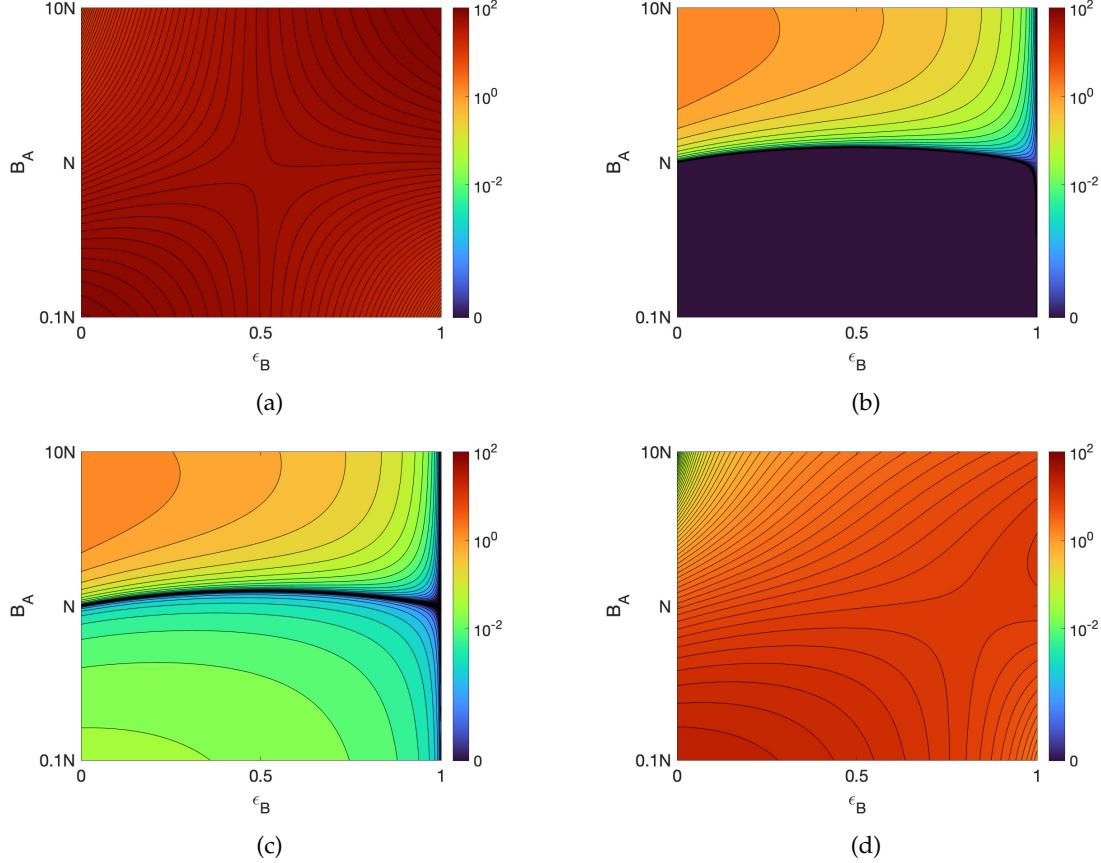
FIGURE 3.4: Figure showing the gain in vote-shares in a star network of size $N = 1000$, obtained by employing the optimal allocation strategy, compared to common heuristics such as (A) targeting the hub, (B) targeting the periphery uniformly, (C) targeting all nodes uniformly and (D) degree-based targeting. The percentage gain in vote-share $\Delta X_A(\%) = ((X_A^*/X_A^h) - 1) \times 100$, is illustrated for each case as controller budget B_A and fraction of competitor budget to the periphery ϵ_B are varied.

Fig. 3.4 illustrates the percentage gain in vote-shares $\Delta X_A\% = ((X_A^*/X_A^h) - 1) \times 100$, where X_A^* is the vote-share obtained from optimal allocations (determined using Eq. (3.10)) and X_A^h is the vote-share obtained when using heuristic approaches. We vary competitor allocations using the parameter ϵ_B as $0 \leq \epsilon_B \leq 1$ and the controller budget between $1 \leq B_A \leq 10N$ to determine the effect of controller budget and competitor allocations on the gain in vote-shares $\Delta X_A\%$. The effect of competitor budget on $\Delta X_A\%$ is considered later in the section. For now we set $B_B = N$.

Fig. 3.4a illustrates the gain in vote-shares when a controller switches from (a) targeting the hub to targeting the network optimally. The gain is maximum $\Delta X_A\% \approx 90.8\%$ when the $\epsilon_B = 1$, i.e. B targets the periphery and has a large budget

$B_A = 10N$. The gain in vote-share implies that nodes on average are twice as likely to adopt state A, which yields a significant advantage to the controller. The minimum gain in this setting is close to $\Delta X_A\% \approx 7.3\%$, which is still a considerable increase in vote-shares.

Figs. 3.4b and 3.4d illustrate the gain in vote-shares when optimal allocations are compared to other approaches, such as (b) targeting the periphery, or (c) targeting all nodes uniformly. In both instances we find that the heuristic approaches work reasonably well compared to the optimal strategy, and yield little to no gain in vote-shares in most instances. Note that $\Delta X_A = 0$ for low budgets $0.1N \leq B_A \leq N$ in Fig. 3.4b, implies that the optimal strategy allocates resources strictly to the periphery, independent of the competitor strategy. In both Figs. 3.4b and 3.4d, we find that the optimal strategy is more effective for large budgets for controller A. The maximum gain is $\Delta X_A\% \approx 2.7\%$ in both instances, obtained under similar settings, $B_A \approx 6N$ and $\epsilon_B = 0$ (B targets the hub).

Fig. 3.4c compares the optimal strategy to the degree-dependent heuristic approach. The optimal strategy measurably outperforms the heuristic and the maximum gain in vote-shares $\Delta X_A\% \approx 25.9\%$ is obtained when $\epsilon_B = 0$ and controller budget is $B_A = 0.1N$.

Overall, we observe that regardless of the competitor strategy, continuous allocations almost always perform better than the discrete approach. Additionally, contrary to past work that highlight the significance of hub nodes in the *influence maximisation* process (Masuda (2015); Kuhlman et al. (2013)), here we observe that peripheral nodes are equally important as hub nodes in the spread process, and external controllers can gain significantly from targeting peripheral nodes.

3.4.5 Effect of competitor budget

We now analyse the impact of competitor budget on the effectiveness of the optimal strategy. Here we compare optimal allocations to the traditional discrete method where only the hub is targeted. We focus on two settings: (i) B targets the hub ($\epsilon_B = 0$) (ii) B targets the periphery. Both controller budgets are varied as $0.1N \leq B_A \leq 10N$ and $0.1N \leq B_B \leq 10N$. Results are shown in Fig. 3.5, where Fig. 3.5a illustrates the setting where B targets the hub ($\epsilon_B = 0$), and B targets the periphery in Fig. 3.5b ($\epsilon_B = 1$).

Observe in Fig. 3.5a, that the gain in vote-shares is highly sensitive to the change in competitor budget B_B. In Fig. 3.5a, we find that the maximum gain in vote-shares can be as high as $\Delta X_A\% \approx 900.8\%$ at $B_A = 0.1N$ and $B_B = 10N$, which is 10 times the vote-share obtained when the controller targets the hub discretely. When competitor B

targets the periphery, as shown in Fig. 3.5b, the maximum gain is $\Delta X_A\% \approx 494.5\%$ at $B_A = B_B = 10N$.

We also observe that gain in vote-shares is always positively correlated with the competitor budget B_B, whereas correlation to controller budget B_A depends on competitor allocations. Gain in vote-shares is negatively correlated with B_A when B targets the hub, and positively correlated with B_A when B targets the periphery. This implies that it is better to target the hub with increase in B_A when B targets the hub. On the other hand, when B targets the periphery, a controller loses more vote-shares by targeting the hub as B_A increases. Recall that we obtained similar results in Sections 3.4.2 and 3.4.3, where we observed that it is optimal to avoid nodes targeted by the competitor unless the controller has sufficient resources.

| (a) | (b) |
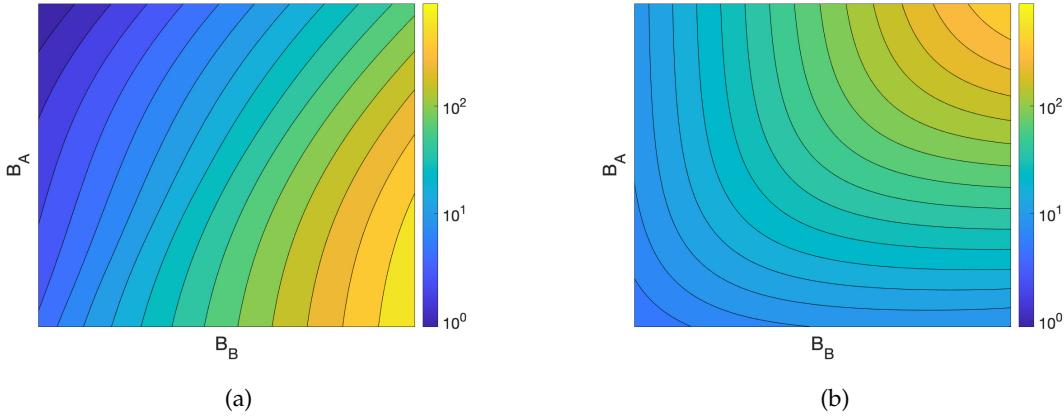
FIGURE 3.5: Figure showing the effect of competitor budget on gain in vote-shares in star network. Here we compare the optimal strategy to a heuristic approach, where only the hub is targeted. Figure (a) illustrates the percentage gain in vote-shares $\Delta X_A(\%)$, when competitor B targets the hub $\epsilon_B = 0$. While Figure (b) shows the percentage gain in vote-shares, when competitor B targets the periphery $\epsilon_B = 1$. Controller budgets in both cases are varied as $0.1N \leq B_A \leq 10N$ and $0.1N \leq B_B \leq 10N$.

In this section, we explore optimal strategies against known competitor allocations in a star network. We define settings where the competitor targets (i) all nodes uniformly, (ii) the hub or (iii) the periphery, and we present how budget conditions and competitor allocations drive optimal allocations in each case. We further highlight the importance of optimal strategies, even in a simplified network structure such as the star network, by comparing its performance to common heuristics. We find that continuous approaches that prioritise peripheral nodes perform better than traditional discrete approaches that preferentially target the hub node. Finally, we show that in the some instances, the loss in vote-shares can significantly increase as competitor budget increases.

We now proceed to examine the problem in game-theoretic settings where no prior knowledge of competitor allocations is available.

3.5 Game-theoretic setting

So far we have presented solutions for optimal allocations in the star network against fixed competitor allocations. However, obtaining prior knowledge of competitor allocations in the real world is rarely possible. With this is mind, we define the *competitive influence maximisation* problem for a more realistic game-theoretic setting, where both controllers actively and simultaneously target the network. Here the game has two players, controllers A and B, who maximise their vote-shares (or utilities), by optimising their strategies over a set of all possible allocations ($\sum_i p_{A,i} = B_A$ and $\sum_i p_{B,i} = B_B$). Controllers play the game iteratively, until it converges to an equilibrium from where neither controller deviates. To better explain the settings considered in our game, we motivate it with an example.

Consider a scenario where two companies seek to maximise the sale of their product in a population. Each company has a budget which they distribute over the network to achieve this. We assume that neither company has any prior knowledge of the competitor's allocation on the network, as there is no incentive to disclose this information in advance. However, companies can gain this knowledge through their interaction with the network during the marketing process, and can then use this knowledge to strategise in future campaigns. We implement it in our model by revealing the allocation vectors (p_A and p_B) at the end of each round. Each player then observes the opponent allocation and plays the best response to this strategy in the following round. We assume multiple iterations of the game are played until an equilibrium state is achieved. In the real world, this corresponds to a stable outcome, from where neither company deviates, i.e. companies are no longer motivated to change their marketing strategies.

To determine the optimal allocations at the equilibrium state, we first parameterise allocation vectors for both controllers A and B as $p_A(\alpha, k_A)$ and $p_B(\beta, k_B)$, where allocations are defined by the number of targeted peripheral nodes (k_A and k_B) and the total amount of resources allocated to the periphery by each controller (α and β respectively). We insert the above allocation vectors in Eq. (3.2) to obtain the relevant vote-share expressions for each controller. Note that, assuming A and B each target k_A and k_B peripheral nodes, we get four disjoint classes of peripheral nodes that should be considered while determining X_A and X_B. These are: (i) $\frac{k_A k_B}{n}$ nodes targeted by both controllers, (ii) $k_A - \frac{k_A k_B}{n}$ nodes targeted exclusively by A, (iii) $k_B - \frac{k_A k_B}{n}$ nodes targeted exclusively by B and finally, (iv) $(n - (k_A + k_B - \frac{k_A k_B}{n}))$ nodes that are not targeted by either controller.

We determine the following partial derivatives $\nabla_{k_A} X_A$ and $\nabla_{k_B} X_B$. We find that the vote-share for each controller increases monotonically with the number of targeted

peripheral nodes[8] ($\nabla_{kA} X_A \geq 0$ and $\nabla_{kB} X_B \geq 0$). We can now eliminate all dominated strategies from our analysis by replacing $k_A = n$ and $k_B = n$ in X_A and X_B, and finally solve $\{\nabla_\alpha X_A, \nabla_\beta X_B\} = 0$ simultaneously, to obtain

$$\{\alpha^*, \beta^*\} = \left\{\frac{an}{n+1}, \frac{bn}{n+1}\right\}. \tag{3.11}$$

This implies that the optimal strategy for both controllers when targeting the network simultaneously is to uniformly distribute their influence over the network. However, this result may be an artefact of the star structure, and thus to determine if the result is independent of network structure, we explore the problem further in a real-world network in the following section.

3.5.1   Numerical results on a real-world network

When considering the game-theoretic setting in star networks, we find that the equilibrium strategy for both controllers is to uniformly target the network, irrespective of controller budgets. While this is an interesting result, it is important to determine if this is an artefact of the distinctive structure of a star graph, or if the results apply equally in real-world networks.

To do this, we first propose a numerical method to deal with the optimisation problem in large, arbitrary networks. Recall that we determine optimal allocations in the case of a star graph by solving $\nabla_{p_A} X_A = 0$. We follow the same approach for arbitrary networks and obtain an analytical expression for the gradient using Eq. (3.2) as follows,

$$\nabla_{p_A} X_A = \frac{1}{N} \vec{1}^T [L + diag(p_A + p_B)]^{-1} (I - diag(x_A)). \tag{3.12}$$

Solving $\nabla_{p_A} X_A = 0$ analytically for larger, arbitrary networks is challenging. Hence we adopt a local search technique based on gradient ascent, which has been shown to work well in similar settings (Lynn and Lee (2016)), to determine the optimal allocations for which vote-share is maximised.

The gradient ascent algorithm (shown in Algorithm 1) iteratively updates the allocation vector p_A by taking incremental steps (determined by step-length η) in the direction of the gradient of the objective function Eq. (3.2), until the vote-share can no longer be improved. The algorithm takes as its input the network structure L, the controller budget B_A, the competitor allocation p_B, the step-length η and an

---

[8]Details of the derivations can be found in Appendix A.2

Algorithm 1: Gradient Ascent (GA) algorithm

Input : $L, B_A, p_B \eta, \mu$;
Result: p_A^*
initialise t=0, $p_A^{(0)}$;
while $p_A^{(t+1)} - p_A^{(t)} > \mu$ **do**
$\quad$ $p_A^{(t+1)} = \mathcal{P}[p_A^{(t)} + \eta \nabla_{p_A^{(t)}} X_A]$;
$\quad$ **if** $X_A^{(t+1)} - X_A^{(t)} < 0$ **then**
$\quad\quad$ $\eta^{(t+1)} = \frac{\eta^{(t)}}{2}$;
$\quad$ **end**
$\quad$ t++;
end

Algorithm 2: Modified GA algorithm for the game-theoretic setting

Input : $L, B_A, B_B, \eta, \mu_{inner}, \mu_{outer}$;
Result: p_A^*, p_B^*
initialise t=0, $p_A^{(0)}, p_B^{(0)}$;
while $(p_A^{(t+1)} - p_A^{(t)}) > \mu_{outer} \ || \ (p_B^{(t+1)} - p_B^{(t)}) > \mu_{outer}$ **do**
$\quad$ $p_A^{(t+1)} = GA(L, B_A, p_B^{(t)}, \eta, \mu_{inner})$;
$\quad$ $p_B^{(t+1)} = GA(L, B_B, p_A^{(t)}, \eta, \mu_{inner})$;
$\quad$ t++;
end

---

approximation factor μ to return a $\mu-$approximated optimal strategy. Note that the budget constraint is systematically imposed at each step, by projecting the resulting vector $[p_A^{(t)} + \eta \nabla_{p_A^{(t)}} X_A]$ onto an N-simplex, by employing the algorithm proposed in (Chen and Ye (2011)), such that $\sum_i p_{A,i} = B_A$. We modify the learning parameter η using backtracking line search[9] to ensure convergence[10] (Boyd et al. (2004)).

On taking a better look at Eq. (3.2), we find that the vote-share (X_A) for the focal controller (A) is concave in p_A (controller's allocation) and convex in p_B (competitor's allocation)[11]. Note that controllers in this setting are identical and hence inter-changeable. Thus the concavity result naturally applies to B as well. Moreover, observe that the constraint sets $\mathcal{P}_A$ and $\mathcal{P}_B$, for allocation vectors p_A and p_B, are both compact and concave. Therefore the current game-theoretic setting belongs to the class of concave-convex games, which characteristically have a unique equilibrium point that can be reliably found using gradient-based search methods (Rosen (1965)).

---

[9]When $X_A(t+1) < X_A(t)$, the solution $p_A(t+1)$ at time step $(t+1)$ is rejected and the allocation vector is optimised again with an updated step-length $\eta(t+1) = \frac{\eta(t)}{2}$.

[10]The time-complexity to obtain a $\mu-$approximated solution is given by $O(N^3/\mu)$. This is because each step in the algorithm involves the inversion of an $N \times N$ matrix which has a complexity of $O(N^3)$.

[11]Details of the proof are shown in Appendix A.3.

Hence we extend the proposed gradient-ascent algorithm to the game-theoretic setting (shown in Algorithm 2). Allocation strategies for both controllers are initialised as random vectors $p_A^{(0)}$ and $p_B^{(0)}$. In each step t, allocation vectors are simultaneously optimised, using the gradient ascent method, for each controller ($p_A^{*(t)}$ and $p_B^{*(t)}$), against the respective competitor's allocations, $p_B^{*(t-1)}$ and $p_A^{*(t-1)}$.

(a)   (b)
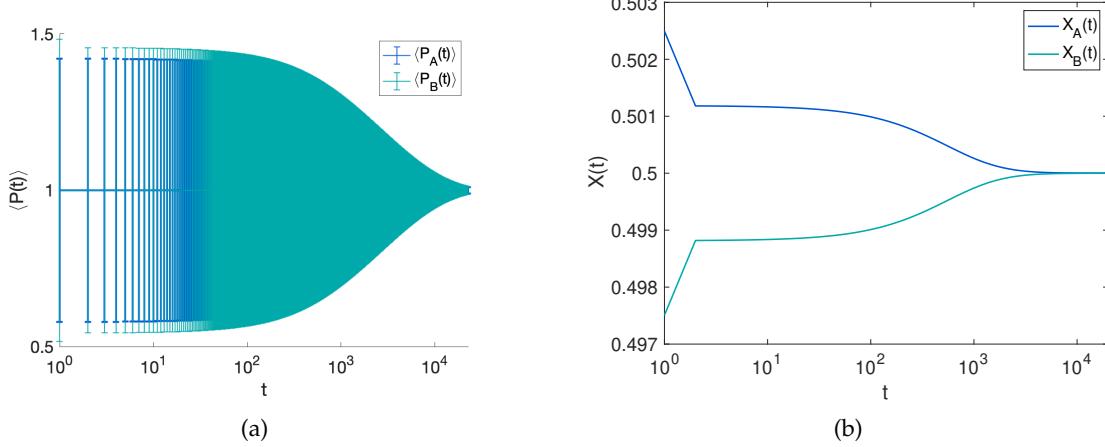
FIGURE 3.6: Figures illustrating the equilibrium strategies obtained using numerical methods in an email network of $N = 1133$. Figures (A) and (B) illustrate results when controllers have equal budgets $B_A = B_B = N$. Figure (a) shows mean allocations to nodes from both controllers along with standard deviations over t iterations of the game. Figure (b) shows the corresponding vote-shares obtained by the controllers.

The modified GA algorithm is then applied to a real-world network. We consider an email-interaction network that maps communications between employees of a university (University Rovira i Virgili (Tarragona)) (Guimera et al. (2003); Rossi and

(a)   (b)
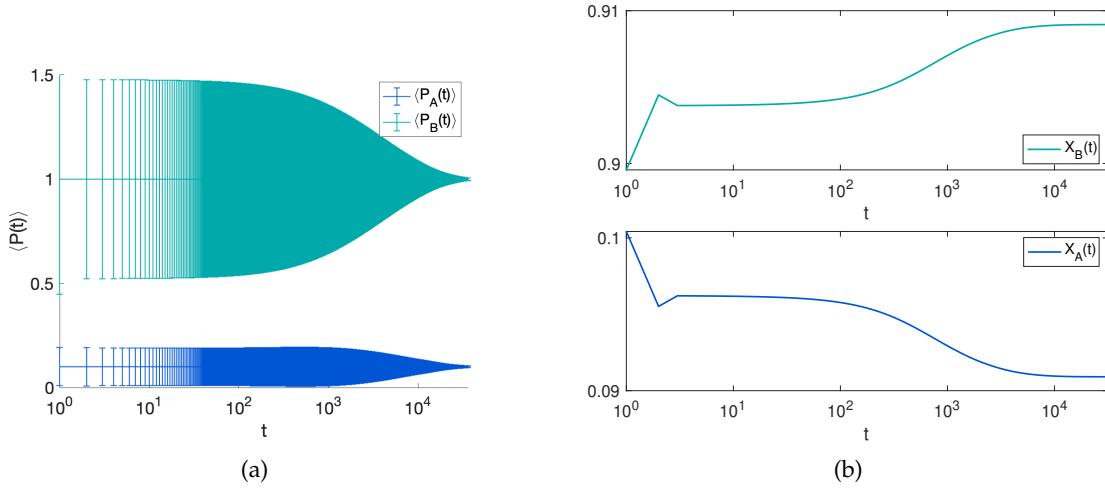
FIGURE 3.7: Figure showing equilibrium strategies obtained using numerical methods in an email network of $N = 1133$. Figures (A) and (B) illustrate results when controllers have unequal budgets $B_A = 0.1N$ and $B_B = N$. Figure (a) shows mean allocations to nodes from both controllers along with standard deviations over t iterations of the game. Figure (b) shows the corresponding vote-shares obtained by the controllers.

Ahmed (2015)). For our experiments, we use the largest connected component of size $N = 1133$ nodes, with an average degree of $\langle k \rangle \approx 9.62$. We consider two instances: (i) where controllers have equal budgets $B_A = B_B = N$, and (ii) where controllers have unequal budgets, $B_A/B_B = 0.1$ and $B_B = N$. In each case we initialise the step-length for the gradient ascent algorithm at $\eta = 10$, and keep the approximation factors $\mu_{outer} = 10^{-10}$ and $\mu_{inner} = 10^{-3}$. The speed of convergence can be controlled by tuning the step-length and approximation factors. Our results show that in both cases the algorithm converges in less than 10^5 iterations.

Figs. 3.6 and 3.7 illustrate numerical results for equal and unequal controller budgets respectively. We find that in both cases the algorithm converges to an equilibrium. Figs. 3.6a and 3.7a show the mean allocations to nodes in the network, and Figs. 3.6b and 3.7b show the evolution of vote-shares (X_A and X_B) with time. Observe how the standard deviation of allocation vectors decreases to 0, as the algorithm reaches convergence. Here too, we find that the equilibrium strategy is to target the network with uniform allocations, irrespective of the budget available to the controller, which generalises our results obtained in the star network to real world settings.

3.6 Summary

In this chapter, we explore the *competitive influence maximisation* problem with continuous allocations in the voter model. Contrary to traditional methods, where nodes are typically targeted in a binary fashion, here we consider continuous allocation of influence to the network where an array of nodes are targeted with varying intensities, based on their perceived importance in the *influence maximisation* process.

We assume two controllers compete to maximise their vote-shares in the network. Two instances for competitor allocations are considered. One where competitor allocations are fixed and known in advance. The other where the competitor is active and both controllers target the network simultaneously. When competitor allocations are unknown, we frame the problem in a game-theoretic setting and obtain the optimal allocation for both controllers by determining the equilibrium point in the game.

For both types of competitor allocations, we obtain analytical closed-form solutions for optimal allocations in a star network and show how the optimal allocations vary with controller budgets and competitor allocations. We find that when the competitor has a large budget, it is better to avoid the nodes targeted by them, and focus instead on the nodes avoided by the competitor. As controller budget increases, we observe that more resources are allocated to nodes targeted by the competitor. We also observe that, irrespective of competitor allocations, allocating influence continuously to the network consistently outperforms the discrete approach, and we show that a

controller can nearly double their vote-share by targeting the network continuously, as opposed to discretely targeting the hub node.

In the game-theoretic setting, we find that the optimal allocation for both controllers in a star network, irrespective of the controller budgets, is to uniformly target the network. We further explore the problem in a larger network, and show that the results for optimal allocations are the same even in a real-world network.

Finally, we recognise that assuming a linear cost function for controllers (i.e. where allocations are linearly proportional to the strength of influence), may be a naïve assumption that is not fully representative of real world settings. Therefore, to generalise our model further, in the following chapter, we consider other nonlinear cost functions, and systematically examine how optimal allocations vary in each case.

Chapter 4

Nonlinear cost of allocations

4.1 Introduction

In this thesis, we explore the problem of competitive *influence maximisation* in social networks. We consider continuous allocation of resources over a network where nodes are targeted with varying strengths —based on their role in the influence spread process. In our work so far, we have assumed that the amount of resources allocated to a node is equivalent to the strength of influence it experiences (i.e. amount of resources allocated to an individual is directly proportional to the effect of influence). As shown in Chapter 3, this yields a convex optimisation problem that can be effectively solved using search methods such as gradient ascent (Lynn and Lee (2016)). However, as argued in Section 3.6 of Chapter 3, a linear relationship between the cost and effect of influence is a naïve assumption, as processes in the real world are rarely ever linear. With this in mind, in this chapter we extend our work further to study competitive *influence maximisation* in the voter model under alternative and more realistic cost functions[1].

First, we consider the instance where a controller pays an additional cost to access nodes in a network with the intention of influencing them. In the real world, the additional cost is analogous to removing adoption barriers which are prevalent in many societies (Mobarak and Saldanha (2022)). For instance, lack of information and understanding about vaccinations often leads to its poor uptake in certain populations (Mills et al. (2005); Esposito et al. (2014)), In such instances, removing this barrier by providing the necessary information to people about vaccinations can improve uptake rates in populations. For example, it was observed that providing free and safe contraception to women (that come at an additional cost to health authorities), particularly in low-income countries, encourage them to visit health centres where they can be informed about the health benefits of vaccinating their children, often

---

[1]These functions define the relationship between cost and effect of influence on individuals.

leading to a rise in vaccination uptakes in the population (Gates (2019)). Alternatively, there may be economic barriers (Knapp et al. (2006); Zhu et al. (2014)), in which case, individuals are unable to adopt a product or lifestyle change due to the lack of financial means. Thus in these instances, subsidies that come at a cost to the controller can make people more open to trialing the product. Motivated by this, we extend our competitive *influence maximisation* model to include an additional cost of access parameter, that a controller is subjected to, when targeting an individual in the network. We explore multiple instances of competitor allocations, and we show that irrespective of competitor allocations, the cost of access parameter regulates the optimal configuration of allocations between discrete and continuous approaches. Thus bridging the gap between our model and the traditional discrete methods of *influence maximisation*.

In addition, we consider instances where the effect of influence on individuals vary nonlinearly with the cost of influence. For instance, studies from marketing research show that more is *not* always better, and that the duration and complexity of advertisements often have a *diminishing returns* effect on customer engagement and interest (Fortin and Dholakia (2005); Goldstein et al. (2011)). Alternatively, individuals may respond to external influence slowly at first and then rapidly, which can occur due to increased popularity or better understanding of a product or technology, such as in the case renewable technologies (Zhang et al. (2015)). Thus yielding a convex relation between the cost and effect of influence (observed as a *delayed effect in influence*). We determine optimal allocations against varied competitor allocations in each case and show that targeting the network optimally is crucial in many settings, where simple heuristics (such as degree-based targeting or uniform allocation of resources) can cause the controller to lose a significant share of votes in the population.

4.2 Outline

The rest of the chapter is structured as follows:

In Section 4.3 we modify the model proposed in Chapter 3 to include the cost of access parameter.

In Section 4.4 we discuss analytical results of optimal allocations for the competitive *influence maximisation* problem with fixed cost of access in a star network.

In Section 4.5 we propose heuristics to optimise vote-shares in arbitrary network structures, and discuss their performance first in a star network for comparison, and then in a real-world network.

In Section 4.6 we discuss nonlinear cost functions and their effect on optimal allocations and vote-shares. We obtain optimal results in both synthetic networks and a real-world network and compare them using numerical and analytical approaches.

In Section 4.7 we summarise our results and discuss possible future directions to expand this work.

4.3 Voter dynamics with fixed cost of access

A population of N individuals is considered. The structure of the network is represented mathematically using a non-negative weighted adjacency matrix W, as shown in Chapter 3. Here too we consider unweighted, undirected graphs where all links have the same weight $w_{ij} = w_{ji} = 1$. We assume that influence flow in the network follows traditional voter dynamics (Holley and Liggett (1975)). As before (shown in Chapter 3), we have two controllers A and B, competing to maximise their influence in the network. In this chapter, we optimise allocations for the focal controller A, only against fixed competitor allocations (from controller B), which can also be perceived as unfavourable bias in the network, against opinion A.

Allocations on the network from both controllers are captured in the vectors p_A and p_B, constrained by their respective budgets B_A and B_B. In addition, here we assume that the focal controller pays a fixed cost c to access each node in the network, such that the total cost of targeting a node is given by $(p_A + c)$. It is worth noting that the cost c is only paid when controller A targets the node with positive allocation i.e. $p_A > 0$. For $p_A = 0$, we assume $c = 0$. Assuming cost of access c is uniform for all nodes in the network, we modify the budget constraint as follows,

$$\vec{1}^T(p_A + c \cdot \mathbb{1}_{p_A > 0}) = B_A. \tag{4.1}$$

Here $\mathbb{1}_{p_A > 0}$ is an indicator vector, where the ith position is 1 if node i has positive allocations ($p_{A,i} > 0$), else the value is 0 for $p_{A,i} = 0$.

Since competitor allocations are fixed, we do not consider a cost of access for B, as any additional cost paid by the competitor to access nodes in the network can be directly accounted for by adjusting the strength of their influence B_B on the network.

Observe that, the vote-share function is the same as before. Given by,

$$\implies X_A = \frac{1}{N}\vec{1}^T x_A = \frac{1}{N}\vec{1}^T[L + diag(p_A + p_B)]^{-1} p_A. \tag{4.2}$$

Finally, the optimisation problem is given by,

$$p_A^* = \arg\max_{p_A \in \mathcal{P}} X_A^*(L, p_B), \tag{4.3}$$

where $\mathcal{P}$ is a set of all possible allocations subject to Eq. (4.1).

We begin with an exploratory analysis of the problem in a simplified network structure, such as the star network. From here we draw insights about optimal allocation patterns in the given setting, which are then employed to propose heuristics that examine larger, arbitrary networks —in an effort to generalise our results.

4.4   Analysis of the star topology

Consider an unweighted and undirected star network of size N nodes. We examine a setting where competitor B uniformly allocates resources to the network, i.e. every individual in the network is equally opposed to adopting opinion A. In the real world, this represents uniform resistance to the adoption of new ideas and innovations (Wiedmann et al. (2011); Heidenreich and Kraemer (2016)), such as, aversion to vaccinations (Streefland (2001); Hobson-West (2003)), and technology (Huang et al. (2011)). Some of this is caused through economic, information and other adoption barriers (Christodoulakis et al. (2017); Vanclay (1992); Butler and Sellbom (2002); Redpath (2012)). Assuming that controller A has to pay an initial cost to prime each node for targeting (possibly by removing any existing barriers to adoption), we determine optimal configurations for which A maximises their opinion share in the network against B.

Competitor allocations to the network are fixed and constrained by the budget available to them $B_B = b$. Here we assume that B targets every node in the network with strength 1, i.e. $b = N$, and hence the competitor allocation vector in this instance is given by $p_{B,i} = 1$ ($\forall i \in \{1, \ldots, N\}$). The allocation vector for controller A is defined using the parameters α and k_A, where α is the amount of resources allocated to the periphery and k_A is the number of targeted peripheral nodes. We solve the partial derivatives $\Delta_\alpha X_A = 0$ and $\Delta_{k_A} X_A = 0$ to obtain the optimal allocations $p_A^*(\alpha^*, k_A^*)$ as,

$$\alpha^* = \frac{k_A^*(b+n+1)(a-(k_A^*+1)c)}{b(k_A^*+1)+(n+1)^2},$$

where k_A^* is given by,

$$k_A^* = \frac{1}{c(b^2 - c(n+1)(b+n+1))}\Big((n+1)(b+n+1)c^2 \quad \text{(4.4)}$$
$$- ((ab+2a+2b)n + (b+1)(a+b)))c$$
$$+ ((n+1)^2 c + ab)\sqrt{c(n+1)(b+n+1)}\Big),$$

subject to $0 \le \alpha^* \le a - (k_A^*+1)c$ and $0 \le k_A^* \le min(n, (\frac{a}{c}-1))$.

Here we make two observations about optimal allocations in star networks. First, we find that when $k_A^* = n$, optimal allocation to the periphery is

$$\alpha^* = \frac{(a-(n+1)c)n}{(n+1)},$$

implying that the optimal configuration for allocations in this case, is to uniformly distribute available resources to the network. Observe that, setting $c = 0$ yields $\alpha^* = an/(n+1)$, which is also the result we obtained in Section 3.4.1 of Chapter 3.

Second, we find that number of targeted peripheral nodes k_A^* is a quadratic in cost of access c. The detailed structure of optimal allocations however, is not immediately obvious from Eq. (4.4), and further analytical exploration of Eq. (4.4) (e.g. under boundary conditions) yields lengthy expressions that offer little insight. Thus to better understand the pattern of optimal allocations, we plot our results in Fig. 4.1 for varying costs of access c, where the optimal number of targeted peripheral nodes k_A^* is shown in Fig. 4.1a, while Fig. 4.1b shows the allocations to the hub node ($i = 1$) as a fraction of total allocations, also given by, $p_{A,1}/\sum p_A = 1 - (\alpha^*/a - (k_A^*+1)c)$.

Our analytical results are complemented with numerical results obtained in a star network of size $N = 1000$. Controllers have equal budgets $B_A = B_B = N^2$ Optimal allocations $p_A^*(\alpha^*, k_A^*)$ are obtained numerically by using a brute-force method where α is discretised between $0 \le \alpha \le (a - (k_A + 1)c)$ using $\Delta\alpha = 0.001$, and $\Delta k_A = 1$ for $0 \le k_A \le min(n, (\frac{a}{c}-1))$.

Overall, Fig. 4.1 shows that our analytical results are in good agreement with the numerical results. In Fig. 4.1a we see that k_A^* gradually decreases as the cost of access c increases. To an extent this makes intuitive sense as fewer nodes are expected to be targeted when the cost of accessing nodes is expensive. However, when we combine this result with Fig. 4.1b, an interesting observation emerges. We find that the cost of access variable c behaves as a parameter that tunes the optimal configuration of

---

[2]We assess the impact of controller budgets on optimal allocations later in the section.

(a)   (b)
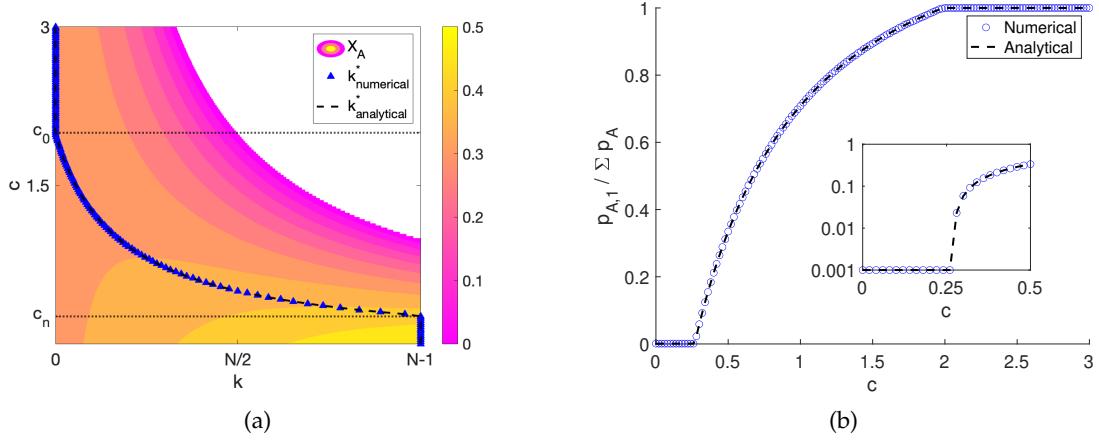
FIGURE 4.1: Figure showing how optimal allocations vary with cost of access c, in a star network of size $N = 1000$. Controller budgets are equal $B_A = B_B = N$. Analytical solutions are obtained using Eq. (4.4) and numerical results are obtained using a brute-force method, where α is discretised as $\Delta\alpha = 0.001$ and k_A as $\Delta k_A = 1$. Figure (a) shows how the optimal number of targeted peripheral nodes k_A^* changes with cost of access c on a heatmap showing vote-shares (X_A). The blank region corresponds to values that do not meet the budget constraint, $0 \leq k_A^* \leq min(n, (\frac{a}{c} - 1))$. Figure (b) shows how allocations to the hub node vary (as a fraction of total allocations $p_{A,1} / \sum_i p_{A,i}$), with change in cost c. The inset shows a closer view of allocation behaviour at low values of c. The inset shows a closer examination of allocation behaviour at low values of c.

allocations to oscillate between continuous and discrete distributions on the network. More specifically, we observe that for low values of c, influence is continuously distributed over the network, i.e. $k_A^* = n$. However, as c increases, a more discrete approach is preferred, where resources are diverted away from the periphery to the hub node.

Additionally, we find that optimal allocations can be divided into three regimes: (i) $k_A^* = n$, (ii) $0 < k_A^* < n$ and (iii) $k_A^* = 0$. Observe that these regimes surface as the cost of access varies from (i) $c \leq c_n$ to, (ii) $c_n < c < c_0$, and finally (iii) $c \geq c_0$. We identify c_n and c_0 as critical points where k_A^* either departs from, or approaches the boundary solutions, i.e. $k_A = \{n, 0\}$. The value of the critical points c_n and c_0 are determined by setting $k_A^* = n$ and $k_A^* = 0$ in Eq. (4.4) respectively and then solving for c. This yields $c_n = 0.27$ and $c_0 = 1.99$, implying that as cost of access exceeds nearly 30% of the per node budget, optimal strategies redirect more resources to the hub node, and less to the periphery, until c is nearly twice the per-node budget —at which point all resources are focused on the hub node (i.e. $k_A^* = 0$).

We now proceed to determine how controller budgets affect optimal allocations against uniform competitor allocations. We do this by determining how critical costs c_n and c_0 vary with controller budgets (shown in Fig. 4.2).

We find that controller budgets significantly impact optimal allocations. We find that both c_n and c_0, shown in Fig. 4.2a and Fig. 4.2b respectively, are positively correlated
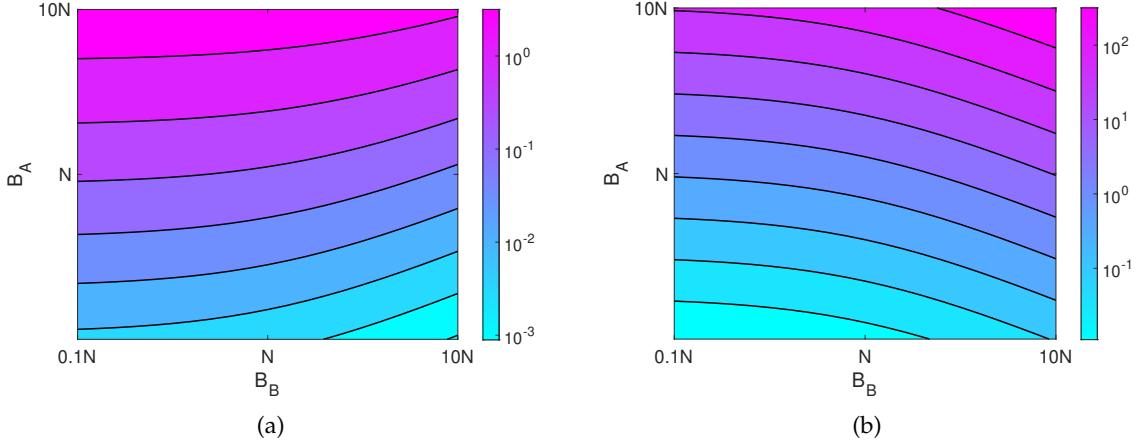
FIGURE 4.2: Figure showing how optimal allocations against uniform competitor allocations respond to change in controller budgets in a star network of size $N = 1000$. We plot the values of c_n and c_0 as budgets are varied between $0.1N \leq B_A \leq 10N$ and $0.1N \leq B_B \leq 10N$. Figure (a) plot c_n against both controller budgets, while Figure (b) illustrates the sensitivity of c_0 to controller budgets.

with the controller budget B_A, implying that, as controller budget increases, it takes longer to depart from the continuous regime and also takes longer to the arrive at the discrete regime. On the other hand, we find that c_n is negatively correlated with B_B and c_0 is positively correlated with competitor budget B_B. This means that as competitor influence on the network increases, optimal allocations depart sooner from the continuous regime but takes longer to reach the discrete regime.

So far, we observe a shift in optimal allocation of resources from peripheral nodes to the hub node as the cost of access increases. We now examine if this strategy where resources are directed to the hub node for high costs of access is the same irrespective of change in competitor allocations. Specifically, we consider the instance where the competitor solely targets the hub node, i.e. $p_{B,1} = B_B = b$ and $p_{B,i} = 0$ ($\forall i \in \{2, \ldots, N\}$). This also presents an interesting real world setting, where we assume only the leader of a social group is biased against opinion A. Considering there is a cost to access each member of the population and the hub node is inherently unfavourably biased, here we are determined to find out if it is still reasonable to focus on the hub node as cost of access increases.

The vote-share expression for controller A in this case is given by,

$$X_A = \frac{k_A(n+1)(a-(k_A+1)c)-(n+1)\alpha^2+((b-(n+1)c)k_A+(n+1)(a-c))\alpha}{(n+1)((a+b-(k_A+1)c)(k_A+1)-1)\alpha}.$$

(4.5)

We solve $\nabla_\alpha X_A = 0$ to obtain,

$$\alpha = \frac{\left(\pm \sqrt{(n+1)^2 + (n+1-k_A)(a+b-c(k_A+1))} - (n+1) \right) k_A}{n+1-k_A}.$$

Naturally, we discard the negative root, as it violates the non-negativity constraint for allocations, which leaves us,

$$\alpha^* = \frac{\left(\sqrt{(n+1)^2 + (n+1-k_A)(a+b-c(k_A+1))} - (n+1) \right) k_A}{n+1-k_A}. \tag{4.6}$$

Observe that, Eq. (4.6) will always yield non-negative, real solutions given that $a > c(k_A + 1)$, i.e. the total budget must always exceed the cost of accessing targeted nodes.

Now, replacing Eq. (4.6) in Eq. (4.5), and solving $\nabla_{k_A} X_A = 0$ we obtain,

$$k_A^* = \{ n+1, \frac{a+b+n+1}{c-1} \}. \tag{4.7}$$

Note that neither of these solutions meet our constraint $0 \le k_A \le n$. The first solution is discarded as k_A^* cannot exceed the total number of peripheral nodes in the network n. While considering the second solution, we look at two possible cases, $c < 1$ and $c > 1$. The expression is undefined at $c = 1$. For $c < 1$, we find that $k_A^* < 0$ and thus, is not a feasible solution. When $1 < c \le a$, the resulting solution breaches the constraint, $k_A^* > n$. Therefore we argue that the solution for k_A^* always lies on the boundary, $k_A \in \{0, n\}$, implying that optimal allocations follow an *all-or-none* strategy.

To get a better understanding of how optimal allocations behave with change in c, we employ numerical methods on a star network of size $N = 1000$. Controllers have equal budgets, $B_A = B_B = N$ and we use the brute-force method (where $\Delta\alpha = 0.001$ and $\Delta k_A = 1$) to obtain the optimal allocations. Optimal allocations are illustrated in Fig. 4.3.

Fig. 4.3a illustrates how the optimal number of targeted peripheral nodes k_A^*, changes with the cost of access c. While Fig. 4.3b illustrates the corresponding allocations to the hub. Overall, we find that allocation patterns still change from the continuous to the discrete approach as c increases. However, unlike the gradual diversion of allocations from the periphery to the hub node, observed in the instance where competitor allocations are uniform, here we observe a sharp change in strategy from continuous to discrete allocations. More specifically, for lower values of c, it is optimal to target the network with continuous allocations, $k_A^* = n$. However when $c \ge c_c$, where c_c is the critical cost, the optimal solution abruptly switches to the discrete regime $k_A^* = 0$.
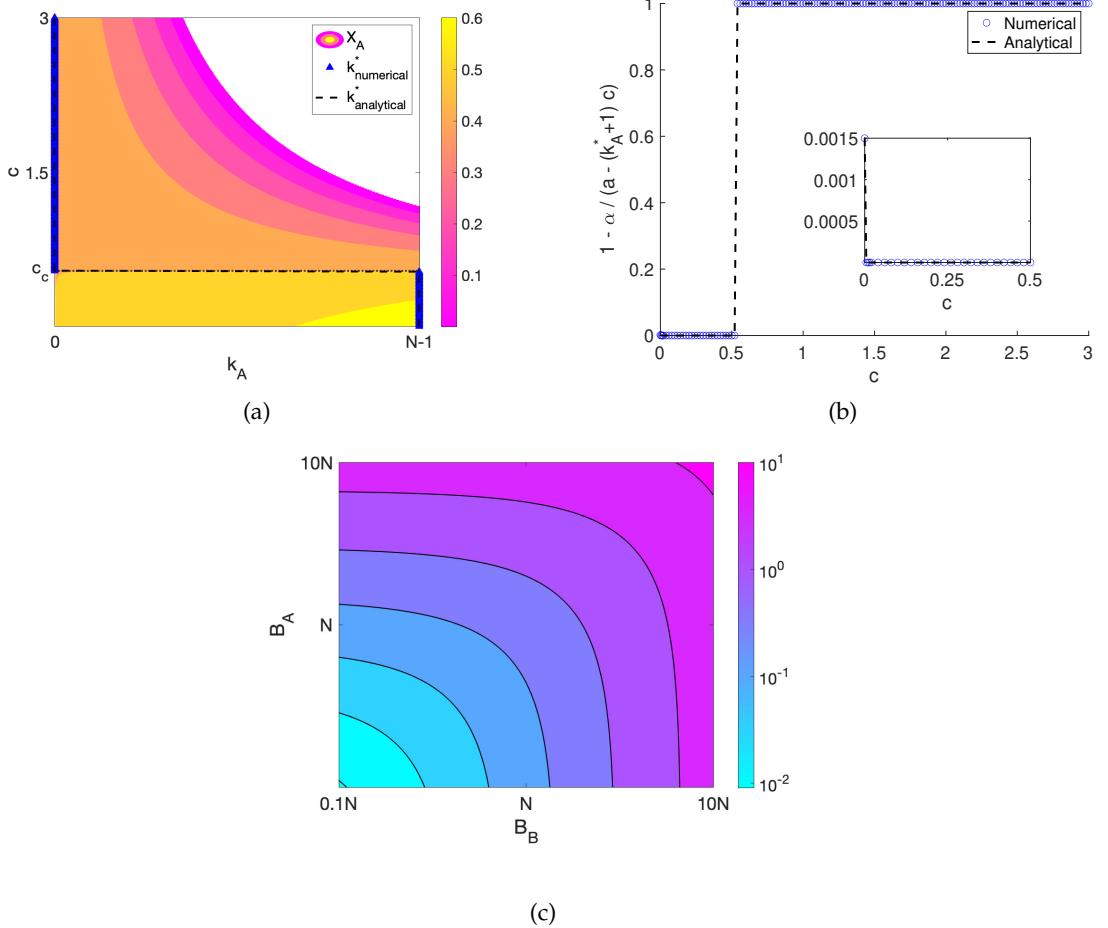
FIGURE 4.3: Figure showing how optimal allocations vary with cost of access c. All results are obtained for a star network of size $N = 1000$. The controller budgets here are equal $B_A = B_B = N$ and competitor B targets the hub node discretely. Analytical solutions for k_A^* are obtained by evaluating Eq. (4.5) at the boundary $k_A \in \{0, n\}$ and α^* is determined using Eq. (4.6). Numerical results are obtained using a brute-force method, where α is discretised as $\Delta\alpha = 0.001$ and k_A as $\Delta k_A = 1$. Figure (a) shows how the optimal number of targeted peripheral nodes k_A^* changes with cost of access c, superimposed on a heatmap illustrating vote-shares (X_A). The blank region corresponds to values that do not meet the constraint budget, $0 \leq k_A^* \leq min(n, (\frac{a}{c} - 1))$. Figure (b) shows how allocations to the hub node vary (as a fraction of total allocations $p_{A,1} / \sum_i p_{A,i}$), with change in cost c. The inset shows a closer examination of allocation behaviour at low values of c. Figure (c) shows how the critical cost c_c varies with controller budgets B_A and B_B.

The critical point c_c is obtained as,

$$c_c = \frac{a + b + 2(n+1) - 2\sqrt{(n+1)(a+b+n+1)}}{n+1}. \tag{4.8}$$

In the current setting, we have $N = n + 1 = 1000$ and $B_A = B_B = N$. Using Eq. (4.8) we obtain the critical cost as $c_c \approx 0.54$, also shown in Fig. 4.3. This implies that when

cost of access exceeds half of the per node budget, it is no longer optimal to target peripheral nodes and thus the optimal allocation switches to discrete targeting.

We are also interested to see how c_c changes with controller budgets. Results are shown in Section 4.4 for $N = 1000$ and controller budgets $0.1N \leq B_A \leq 10N$ and $0.1N \leq B_B \leq 10N$. We find that c_c is positively correlated with both controller budgets B_A and B_B, i.e. the optimal allocation strategy takes longer to switch from the continuous to the discrete regime as budgets for both controllers increase.

4.5   Heuristic approaches for arbitrary network structures

In the previous section we presented optimal solutions for the competitive *influence maximisation* problem with cost of access c, using both analytical and numerical methods, in a simplified network structure such as a star graph. Our results reveal a shift in allocation strategy - from continuous to discrete allocations focused on the high-degree hub (irrespective of competitor allocations), as cost of access c is increased. We now ask ourselves how this result extends to other arbitrary, more realistic network structures. Observe that, solving $\Delta_{p_A} X_A = 0$ analytically for large networks with heterogeneous degree distributions can be very challenging. We therefore propose and evaluate heuristic approaches to quantify optimal allocations.

For this purpose, the following heuristic approaches are considered:

1. **Forward Greedy**: We use the traditional greedy method as a benchmark for our proposed heuristics, and modify it slightly for our purpose. More specifically, we exploit the gradient of the vote-share function wherever possible to mitigate the computational challenges of combinatorics used in the traditional greedy method.

   The algorithm begins with an empty set of targeted nodes (i.e. the allocation vector is a zero vector $p_{A,i} = 0, 1 \leq i \leq N$), and then iteratively allocates positive influence to the node that yields the maximum marginal gain in vote-shares, without breaching the total budget B_A —until the vote-share can no longer be improved. It is important to note that, we strictly allocate positive influence to new nodes, and at no point do we re-evaluate the existing set of targeted nodes (or remove allocations from a targeted node).

   At any given time-step t, the node that improves the vote-share maximally is determined by iterating through all nodes with zero allocations ($\forall i$ where $p_{A,i} = 0$). While considering a candidate node i at time step t, a new allocation vector is initialised where resources are uniformly allocated to node i and all nodes targeted at time t. The gradient ascent algorithm described in Chapter 3,

is then used to obtain the optimal allocation configuration $p_A^*(t)$ and the corresponding vote-share X_A^*. The allocation vector that improves the vote-shares maximally is then selected as $p_A^*(t+1)$. This process continues until the vote-share no longer improves.

2. **Backward Greedy**: For the Backward-Greedy approach, we begin with $k_{max} = min(N, (B_A/c) - 1)$ number of nodes targeted at time $t = 0$. Selecting k_{max} nodes to target, where $k_A < N$, is a combinatorial task that can be expensive even for small networks. We get around this problem by exploiting the gradient of the vote-share function. We consider the linear unbudgeted case (where $p_A = \vec{1}$) and determine the gradient at this point [3], and subsequently select the top k_{max} nodes with the highest gradients to initialise our target set of nodes. The algorithm then proceeds by iterating through each node in the target set, and greedily removing one node at every time step, i.e. such that vote-share is increased maximally —until the vote-share can no longer be improved. When iterating through each node i in the target set allocations $p_{A,i}$ to the candidate node i is set to zero. The allocation vector is then normalised to meet the budget constraint, and ultimately optimised using the gradient ascent method.

3. **Lowest vote-share**: We begin with positive allocations to k_{max} nodes in the network. As in Backward-Greedy, we use the gradient to determine which k_{max} nodes to target. Following this, in every iterated step, we remove the node i (set $p_{A,i} = 0$), with the lowest $X_{A,i}$, until the vote-share can no longer be improved. Once allocations to the node is removed at each time step, we normalise the allocation vector and use gradient ascent to optimise the allocation vector.

4. **Lowest allocation**: Following the same approach as in Backward-Greedy and the lowest vote-share methods, we begin with positive allocations to k_{max} nodes and then iteratively remove the node with the lowest allocation $min(p_A)$ until the vote-share can no longer be improved. After removing a node i (setting $p_{A,i} = 0$), we normalise the resulting allocation vector and optimise it using gradient ascent.

5. **Degree-based allocations**: We pick the top k_{max} nodes with the highest degrees in the network and allocate influence in proportion to their degrees, $p_{A,i} \propto deg_i$. The resulting allocation vector is then normalised to meet the budget constraint.

6. **Random**: We randomly select k_{max} nodes from the network and uniformly allocate the available budget $a - (k_{max} \cdot c)$ to the selected nodes.

---

[3]We note that using a uniform allocation vector ($p_A = \vec{1}$) to determine the optimal set of k_{max} nodes works better than any other instance of p_A.

We now examine the performance of our heuristics systematically, first in a star network, then in other synthetic networks such as the random graph and the small world network, and finally in a real-world network.

4.5.1   Analysis of a star network

We first analyse the heuristics in a star network of size $N = 100$ and compare their performance to the analytical solutions obtained in Section 4.4. The size of the network considered here is comparably small as the complexity of greedy methods increase rapidly with network size [4]. For each heuristic method, we present the mean results over 10 simulations in Fig. 4.4.

Fig. 4.4a shows the performance of each heuristic, in terms of final vote-shares obtained. We find that for the most part, sophisticated heuristics like the greedy methods, lowest allocation and lowest vote-share heuristics, outperform naïve approaches like the *degree-based* heuristic and the *random* approach. The Forward-Greedy algorithm consistently yields optimal vote-shares, whereas the Backward-Greedy algorithm performs optimally only in the region $c > 1$. The lowest allocation method performs comparably to the Forward-Greedy method, however the lowest vote-share heuristic yields considerably less vote-shares in comparison. We further note in Fig. 4.4c that the final vote-shares rely more on the number of peripheral nodes k_A targeted, and less on the resource distribution over the targeted nodes.

In terms of time-complexity (as shown in Fig. 4.4b), we find that greedy methods are always more expensive than the lowest vote-share and lowest allocation methods. This is expected as greedy methods are required to iterate through all possible candidate nodes when updating the target set of nodes and thus they have a complexity of $O(\frac{N^5}{\mu})$. The lowest vote-share and lowest allocation methods on the other hand have a complexity of $O(\frac{N^4}{\mu})$ given that the time complexity for the gradient ascent method to converge to a μ approximation of the optimal solution is $O(\frac{N^3}{\mu})$ and the complexity of iteratively removing nodes with the lowest allocation or vote-shares is $O(N)$.

For completion, we also examine the instance where the competitor allocates all its budget to the hub node. We explore this setting specifically to determine how change in competitor allocations affects the performance of our heuristics. Results are summarised in Fig. 4.5.

---

[4] The complexity of the greedy heuristics (forward and backward) in Section 4.5 is given by $O(\frac{N^5}{\mu})$. The complexity for greedily removing or adding nodes to the solution set is $O(N^2)$. Additionally, the complexity of the gradient ascent method can be derived as $O(\frac{N^3}{\mu})$ (as shown in Section 3.5.1), where μ is the approximation-factor used to terminate the algorithm
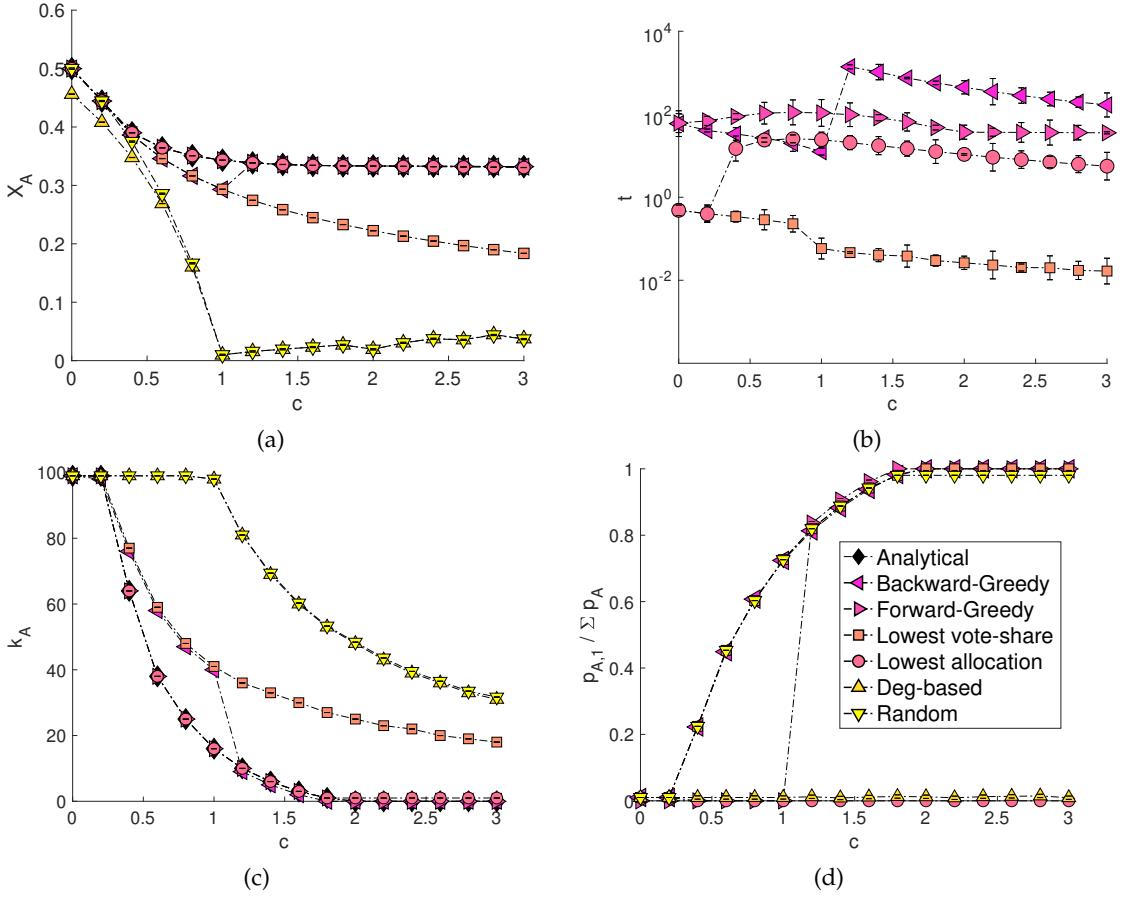
FIGURE 4.4: Figure showing the performance of various heuristics in the star network with $N = 100$ nodes. Competitor allocations are uniform across all nodes. Figure (a) illustrates the final vote-share (X_A) obtained by each heuristic as c is varied. The time complexity of the heuristics (measured in seconds) is shown in Figure (b). Figure (c) and (d) illustrate the configuration of allocations obtained by the heuristics. The number of targeted peripheral nodes k_A is shown in (C), while the allocations to the hub node as a fraction of total allocations ($p_{A,1} / \sum_i p_{A,i}$) is shown in (d). Results are averaged over 100 realisations of the random heuristic, and errorbars show the 95% confidence interval. All other heuristic approaches are deterministic and hence we only plot the mean values.

Fig. 4.5a illustrates the performance of the heuristics. We observe similar trends as before, i.e. the more advanced methods significantly outperform crude heuristics (degree-based and random), especially as cost of access increases. However, contrary to what we observed in Fig. 4.4a, here we find that the lowest-allocation method performs significantly worse than before, as do the other heuristics (Backward-Greedy and lowest vote-share) that use the gradient method to initialise their allocation vectors (with positive allocations to k_{max} nodes). This happens because peripheral nodes are prioritised when the competitor focuses on the hub node, and thus allocations to the hub node are set to 0 when initialising a vector with allocations to $k_{max} < n$ nodes. However, analytical results show that, irrespective of controller allocations, resources should be diverted to the hub node as cost of access is increased. Therefore methods that do not allocate any resources to the hub node when
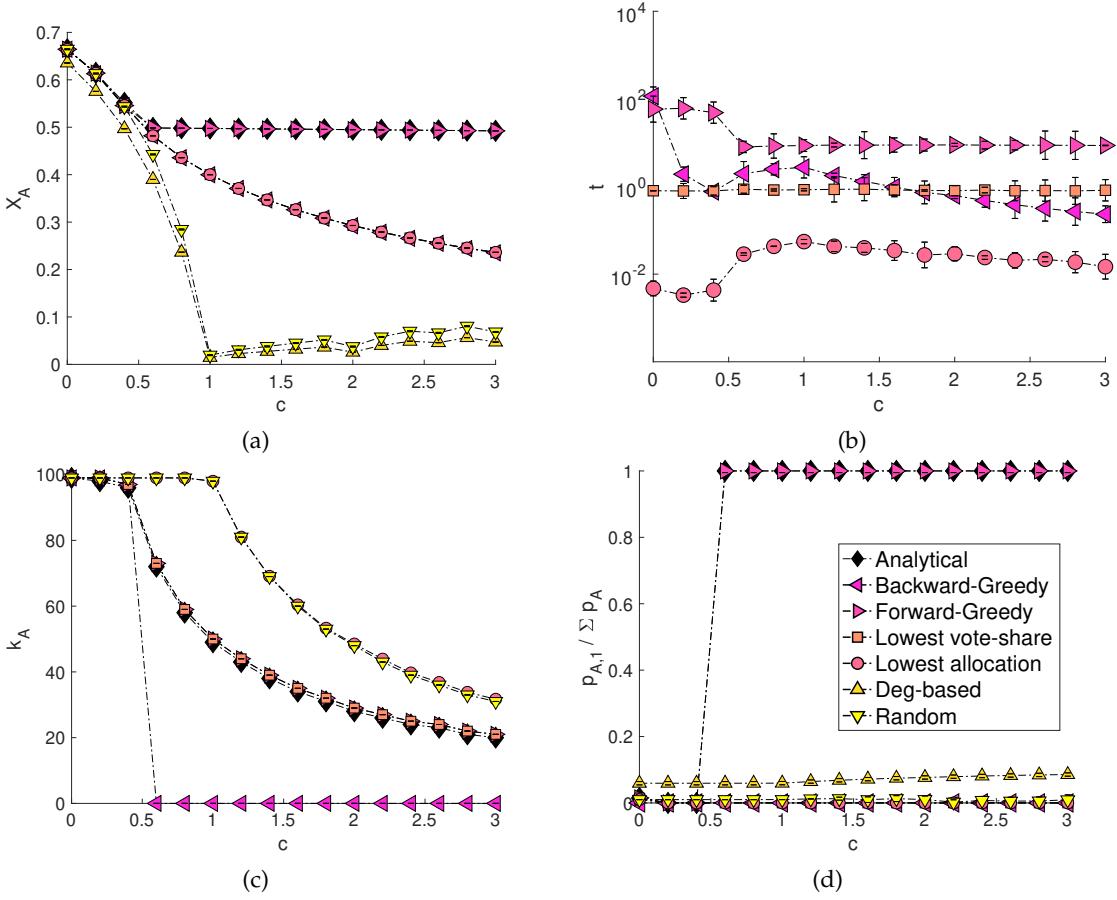
FIGURE 4.5: Figure showing the performance of various heuristics in the star network with $N = 100$ nodes. The competitor targets the hub node discretely. Figure (a) illustrates the final vote-share (X_A) obtained by each heuristic as c is varied. The time complexity of the heuristics (measured in seconds) is shown in Figure (b). Figure (c) and (d) illustrate the configuration of allocations obtained by the heuristics. The number of targeted peripheral nodes k_A is shown in (C), while the allocations to the hub node as a fraction of total allocations ($p_{A,1} / \sum_i p_{A,i}$) is shown in (d). For the random heuristic, results are averaged over 100 realisations and errorbars indicate the 95% confidence interval. All other heuristic approaches are deterministic and hence we only plot the mean results.

initialising the allocation vector get stuck at a local maxima.

4.5.2   Performance of the heuristics in synthetic networks

To generalise our results further, we examine how the heuristics perform in other arbitrary network structures. For this purpose, we explore random (ER) graphs and small-world (SW) graphs that closely resemble several real world networks (Newman (2000)).

Here networks are of size $N = 10$ with average degree $\langle k \rangle = 4$. We consider small networks since analytical solutions are difficult to obtain in network structures with high degree heterogeneity, and thus we rely on combinatorial enumeration to

determine optimal allocations, which is computationally expensive (Aigner and Axler (2007)). More specifically, this is a brute-force technique, where optimal allocations are determined by first enumerating all possible combinations of nodes that can be targeted without breaching the budget, which is followed by using the gradient ascent method to determine the optimal distribution of resources over every combination of nodes. The allocation vector yielding the highest vote-shares is then the optimal solution.

We present our findings in Fig. 4.6. Here competitor B targets the network uniformly. Results are averaged over 100 realisations, where each heuristic is run 10 times over 10 instances of random and small-world graphs. Heuristic methods used are the same as before (as described in Section 4.5).

(a)

(b)

(c)

(d)
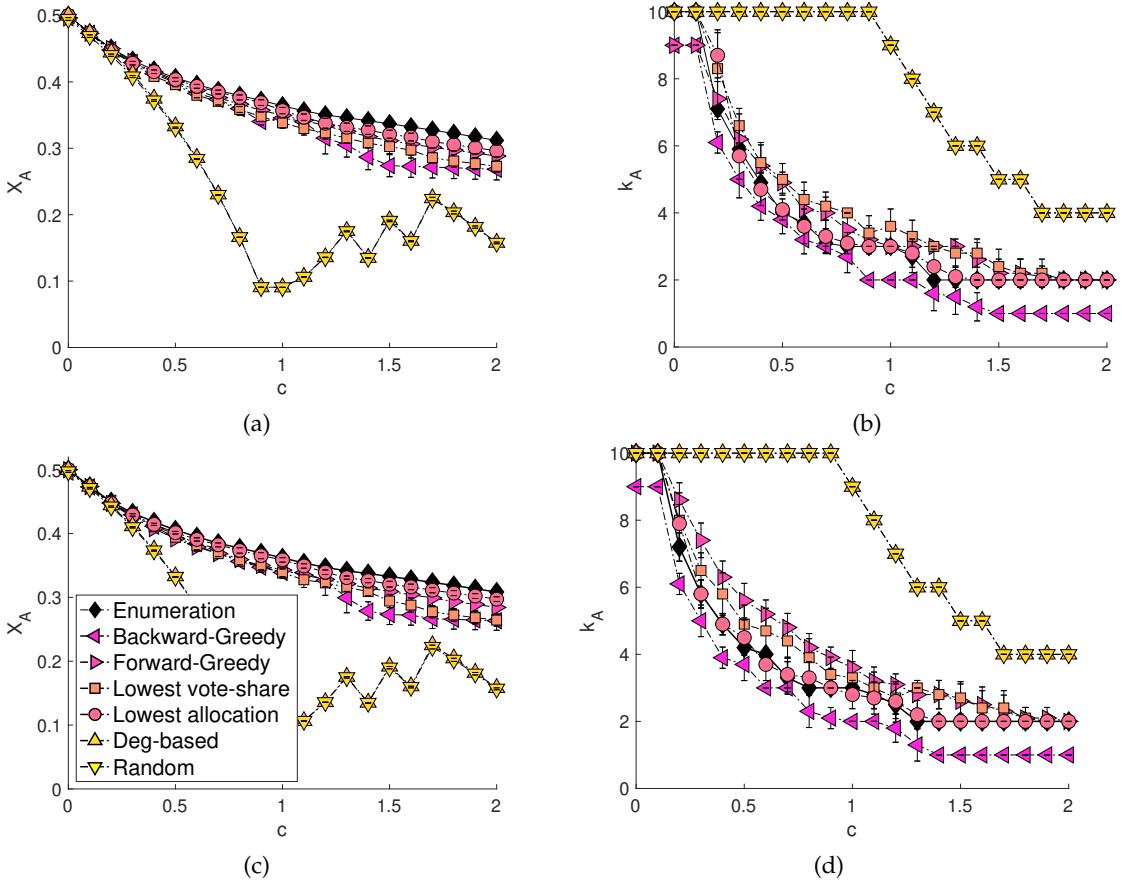
FIGURE 4.6: Figure showing the performance of the proposed heuristics in random (ER) graphs and small-world (SW) graphs. Figure (a) shows the final vote-shares X_A obtained by each heuristic method, and Figure (b) illustrates the optimal number of nodes targeted k_A in the ER networks. Similarly, Figure (c) illustrates the performance of the heuristics in the SW networks and Figure (d) shows the corresponding optimal number of targeted nodes k_A. The approximation-factor used to terminate the gradient ascent method is $\mu = 10^{-7}$. Here errorbars show the 95% confidence interval.

Firstly, in Fig. 4.6 we find that the performance of the heuristics are independent of network structures, i.e. results follow similar trends on both random (ER) graphs as well as small-world (SW) networks. Observe that, the lowest-allocation heuristic

consistently yields the highest vote-shares, closely followed by the Forward-Greedy approach. The Backward-Greedy algorithm and the lowest vote-share method are in comparison much less effective in optimising vote-shares on each network. We also find in Figs. 4.6a and 4.6c that vote-shares for the random and degree-based methods appear noisy in the region $c > 1$. The decrease in vote-shares coincide with the plateaus in Figs. 4.6b and 4.6d, where the number of targeted nodes remains fixed, despite the increase in cost c, and thus vote-share X_A sharply decreases in such an instance.

4.5.3   Performance of heuristics in a real world network

Next, to examine how these heuristics perform in the real world, we explore the heuristic methods in a real-world network depicting collaboration and co-authorship among researchers in the field of network science (Rossi and Ahmed (2015); Guimera et al. (2003)). The nodes represent network scientists and an edge between any two nodes indicates co-authorship. The network is of size $N = 379$ with an average degree of $\langle k \rangle = 4$.

Given that the network is quite large, it is infeasible to use methods such as enumeration to obtain optimal solutions. Thus instead of focusing on how well heuristics perform in these networks, here we determine how much vote-share a controller can gain from employing the proposed heuristics in the real world, as opposed to following naïve approaches such as random or degree-based methods. As greedy methods are expensive to run on large networks, here we only consider the Forward-Greedy algorithm. Additionally, for comparison we introduce a new heuristic *lowest-allocation* $(\Delta = 10)$, which is a modified version of the lowest-allocation heuristic, where the algorithm removes 10 nodes with lowest allocations, instead of removing just one node at every time step. This heuristic has a lower time-complexity, and helps us determine how much vote-share a controller has to sacrifice for faster results.

From Fig. 4.7, we find that the lowest-allocation method consistently yields the highest vote-shares, while significantly outperforming the random and degree-based heuristics. In addition, from Fig. 4.7b, we find that modifying the lowest allocation heuristic reduces the time complexity considerably while affecting the performance of the algorithm only marginally, and therefore can be used alternatively for larger networks.

To gain more insight into our results, we examine the patterns of allocations in more detail. Recall that in star networks, resources were diverted to the high-degree hub node with increase in the cost of access. We now examine if this result is still true for larger, arbitrary networks.
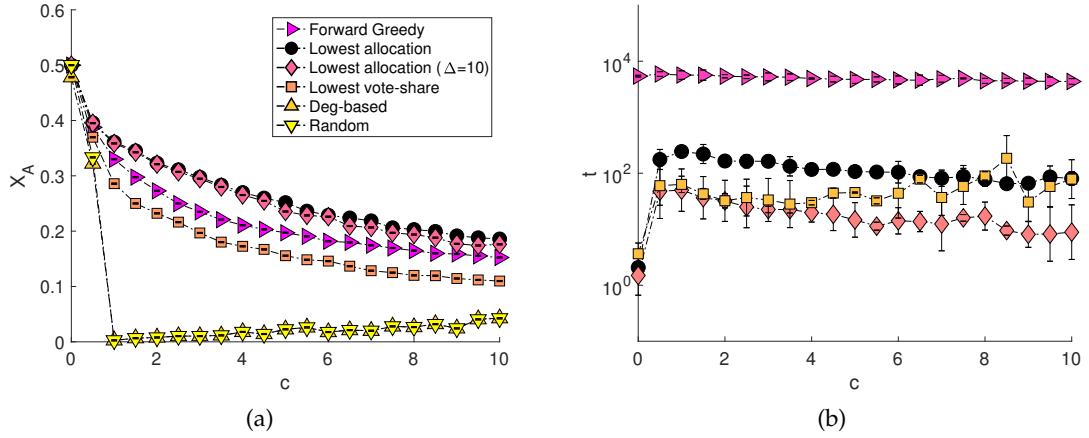
FIGURE 4.7: Figure showing the performance of the proposed heuristics in a real-world network of research collaborations within the network science community ($N = 379$ and $\langle k \rangle = 4$). Results are averaged over 10 simulations for each heuristic. Figure (a) shows the final vote-shares X_A obtained, while Figure (b) illustrates the average time taken (in seconds) by the heuristic approaches. Errorbars indicate the 95% confidence interval. An approximation factor of $\mu = 10^{-7}$ is used to terminate the gradient ascent algorithm.
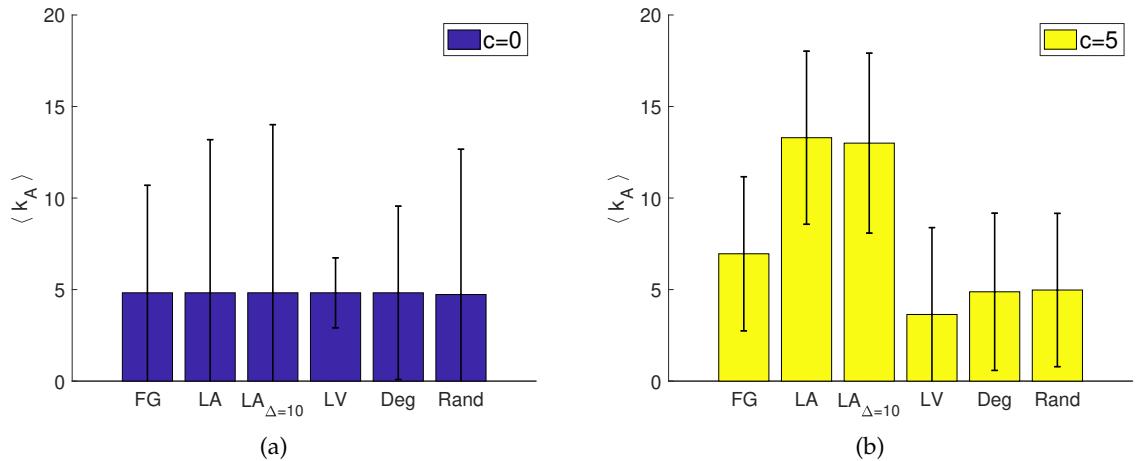
FIGURE 4.8: Figure showing the optimal allocation patterns of heuristics for multiple costs of access in a real-world network. Figure (a) shows the average degree of targeted nodes $\langle k_A \rangle$ as per the final optimal allocation vector given by each heuristic when cost of access $c = 0$. Figure (b) shows the average degree of targeted nodes $\langle k_A \rangle$ of each heuristic when cost of access $c = 5$. Errorbars represent standard errors.

To do this we examine the average degree of targeted nodes in the following *two* settings[5] (illustrated in Fig. 4.8): (i) no cost of access $c = 0$, and (ii) high cost of access $c = 5$, and we find that heuristics that preferentially allocate resources to high degree nodes for high cost of access yield higher vote-shares than heuristics that focus on low degree nodes. Thus generalising our results obtained in star networks to real-world settings.

---

[5]More settings examined in Appendix B.2.
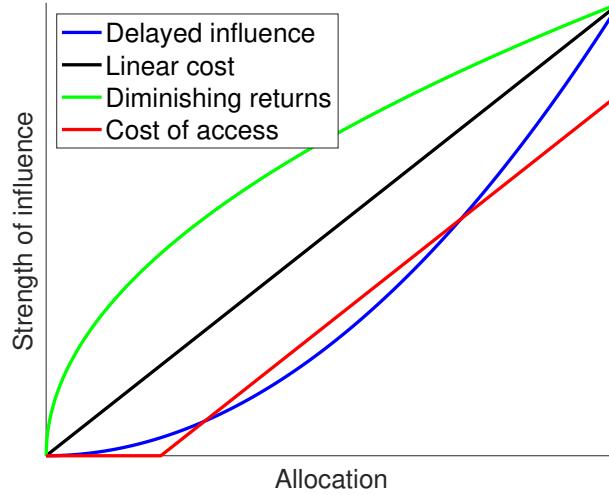
FIGURE 4.9: Figure showing the various cost functions we consider in our model. Linear cost here illustrates the setting examined in Chapter 3 where the strength of influence experienced by an individual is directly proportional to the amount of budget allocated to them. The piecewise-linear function illustrates the cost of access setting where individuals experience influence only after a certain threshold c is breached. The nonlinear cost functions correspond to the diminishing returns setting and the delayed influence setting (described in the text).

4.6   Nonlinear cost functions

So far in this chapter, we examine the competitive *influence maximisation* problem under the assumption of a fixed cost of access paid by the controller to influence any node in the network. We further extend our work in this direction by generalising the cost function to include other nonlinear relations between the cost and effect of influence on individuals. More specifically, we vary the cost of allocation for the controller nonlinearly (see Fig. 4.9) such that the allocation vector p_A is now constrained as $\sum_i p_{A,i}^\gamma = B_A$.

Following this, the optimisation problem can now be written as,

$$p_A^* = \arg\max_{p_A \in \mathcal{P}} X_A^*(L, p_B), \tag{4.9}$$

where $\mathcal{P}$ is a set of all possible allocations subject to the constraint $\sum_i p_{A,i}^\gamma = B_A$.

Depending on the value of γ, the above relation yields two settings: (i) where $\gamma > 1$, we observe a *diminishing returns* in the strength of influence experienced by nodes with increase in allocations, and (ii) where $0 < \gamma < 1$, we find that individuals take longer (or more resources) to respond to external influence and hence this results in a *delayed influence* effect.

4.6.1   Case $\gamma > 1$: Diminishing returns

We first consider the instance where $\gamma > 1$. The strength of influence experienced by a node in this case, is given by the γ-th root of allocated resources $p_{A,i}^{\gamma}$. Mathematically, the allocation vector is constrained by the budget as $\sum_i p_{A,i}^{\gamma} = B_A$, i.e. the norm of the allocation vector is constrained by $B_A^{1/\gamma}$, which implies that the constraint set here is convex. Given that the vote-share X_A is a concave function of allocations p_A (see Appendix A.3), the problem we have at hand is a convex optimisation problem.

A common approach taken to solve this type of constrained optimisation problem uses the Lagrange method (Boyd et al. (2004)), where the objective function is maximised by iteratively stepping in the optimal direction, within the constraint space. To employ this method in our work we follow the approach discussed in Tanaka and Aoyagi (2008) which yields the optimal direction as [6],

$$p_{A,i}(t+1) \leftarrow p_{A,i}(t) + \eta \frac{\nabla_{p_A(t)} X_{A,i}(t)}{p_{A,i}^{\gamma-2}(t)}, \tag{4.10}$$

where η is the step-length.

We therefore first initialise a random, feasible allocation vector and then use Eq. (4.10) to iteratively update the allocations until the total vote-share can no longer be improved, or alternatively we obtain a $\mu-$approximation of the optimal allocation configuration. Note that the allocation vector is also normalised at every time step to satisfy the budget constraint $\sum_i p_{A,i}^{\gamma} = B_A$. The step-length is adjusted using backtracking line search to ensure convergence (Boyd et al. (2004)), and given that the problem at hand is a convex optimisation problem, the algorithm is guaranteed to converge at the global maximum.

4.6.2   Case $0 < \gamma < 1$: Delayed influence

We now consider the setting where $0 < \gamma < 1$. The constraint set in this case is clearly nonconvex which makes the optimisation problem more challenging than before. This also means that the Lagrange method can no longer be applied to solve the optimisation problem.

In general, nonconvex optimisation problems are difficult to solve. However in some cases, the structure of the objective function and the constraint function can be exploited to design polynomial-time algorithms that obtain near-optimal solutions (Jain et al. (2017)). Given that our objective function is convex and the constraint space

---
[6]More details of the method can be found in the Appendix B.3.

is an ℓ_p-norm ball where $0 < p < 1$, we employ the projected gradient ascent algorithm discussed in Chapter 3, and modify the projection step to meet the nonlinear budget constraint. For the projection method, we use the IRBP algorithm which is an instance of a majorisation-minimisation algorithm (Lange (2016)), where the algorithm iteratively alternates between a majorisation step and a minimisation step until it converges (Yang et al. (2021)). In our case, the majorisation step first relaxes and linearises the ℓ_p ball to obtain a weighted ℓ_1 ball, followed by the minimisation step which obtains the projection of the point on the ℓ_1 ball. As this is a nonconvex problem, we do not have any theoretical guarantees of reaching the optimal solution, and therefore we run the algorithm for multiple initialisations of the allocation vector and consider the mean result obtained over all simulations [7].

4.6.3   Results

We first analyse the problem in synthetic networks where optimal allocations are determined analytically and compared to the results obtained using the above numerical approaches. We then examine the problem numerically in a real-world network and discuss our observations.

4.6.3.1   Mean-field approximation

To obtain a benchmark for our numerical results, we propose an analytical method that determines optimal allocations in simplified network structures, such as a core-periphery network. A core-periphery network usually has a core of highly connected nodes and other sparsely connected peripheral nodes. The bimodal degree-distribution resembles many real-world leader-follower type networks, and the simplified network structure limits the degrees of freedom in the problem which allows us to employ analytical methods to determine optimal solutions.

To determine optimal allocations analytically, here we use a degree-based mean-field approximation method. The degree-based mean-field approximation assumes that nodes with the same degree have the same behaviour (or state) and therefore groups nodes based on their degrees. The behaviour of the population is then approximated by averaging over the behaviours of each group of nodes. Although such an approximation is not always effective, it has been observed to work well in networks where there are no degree correlations (Pastor-Satorras and Vespignani (2001)).

We obtain a degree-based mean-field approximation by following the approach taken in Romero Moreno et al. (2021.), and modifying it to reflect our budget constraint as follows,

---

[7]For uniformity, we also run multiple simulations for the convex case where $\gamma > 1$ and show the average results

$$X_{MF} = \frac{\left(\sum_k \frac{P_k k}{k+a_k^{\frac{1}{\gamma}}+b_k^{\frac{1}{\gamma}}}\right)\left(\sum_k \frac{P_k k a_k^{\frac{1}{\gamma}}}{k+a_k^{\frac{1}{\gamma}}+b_k^{\frac{1}{\gamma}}}\right)}{\sum_k \frac{P_k k(a_k^{\frac{1}{\gamma}}+b_k^{\frac{1}{\gamma}})}{k+a_k^{\frac{1}{\gamma}}+b_k^{\frac{1}{\gamma}}}} + \sum_k \frac{P_k a_k^{\frac{1}{\gamma}}}{k+a_k^{\frac{1}{\gamma}}+b_k^{\frac{1}{\gamma}}}. \tag{4.11}$$

Here P_k is the degree-distribution of the network and k is the degree of nodes in the network. Additionally, b_k is the competitor's allocation to the group of nodes with degree k and a_k is the controller's allocation to the same group of nodes. Therefore, the influence experienced by the nodes in the group from both controllers are $a_k^{1/\gamma}$ and $b_k^{1/\gamma}$. Note that as node behaviours are considered to be the same within any group, allocations from both controllers are also assumed to be uniform across all nodes in the group. We can now use Eq. (4.11) to determine the optimal allocation patterns in any large, arbitrary core-periphery network structure.

4.6.3.2 Analysis of synthetic networks

Core-periphery networks of size $N = 1000$ are considered, with a core formed by $P_1 = p_r = 0.25$ (or 25% of the total nodes in the network). Nodes in the highly clustered core have a degree of $k_1 = 30$ and the sparsely connected peripheral nodes have a degree $k_2 = 3$. We examine two settings, one where B targets the hub nodes and another where they target the peripheral nodes.

Competitor allocations to the core

We first consider the instance where B targets the hub nodes. In this case the competitor budget is distributed uniformly over all nodes in the core of the network ($b_{k_1} = \frac{B_B}{p_r N}$ and $b_{k_2} = 0$). We consider ϵ_A as the fraction of total budget allocated to hub nodes by controller A. Using Eq. (4.11) we obtain,

$$X_{MF} = \left(\frac{p_r k_1}{k_1 + \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}} + b_{k_1}^{\frac{1}{\gamma}}} + \frac{(1-p_r)k_2}{k_2 + \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}} + b_{k_2}^{\frac{1}{\gamma}}} \right) \cdot$$

$$\left(\frac{p_r k_1 \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}}}{k_1 + \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}} + b_{k_1}^{\frac{1}{\gamma}}} + \frac{(1-p_r)k_2 \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}}}{k_2 + \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}} + b_{k_2}^{\frac{1}{\gamma}}} \right) \cdot$$

$$\left(\frac{p_r k_1 \left(\left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}} + b_{k_1}^{\frac{1}{\gamma}} \right)}{k_1 + \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}} + b_{k_1}^{\frac{1}{\gamma}}} + \frac{(1-p_r)k_2 \left(\left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}} + b_{k_2}^{\frac{1}{\gamma}} \right)}{k_2 + \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}} + b_{k_2}^{\frac{1}{\gamma}}} \right)^{-1}$$

$$+ \left(\frac{p_r \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}}}{k_1 + \left(\frac{\epsilon B_A}{p_r N}\right)^{\frac{1}{\gamma}} + b_{k_1}^{\frac{1}{\gamma}}} + \frac{(1-p_r) \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}}}{k_2 + \left(\frac{(1-\epsilon)B_A}{(1-p_r)N}\right)^{\frac{1}{\gamma}} + b_{k_2}^{\frac{1}{\gamma}}} \right) \cdot$$

We now use semi-analytical methods to determine the optimal allocation ϵ_A^* using the above expression that maximises vote-shares X_A^*. Additionally, we use methods described in Sections 4.6.1 and 4.6.2 to determine optimal allocations numerically. For numerical results, we consider 10 instances of core-periphery networks, each of size $N = 1000$, $p_r = 0.25$, $k_1 = 30$ and $k_2 = 3$. Networks are generated using the configuration model (Newman (2018)). We consider three budget scenarios: (i) insufficient budget $B_A/B_B = 0.1$, (ii) equal budget $B_A/B_B = 1$ and finally (iii) excess budget $B_A/B_B = 10$. In each case, $B_B = N$. Fig. 4.10 shows optimal allocations and vote-shares, against competitor allocations to the hub nodes averaged over 10 simulations.

Fig. 4.10a illustrates the fraction of total resources allocated to the hub nodes in the network, as γ is varied. Fig. 4.10b shows the corresponding vote-shares obtained by the controller. Immediately, we see a stark difference in optimal allocations and vote-shares between the two regimes: (i) $0 < \gamma < 1$ and (ii) $\gamma > 1$. We find that allocation strategies are highly sensitive to controller budgets when $0 < \gamma < 1$. Whereas when $\gamma > 1$, optimal strategies are roughly the same even when budgets differ considerably. A similar phenomenon is also reflected in Fig. 4.10b, where controller budgets significantly affect vote-shares for $0 < \gamma < 1$, and have little to no effect on vote-shares when $\gamma > 1$. The linear case, $\gamma = 1$, clearly acts the transition point between the two regimes.

Taking a closer look at the region where $0 < \gamma < 1$, we find that optimal allocations oscillate between discrete and continuous configurations, particularly under limited budget, as γ changes. These fluctuations are more abrupt in analytical results, which
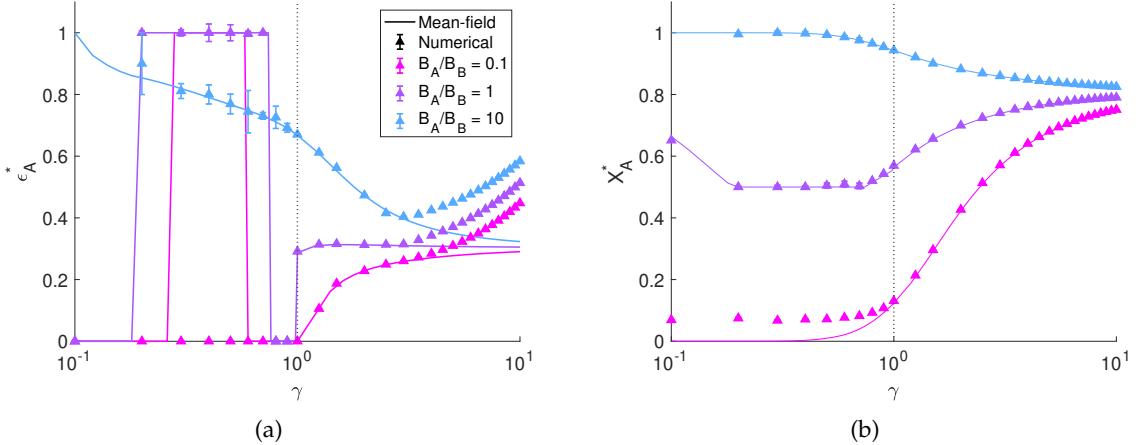
FIGURE 4.10: Figure showing analytical and numerical results for optimal configurations of allocations in a core-periphery network of size $N = 1000$ and average degree $\langle k \rangle = 9.75$. Analytical results are obtained using the mean-field approach shown in Eq. (4.11), while numerical results are obtained using approaches discussed in Sections 4.6.1 and 4.6.2. Competitor B targets hub nodes with degree $k_1 = 30$, which make up for 25% of the total nodes. The remaining nodes all have degree $k_2 = 3$. Figure (a) shows the optimal fraction ϵ_A^* of the total budget allocated to hub nodes as γ is varied. Figure (b) shows the corresponding optimal vote-shares X_A^* obtained by the controller. Numerical results obtained are a $\mu-$approximation of the optimal solutions where $\mu = 10^{-10}$. Results are averaged over 100 simulations, where algorithms for both the convex and nonconvex optimisation problem are run 10 times with random initialisations on 10 realisations of core-periphery networks. Errorbars show the 95% confidence intervals. The result for $\gamma = 0.1$ and $B_A = 10N$ is missing due to runtime errors caused by numerical overflows.

maybe due to the *all-or-none* strategy adopted in the mean-field approximation, i.e. all nodes with the same degree are uniformly targeted.

This assumption also results in discrepancies in the analytical and the numerical results. For instance, where $B_A/B_B = 0.1$ we find that while the analytical solution opts for a discrete strategy (targets only the hub nodes) when $0.3 \leq \gamma \leq 0.5$, the numerical solution does not allocate any resources to the hub nodes in these settings and only targets a fraction of the peripheral nodes. Whereas from Fig. 4.10b we find that the numerical method yields higher vote-shares than the analytical results in this region, thus highlighting the limitations of the mean-field approximation in such instances. Finally, we also observe inconsistencies in Fig. 4.10a for very high values of γ, i.e. $\gamma > 4$ likely caused by numerical instabilities (Barbero and Sra (2014)).

Competitor allocations to the periphery

We now consider the instance where competitor B targets the periphery, i.e. $b_{k_1} = 0$ and $b_{k_2} = \frac{B_B}{(1-p_r)N}$, and illustrate our results in Fig. 4.11. Once again, we find that controller budgets heavily impact optimal allocations and vote-shares in the region where $0 < \gamma < 1$. The variation in allocations decrease as γ crosses over into the region where $\gamma \geq 1$. We also find that, for the most part numerical results closely replicate analytical results, with some exceptions. As before, numerical inconsistencies
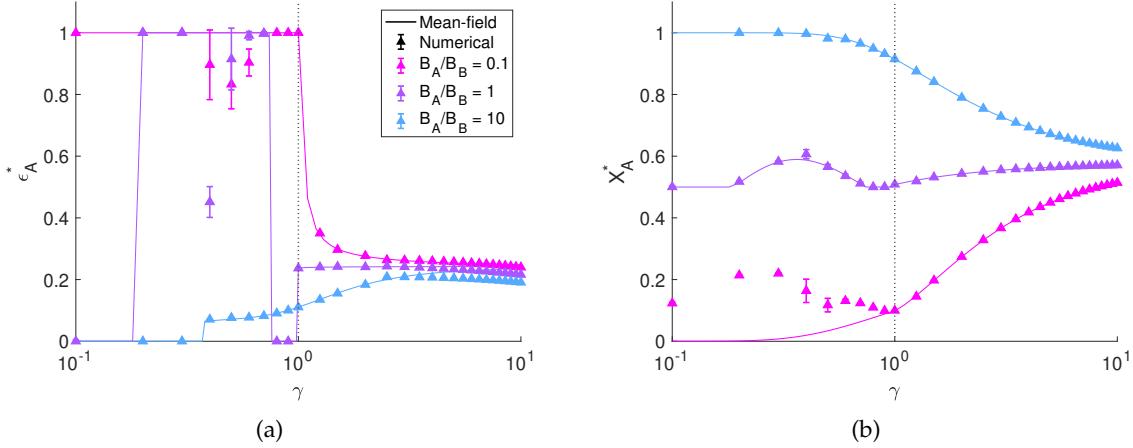
FIGURE 4.11: Figure showing analytical and numerical results for optimal configurations of allocations in a core-periphery network of size $N = 1000$ and average degree $\langle k \rangle = 9.75$. Analytical results are obtained using the mean-field approach shown in Eq. (4.11), while numerical results are obtained using approaches discussed in Sections 4.6.1 and 4.6.2. Competitor B targets peripheral nodes of degree $k_2 = 10k_1$, uniformly. Hub nodes here have degree $k_1 = 30$. Figure (a) shows the optimal fraction ϵ_A^* of the total budget allocated to hub nodes as γ is varied. Figure (b) shows the corresponding optimal vote-shares X_A^* obtained by the controller. Numerical results obtained are a $\mu-$approximation of the optimal solutions where $\mu = 10^{-10}$. Results are averaged over 100 simulations, where algorithms for both the convex and nonconvex optimisation problem are run 10 times with random initialisations on 10 realisations of core-periphery networks. Errorbars show the 95% confidence intervals. The result for $\gamma = 0.1$ and $B_A = 10N$ is missing due to runtime errors caused by numerical overflows.

are observed when $\gamma > 4$. We also observe discrepancies between the analytical and the numerical results in the region $0.4 \leq \gamma \leq 0.6$, and argue that such disparities exist as allocations are artificially constrained in the analytical method, whereas numerical approaches are more flexible, thus yielding higher vote-shares.

4.6.3.3   Analysis of a real world network

Next, to explore how these results apply to the real world, we explore a collaboration network among network scientists, also discussed in Section 4.5.3, consisting of $N = 379$ scientists connected through coauthorship of papers.

We consider two instances of competitor allocations: (i) B targets hub nodes, and (ii) B targets peripheral nodes. Identifying hub nodes in a heterogeneous network is not a straight-forward problem, but here for simplicity, we use the degree centrality measure to distinguish between hubs and peripheral nodes. Given that the average degree of the network is $\langle k \rangle = 4.8$, we assume nodes with degrees above $k > 5$ are hubs, and those with $k \leq 5$ form the periphery of the network. We find that this method classifies nearly 30% of the network as hubs and the rest as the periphery.
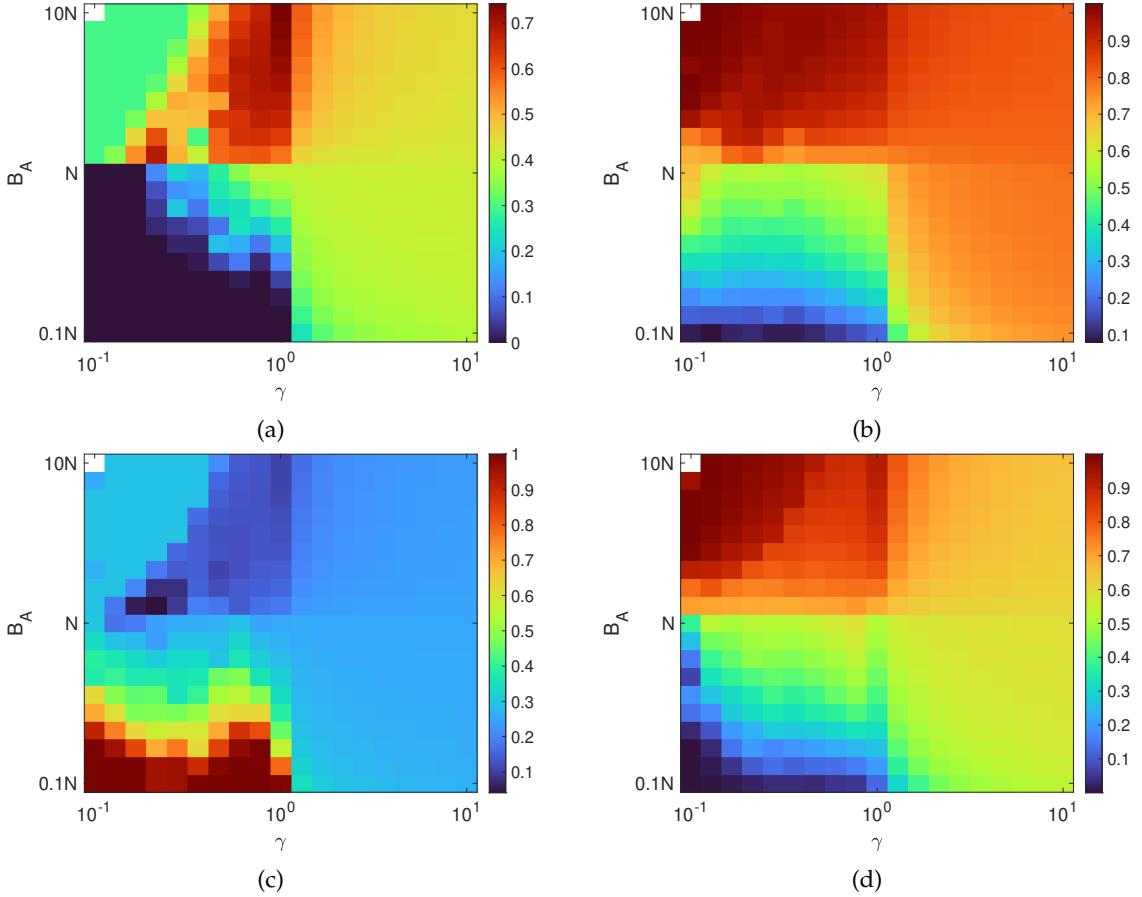
FIGURE 4.12: Figure showing optimal allocations and vote-shares in a real-world network depicting collaborations among network scientists. The network is of size $N = 379$ and average degree $\langle k \rangle = 4.8$. The top panel (Figures (a) and (b)) shows results for the instance where the competitor B targets hubs nodes. Figure (a) illustrates the optimal fraction of total budget that is allocated to the hub nodes, and Figure (b) shows the corresponding optimal vote-shares for varying γ and B_A. The bottom panel (Figures (c) and (d)) illustrates results for the instance where competitor B targets the periphery of the network. Figure (c) shows the optimal fraction of budget allocated to hub nodes and Figure (d) shows the corresponding optimal vote-shares for varying γ and B_A. Numerical results obtained are a $\mu-$approximation of the optimal solutions where $\mu = 10^{-10}$. Results are averaged over 10 simulations, where algorithms for both the convex and nonconvex optimisation problems are run 10 times with random initialisations. The missing values for $\gamma = 0.1$ and $B_A = 10N$ across all four heatmaps are due to numerical overflows.

As it is infeasible to apply analytical methods to highly heterogeneous network structures, here we rely on numerical approaches to determine optimal allocations. We systematically examine each setting of competitor allocations and determine optimal allocations, as the nonlinearity constraint of the allocation vector changes between $0.1 \leq \gamma \leq 10$ and the controller budget is varied between $0.1 \leq B_A/B_B \leq 10$, where $B_B = N$. Results are averaged over 10 simulations, and shown in Fig. 4.12.

Figs. 4.12a and 4.12c illustrate the optimal fraction of the total budget allocated to the hub nodes, against competitor allocations to the hub nodes and to the peripheral

nodes respectively. Once again, in both instances we observe that optimal allocation patterns are highly sensitive to change in budget conditions in the region $0 < \gamma < 1$, whereas optimal allocation patterns are more uniform when $\gamma > 1$. Furthermore, we observe that under low budget conditions, optimal allocation configurations avoid nodes targeted by the competitor and focus more resources on nodes avoided by the competitor. This trend changes when the budget increases $B_A > N$, and more resources are allocated to the nodes targeted by the competitor. For instance, observe in Fig. 4.12a, where the competitor targets the hub of the network, optimal allocations under low budget conditions $B_A \le N$ are focused more on the periphery and less on the hub nodes. However, as the budget B_A increases, more resources are diverted to the hub nodes. Recall from Chapter 3 that this is the same trend we observed in the linear case.

Figs. 4.12b and 4.12d illustrate the optimal vote-shares obtained against a competitor targeting the hub and the periphery respectively. Similar to results in Section 4.6.3.2, we find that vote-shares vary significantly with budget settings when $0 < \gamma < 1$ for both competitor allocations, but not as much when $\gamma \ge 1$.

So far we have observed patterns of optimal allocations in settings with nonlinear budget constraints. We now determine how much vote-share a controller would gain from optimally targeting a network as opposed to employing a naïve strategy. For comparison, we consider two simple heuristics: (i) the degree-based approach and (ii) uniform allocation.

Once again we consider the collaboration network among network scientists for our simulations. We vary γ and B_A, and in each instance determine the optimal allocation and vote-shares numerically, using approaches defined in Sections 4.6.1 and 4.6.2. For comparison, we define the allocation vector for the degree-based approach as $p_{A,i} \propto k_i$, where k_i is the degree of a node i, normalised to meet the budget constraint. Additionally, the allocation vector for uniform allocation is given by $p_{A,i} = (B_A/N)^{1/\gamma}, 1 \le i \le N$. The corresponding vote-shares are determined for each case as X_A^{deg} and X_A^{uni} respectively. Gain in vote-shares is then measured as $[X_A^* / X_A^{deg} - 1]$ and $[X_A^* / X_A^{uni} - 1]$, and results are shown in Fig. 4.13.

While Figs. 4.13a and 4.13b illustrate gain in vote-shares against competitor allocations to the hub, Figs. 4.13c and 4.13d show gain in vote-shares against competitor allocations to the periphery. In both cases, we find that the gain in vote-shares show similar patterns. The sensitivity of gain in vote-shares to controller budget is more in the *delayed influence* setting, as compared to the *diminishing returns* setting. In particular, we observe that the controller can gain significantly ($\approx 10^{10}$ times) by targeting the network optimally for low values of γ and a low budget B_A. This implies that the effectiveness of the optimal strategy in comparison to other heuristics is higher when individuals take longer (or more resources) to experience or respond to
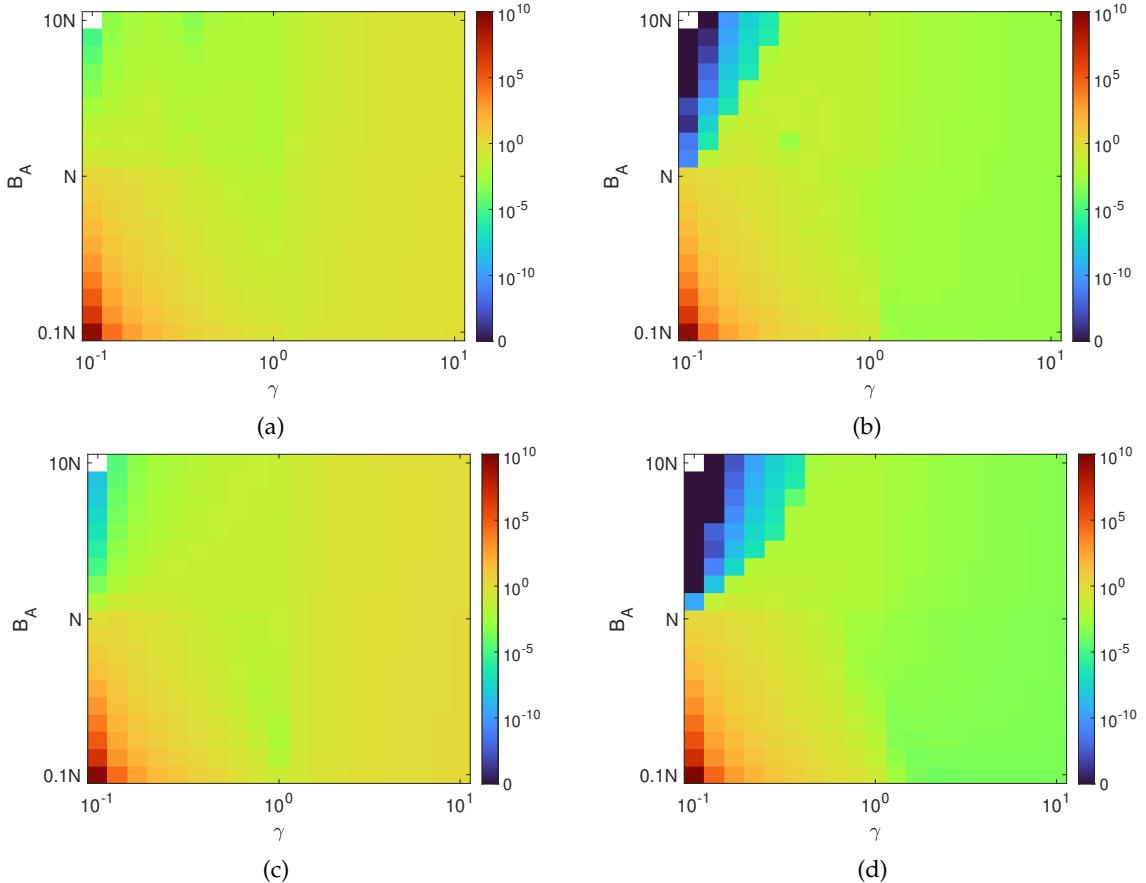
FIGURE 4.13: Figures showing the gain in vote-shares obtained when employing optimal allocation strategies in a real-world collaboration network of size $N = 379$ and average degree $\langle k \rangle = 4.8$, as opposed to simple heuristics such as (i) degree-based targeting or (ii) uniform allocations. We consider two instances of competitor allocations. The top panel consisting of Figures (a) and (b) illustrate results for the setting where the competitor targets the hub nodes. Figure (a) shows the gain in vote-shares $[X_A^* / X_A^{deg} - 1]$ when the optimal strategy is compared to degree-based targeting, whereas Figure (b) shows the gain in vote-shares $[X_A^* / X_A^{uni} - 1]$ when the optimal strategy is compared to uniform allocations. The bottom panel shows results for the instance where the competitor targets the periphery. Figure (c) compares the optimal strategy to degree-based targeting $[X_A^* / X_A^{deg} - 1]$ and Figure (d) compares the optimal strategy to uniform allocations $[X_A^* / X_A^{uni} - 1]$. Numerical results obtained are a $\mu-$approximation of the optimal solutions where $\mu = 10^{-10}$. Results shown are mean values obtained over 10 simulations, where algorithms for both the convex and nonconvex optimisation problems are run 10 times with random initialisations. The missing values for $\gamma = 0.1$ and $B_A = 10N$ across all *four* heatmaps are due to numerical overflows.

external influence under low budget conditions. We further observe that for larger budgets and low γ, the optimal strategy, irrespective of competitor allocations is to target the network uniformly.

4.7   Summary and conclusions

In this chapter, we explore the competitive *influence maximisation* problem under nonlinear cost of influence. Traditionally, the *influence maximisation* problem has been studied in a linear setting where the cost of influence (or allocations) is analogous to the strength of influence experienced by individuals, and thus the relationship between the cost and effect of influence was expressed using a linear function. However, assuming a linear cost function may be an over-simplification that may not apply to all real-world settings. Thus, here we study the competitive *influence maximisation* problem for other nonlinear cost functions.

In particular, we consider two settings. First, we consider the setting where the controller pays a fixed cost of access to positively influence a node in the network. We explore optimal allocation patterns for this setting against different competitor allocations and find that the cost of access parameter always regulates the optimal configuration of allocations between the continuous and the discrete approach. More specifically, we find that resources are diverted to high-degree hub nodes in the network as cost of access increases. For low costs of access, low-degree nodes are preferred. Thus bridging the gap between our proposed continuous approach and the traditional discrete method of influencing networks where nodes are either targeted or not.

Second, we consider more general settings where the effect of influence varies nonlinearly with the cost of influence. Specifically, we consider two scenarios: (i) where increasing allocations diminishes the marginal strength of influence experienced by nodes, and (ii) where nodes take longer or more allocations to start experiencing the effect of influence (observed as the delayed effect of influence). We find that optimal allocations and vote-shares are highly sensitive to budget conditions where allocations have a delayed effect on influence experienced by nodes. On the contrary, when allocations have a diminishing effect on influence experienced by nodes, optimal allocations and vote-shares show limited sensitivity to budget conditions. We further show that targeting the network optimally under low budget conditions and under nonlinear budget constraints can result in significant gain in vote-shares, when compared to more naïve approaches.

In this chapter, we illustrate examples where allocations do not exhibit a linear relationship with influence, e.g. adoption barriers and we argue the importance of considering nonlinear cost functions when studying competitive *influence maximisation*

in social networks. Our results discuss optimal allocations only for known competitor allocations, and an interesting future direction for this work would be to study the problem assuming incomplete knowledge of competitor strategy, in a game-theoretic framework.

Chapter 5

Presence of negative ties

5.1 Introduction

So far in the thesis, we have observed that the modification of collective opinions through external influence poses an interesting research question that has far-reaching societal and commercial implications (Wilder et al. (2018a); Ranganath et al. (2016); Watts et al. (2007); Kiss et al. (2017); Pastor-Satorras et al. (2015); Galam (1999); Easley et al. (2010); Jackson (2010)). However, the majority of this literature strictly investigates friendship networks where influence propagates based on positive recommendations and endorsements (Newman (2003)). Effects of negative relationships on opinion propagation have received limited attention and have been historically discounted from network dynamics given their sparse presence (Offer (2021)) and association with avoidance behaviour (Harrigan and Yap (2017)) i.e. people who dislike (or distrust) each other are unlikely to communicate (or be connected within a network).

However, the scope of anonymity and prevalence of fake profiles on social media platforms have made negative ties increasingly ubiquitous (Bae and Lee (2012); Pfeffer et al. (2014)), and even typical of many recommendation and trading networks (Guha et al. (2004); Leskovec et al. (2010)), thus demanding a more robust understanding of the impact of negative ties on network dynamics. More so, since negative ties are governed by a unique set of properties that make them distinctly different from positive ties (Easley et al. (2010)).

Here we take the *competitive influence maximisation* model proposed in Chapter 3 and apply it to networks with negative ties. We consider negative ties as antagonistic relationships that negatively influence social neighbours, and persuade them to adopt an opposing position (or opinion). Such relationships pose a unique challenge when trying to maximise influence in a network. For example, relationships of distrust

(negative reviews) on e-commerce platforms (e.g. eBay) following below-par experiences, can negatively impact future transactions or communication (Borgs et al. (2010); Chen et al. (2011)). Therefore requiring effective navigation of such ties when influencing a network externally. Particularly in competitive environments, as neglectfully targeting individuals that propagate net negative influence, can in turn, facilitate the spread of undesirable opinions.

In this chapter, we demonstrate the need for a negative-tie aware approach to maximise influence in social networks under competitive settings. In doing so, we make several contributions to the existing field of research. First, we modify the model presented in Chapter 3 to explore the problem in signed networks, and subsequently present an algorithm that maximises influence in networks with negative ties under competitive settings. We also demonstrate the importance of considering negative ties while maximising influence in real-world networks, and further show how the effectiveness of a negative-tie aware approach varies with network topology, availability of budget and competitor allocations. Finally, in addition to examining the problem for known competitor allocations, we also study the problem in a game-theoretic framework, where competitor strategies are unknown.

5.2   Outline

The chapter is structured as follows.

We modify the model presented in Chapter 3 to handle signed edges in the network, which we present in Section 5.3.

In Section 5.4, we investigate the problem in real-world settings to highlight its relevance and perform robust analyses using numerical methods in Section 5.5.

In Section 5.6, we present analytical support for our numerical results and further provide analytical expressions for optimal allocations in complex networks.

Finally, in Section 5.7 we explore the problem more deeply in game-theoretic scenarios.

5.3   Voting dynamics in signed networks

As before, we consider a population of N individuals interacting with one another in a social network, this time through congenial (positive) as well as hostile (negative) relationships. The structure of the network in this instance is therefore represented using a signed graph $G(V, E, W)$. Vertices $V = \{1, 2, \ldots, N\}$ denote individuals in the population connected through a set of edges E, depicting social connections (which

unlike the models considered in Chapters 3 and 4 can be negative). We therefore obtain $W \in R^{N \times N}$ as the corresponding *signed* weighted adjacency matrix. The weight of an edge w_{ij} here not only determines the strength of influence that flows from $i \longrightarrow j$, but also the type of influence propagated in that direction (positive or negative). A negative weight $w_{ij} < 0$ symbolises a negative edge and therefore implies negative influence from i on j, whereas $w_{ij} > 0$ suggests node j experiences positive influence from i. Edge weights w_{ij} are independent of each other in directed networks and nodes may experience different strengths and types (positive or negative) of influence from one another, i.e. $w_{ij} \neq w_{ji}$. In undirected networks, $w_{ij} = w_{ji}$.

As shown in Chapter 3, state variables $x_{A,i}$ and $x_{B,i} = 1 - x_{A,i}$, characterise the propensity of a node i to adopt opinion state A or B, and opinion propagation follows voter dynamics. In signed graphs, node i copies the opinion state of j if the edge from j to i is positive ($w_{ji} > 0$). Conversely, i adopts the opposing view if the edge weight is negative ($w_{ji} < 0$). We assume that external controllers influence the network only positively. Thus, if a node picks an external controller it strictly copies their opinion state. We quantify external influence on the network in terms of resource distribution vectors $\{p_A, p_B\} \in \mathbb{R}^N_+$, where any element $p_{A,i} \geq 0$ (or $p_{B,i} \geq 0$) represents the amount of influence experienced by a node i from controller A (or B). The vectors are constrained linearly by the respective budgets (B_A or B_B) available to each controller, i.e. $\sum_i p_{A,i} = B_A$ (or $\sum_i p_{B,i} = B_B$).

The rate at which a node i adopts opinion A is given by,

$$\frac{dx_{A,i}}{dt} = (1 - x_{A,i}) \cdot \phi_i(A) - x_{A,i} \cdot \phi_i(B). \tag{5.1}$$

The terms $\phi_i(A)$ and $\phi_i(B)$ represent the fraction of total influence experienced by node i in favour of opinion A and B respectively, defined as follows,

$$\phi_i(A) = \frac{\sum\limits_{j \in \mathcal{N}_i^+} w_{ji} x_{A,j} - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji}(1 - x_{A,j}) + p_{A,i}}{\sum\limits_{j \in \mathcal{N}_i^+} w_{ji} - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} + p_{A,i} + p_{B,i}};$$

$$\phi_i(B) = \frac{\sum\limits_{j \in \mathcal{N}_i^+} w_{ji}(1 - x_{A,j}) - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} x_{A,j} + p_{B,i}}{\sum\limits_{j \in \mathcal{N}_i^+} w_{ji} - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} + p_{A,i} + p_{B,i}}.$$

Here, A is chosen as the focal controller. Similar expressions can be derived for opinion B.

In signed networks, nodes experience both positive and negative influence from their neighbours. The collective positive influence is given by $\sum_{j \in \mathcal{N}_i^+} w_{ji} x_{A,j}$, and the total

negative influence (from neighbours in state B) is $\sum_{j \in \mathcal{N}_i^-} w_{ji}(1 - x_{A,j})$, where $\mathcal{N}_i^+$ and $\mathcal{N}_i^-$ are the sets of positive and negative neighbours of a node i. Edge weights w_{ji} strictly refer to incoming edges, and allocations from external controllers A and B on node i are $p_{A,i}$ and $p_{B,i}$ respectively.

Given the above, the steady-state equation for the system can be evaluated by replacing $\frac{dx_{A,i}}{dt} = 0$ in Eq. (5.1), to obtain

$$x_{A,i}^* = \frac{p_{A,i} - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} + \sum\limits_{j \in \mathcal{N}_i^+} w_{ji} x_{A,j} + \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} x_{A,j}}{\sum\limits_{j \in \mathcal{N}_i^+} w_{ji} - \sum\limits_{j \in \mathcal{N}_i^-} w_{ji} + p_{A,i} + p_{B,i}}. \tag{5.2}$$

Here, $x_{A,i}^*$ is the probability a node i has opinion state A at equilibrium. For a network of size N nodes, we obtain a system of N equations which can be summarised as follows,

$$\left[L + diag(p_A + p_B)\right] x_A = p_A + \vec{1}^T W^-,$$

$$\implies X_A = \frac{1}{N} \vec{1}^T x_A = \frac{1}{N} \vec{1}^T [L + diag(p_A + p_B)]^{-1} (p_A + \vec{1}^T W^-). \tag{5.3}$$

X_A in Eq. (5.3) denotes the total vote-share obtained by controller A at equilibrium. Assuming W^+ and W^- are the weighted adjacency matrices of the positive edges and negative edges of the network, $\vec{1}^T W^-$ captures the total strength of negative influence on every node in the network. L is an $N \times N$ matrix given by $L = diag(\vec{1}^T (W^+ - W^-)) - (W^+ + W^-)$. Here diagonal elements represent the absolute sum of all edge weights of a node i, $L_{ii} = \sum_{j \in \mathcal{N}_i^+} w_{ji} - \sum_{j \in \mathcal{N}_i^-} w_{ji}$ and off-diagonal elements are $L_{ij} = -w_{ji}$. For unweighted graphs, we get $L = diag(D) - (A^+ - A^-)$, where D is the degree-vector[1]. A^+ and A^- are the respective adjacency matrices of the positive and negative components. Note that Eq. (5.3) offers a unique solution for X_A as $[L + diag(p_A + p_B)]$ is diagonally-dominant, and hence invertible.

The formal optimisation problem can therefore be stated as

$$p_A^* = \text{argmax}_{p_A \in \mathcal{P}} \quad X_A^*(L, p_B), \tag{5.4}$$

---

[1] diag(D) is an $N \times N$ matrix where the diagonal elements capture the degrees of nodes in the network.

here $\mathcal{P}$ is a set of all possible allocations p_A such that $0 \leq p_{A,i} \leq B_A$
($\forall i \in \{1, 2, \ldots, N\}$), and $\sum_{i=1}^{N} p_{A,i} = B_A$.

For a passive opponent B (where p_B is fixed and known), controller A maximises its
vote-shares at equilibrium using Eq. (5.4). Closed-form solutions can be readily
obtained in networks with simplified structures (e.g. star networks) by solving the set
of equations in Eq. (5.3). To do this, we first determine the gradient $\nabla_{p_A} X_A$ by
differentiating Eq. (5.3) wrt to the allocation vector p_A as

$$\nabla_{p_A} X_A = \frac{1}{N} \vec{1}^T [L + diag(p_A + p_B)]^{-1} (I - diag(x_A)), \tag{5.5}$$

and then solving $\nabla_{p_A} X_A = 0$ to obtain optimal allocations p_A^* that yields maximum
vote-shares X_A^*. Obtaining analytical solutions for $\nabla_{p_A} X_A = 0$ in larger, more complex
networks however can be considerably challenging. We therefore use the gradient
ascent technique, discussed in Chapter 3, to determine optimal allocations in arbitrary
networks. Observe that Eq. (5.5) is the same expression we obtained for the gradient
in Chapter 3. This is because the new additional term $\vec{1}^T W^-$ in Eq. (5.3) (which is an
artefact of signed edges in the network), is independent of allocations p_A. Thus we
can use the gradient method proposed in Chapter 3 without having to modify it any
further.

Recall that the algorithm works by initialising the allocation vector as a uniformly
distributed random vector $p_A^{(0)}$ such that $\sum_i p_{A,i}^{(0)} = B_A$, and subsequently updating it
with the help of the gradient $\nabla_{p_A} X_A$, until a μ-approximated optimal allocation p_A^* is
obtained, given a budget B_A, opponent strategy p_B and network structure L. The
budget constraint is systematically imposed at each step, by projecting the resulting
allocation vector $p_A^{(t+1)}$ onto an N-simplex using the algorithm in (Chen and Ye
(2011)), such that $\sum_i p_{A,i} = B_A$. The learning parameter η is adjusted using
backtracking line search to ensure convergence (Boyd et al. (2004)).

To measure the effectiveness of the negative-tie aware method, we compare it to a
mechanism that discounts negative ties from network dynamics. Two commonly used
approaches to achieve this are (Li et al. (2013)): (i) using absolute values of edge
weights ($W_{ij} = |w_{ij}|$), i.e. all edges are assumed to be positive, and (ii) ignoring edges
with negative weights, i.e. $W_{ij} = max(0, w_{ij})$. In the rest of the chapter, we strictly
compare our proposed method to the first approach. This is because considering all
edges to be positive consistently yields equal or higher vote-shares than when
negative ties are removed, in undirected networks and directed network respectively
(See Appendix C.1 for more details).

Observe that when considering the absolute values of edge weights, the model is identical to the one proposed in Chapter 3. Recall that the expression for vote-shares at equilibrium is given by,

$$x_{A,i}^{*(+)} = \frac{p_{A,i} + \sum\limits_{j \in} w_{ji} x_{A,j}^{(+)}}{\sum\limits_{j \in} w_{ji} + p_{A,i} + p_{B,i}}, \tag{5.6}$$

$$\implies X_A^{(+)} = \frac{1}{N} \vec{1}^T x_A^{(+)} = \frac{1}{N} \vec{1}^T p_A [L^{(+)} + diag(p_A + p_B)]^{-1}. \tag{5.7}$$

Here $X_A^{(+)}$ indicates the final vote-share obtained by a controller to whom all edges appear strictly positive. $L^{(+)}$ is the $N \times N$ Laplacian matrix of the directed graph given as $L = diag(\vec{1}^T W) - W$, where W is the corresponding weighted adjacency matrix. Allocation vectors of both controllers remain unchanged and are given by p_A and p_B. The optimisation problem in this case can be defined as,

$$p_A^* = \text{argmax}_{p_A} X_A^{*(+)}(L^{(+)}, p_B, B_A). \tag{5.8}$$

The budget constraints apply as before.

We find that the gradient $\nabla X_A^{(+)}$ in this variant is similar to Eq. (5.5), and can be written as $\nabla_{p_A} X_A^{(+)} = 1/N \vec{1}^T [L^{(+)} + diag(p_A + p_B)]^{-1} (I - diag(x_A^{(+)}))$. Using this expression, we can employ gradient ascent steps $GA^{(+)}$ to numerically arrive at a global optimum p_A^*, for any given budget (B_A), competitor allocation (p_B) and unsigned network ($L^{(+)}$).

5.4   Relative gain in opinion shares in real-world networks

We begin by determining the relevance of the problem in the real world. To do this, we consider a real-world bitcoin trust network (Kumar et al. (2016, 2018)), where edges are labelled with positive or negative signs. The network is formed by members of the Bitcoin OTC platform who rate each other based on their interactions or experience during transactions. Ratings (or votes) exhibit trust and enable members to form their reputation on the platform for future transactions. Here we examine the extent to which relationships of distrust, if unobserved, can jeopardise *influence maximisation* efforts for an external controller while competing to promote the adoption of a new idea (e.g. financial product) in the community.

The Bitcoin network is a weighted-directed graph of 5.8K nodes. Edges are weighted between -10 and +10, where -10 implies complete distrust and +10 demonstrates ultimate trust. For our experiment, we focus on the single largest connected component of size 4.7K nodes, containing $\approx$ 94% of all edges. Within this component, $\approx$ 8.6% of edges are negatively weighted. In our model, influence flows in the direction opposite to that of rating, i.e. if a user i gives a score of +10 to user j, we assume that user i would copy the opinion state of user j with strength 10. In case of a negative score (-10), user i adopts the opposing state with equal strength.

We then perform optimisation processes GA and $GA^{(+)}$ on the network, and quantify the advantage gained from awareness of negative ties in terms of relative gain in vote-shares, given as $[X_A^{(GA)}/X_A^{(GA+)} - 1]$. Throughout the rest of the chapter, we will use this metric as a standard to measure the inefficiency of the naïve $GA^{(+)}$ algorithm in presence of negative ties, and in doing so, determine the fraction of vote-shares a controller can lose from ignoring negative edges in the network.

We initialise simulations with a uniformly random allocation vector p_A (constrained by the budget B_A) and learning rate $\eta = N$. An approximation factor of $\mu = 10^{-10}$ is used to terminate both algorithms. Here controller B naively targets all nodes in the network with uniform allocations $p_{B,i} = 1$ such that budget $B_B = N$. Controller A on the other hand, adopts GA or $GA^{(+)}$, depending on their knowledge of the network structure, to independently optimise allocations of their budget B_A. We summarise simulation results in Fig. 5.1 for the instance $B_A = N/4$ as it yields the highest gain in vote-shares (See Appendix C.1).

In Fig. 5.1a we illustrate the final 500 steps of the GA algorithm until convergence. The allocation vector $P_A^{(t)}$ at time-step t is incremented by stepping $\eta \nabla_{P_A^{(t)}} X_A$ in the direction of the gradient. Here η scales the step-length. The resulting vector is then projected back onto the N-simplex to preserve the budget constraint $\sum_i^N P_{A,i} = B_A = N/4$. The axes here represent the fraction of resources given to nodes with negative links ϵ_A and to the rest of the network with strictly positive edges $(1 - \epsilon_A)$.

In Fig. 5.1b we demonstrate the evolution of vote-shares when using both algorithms over time t. We find that vote-share monotonically increases when using GA to a final configuration of $X_A^{*(GA)} = 0.3908$. While $GA^{(+)}$ first yields an increase in vote-shares, it is followed by a steady decline in vote-shares that finally converges to $X_A^{*(GA+)} = 0.3582$, resulting in a gain of $\approx$ 9.11%. Thus implying that the awareness of negative edges can critically affect *influence maximisation* efforts in real-world networks and needs more robust exploration leading us to the following sections where we analyse the problem more thoroughly under controlled settings.
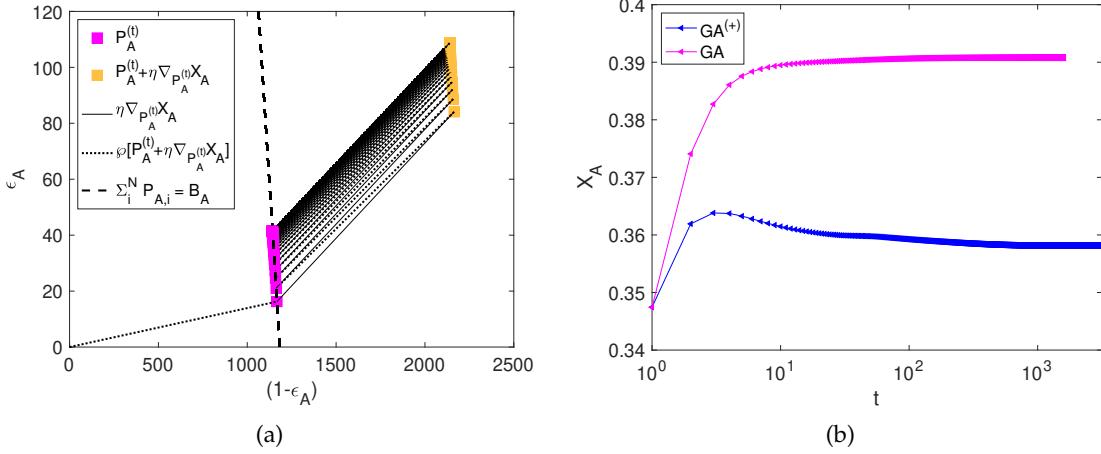
FIGURE 5.1: Simulations are performed on the single-largest connected component ($N = 4734$ nodes) of the Bitcoin OTC network. Figures show (a) step-wise allocations in the gradient ascent algorithm with negative ties (GA) and (b) change in vote-shares (X_A) through several iterations of both algorithms (GA and $GA^{(+)}$). Controller B passively targets all nodes in the network, uniformly, with a budget $B_B = N$. Controller A here has only a quarter of the resources available to B, $B_A = B_B/4 = N/4$. Algorithms are terminated using an approximation factor of $\mu = 10^{-7}$. The learning rate is initialised at $\eta = N$.

5.5   Numerical analysis

As per our model, there are three salient factors that can affect the difference in performance of both algorithms (GA and $GA^{(+)}$). These are the (a) topology of the network (L and $L^{(+)}$), (b) resource conditions (B_A/B_B) and (c) competitor strategies (p_B). We now systematically explore the effect of each of these factors on the gain in vote-shares.

5.5.1   Role of network topology

First, we look at the effect of network topology. Specifically, we examine how the placement and distribution of negative edges in the network can affect the gain in vote-shares.

Synthetic networks of size N=1000 nodes with an average positive degree $\sum_i^N k_{a,i}/N = \langle k_a \rangle = 16$ and average negative degree $\sum_i^N k_{b,i}/N = \langle k_b \rangle = 4$ are used in the following experiments where $k_{a,i} = \sum_{\mathcal{N}_i^+} w_i$ and $k_{b,i} = \sum_{\mathcal{N}_i^-} w_i$ respectively represent the total number of positive and negative edges of any node i. Here $\langle k_a \rangle$ and $\langle k_b \rangle$ are chosen to ensure that only 20% of all edges in the network are negative. We keep this fraction relatively low, first and foremost, to imitate real-world networks, as well as to demonstrate the importance of heeding negative ties even when they are sparingly present.

We assume negative relationships in the network are not correlated to positive ties. Thus we form synthetic signed networks by merging independently generated positive and negative graphs. The positive component has a size of N=1000 nodes with $\langle k_a \rangle = 16$. The size of the negative graph is given by $p \cdot N$, where p indicates the fraction of nodes in the network with negative edges. Varying p changes the assortativity of negative ties, and also allows us to alter the distribution of negative edges in the network. For lower values of p, negative edges are concentrated on fewer nodes, whereas for $p = 1$, every node in the network has at least one negative edge. Once the negative component is generated, $p \cdot N$ nodes are chosen from the positive component to combine both subgraphs, i.e. each node from the negative graph is mapped to a chosen node in the positive component, such that the node retains both sets of edges (from the positive and negative subgraphs) in the resulting network. The placement of negative edges is controlled by the way in which nodes are selected from the positive component for merging[2]. While combining both components, if there are multiple edges between two nodes, we keep the positive edge and ignore the negative tie. We show some examples of how we generate signed networks for for different values of p, in Fig. 5.2.

We vary the heterogeneity of degree distribution in the subgraphs to generate a broad spectrum of network structures for analysis. To that effect, subgraphs can either be homogeneous regular graphs (Newman et al. (2011)) or heterogeneous core-periphery type networks (Rombach et al. (2017)) with bi-modal degree distributions. Here we use the configuration model (Newman (2018)) to generate subgraphs.

Reg-Reg : As a benchmark case we use a homogeneous network where both positive and negative subgraphs are regular graphs (Reg-Reg). The positive component is a k_+-regular graph where $k_+ = \langle k_a \rangle = 16$. For the negative subgraph, we generate homogeneous networks of size $p \cdot N$ and degree $k_- = \langle k_b \rangle / p$, where p is varied between $0.1 \leq p \leq 1$. For those values of p where $k_- \notin \mathbb{Z}_+$, we first generate a k_--regular graph and then add ties to a fraction of randomly selected nodes such that they have $(k_- + 1)$ negative edges. This preserves the average degree $\langle k_b \rangle \approx 4$ and consequently the percentage of negative edges in the network. Here subgraphs are combined by merging randomly chosen nodes from the positive component.

CP-Reg : In this class of networks, the positive graphs have a heterogeneous core-periphery structure and is split evenly ($p_1 = p_2 = 0.5$), where half of all nodes (N/2) have high-degrees $k_{a,high} = 30$ and the rest have low-degrees $k_{a,low} = 2$. The negative component here has a regular or pseudo-regular structure (as described for Reg-Reg networks).

We adopt three different approaches while merging the negative and positive components. This varies the placement of negative ties in the network. For instance,

---

[2]Discussed in detail later in the section.
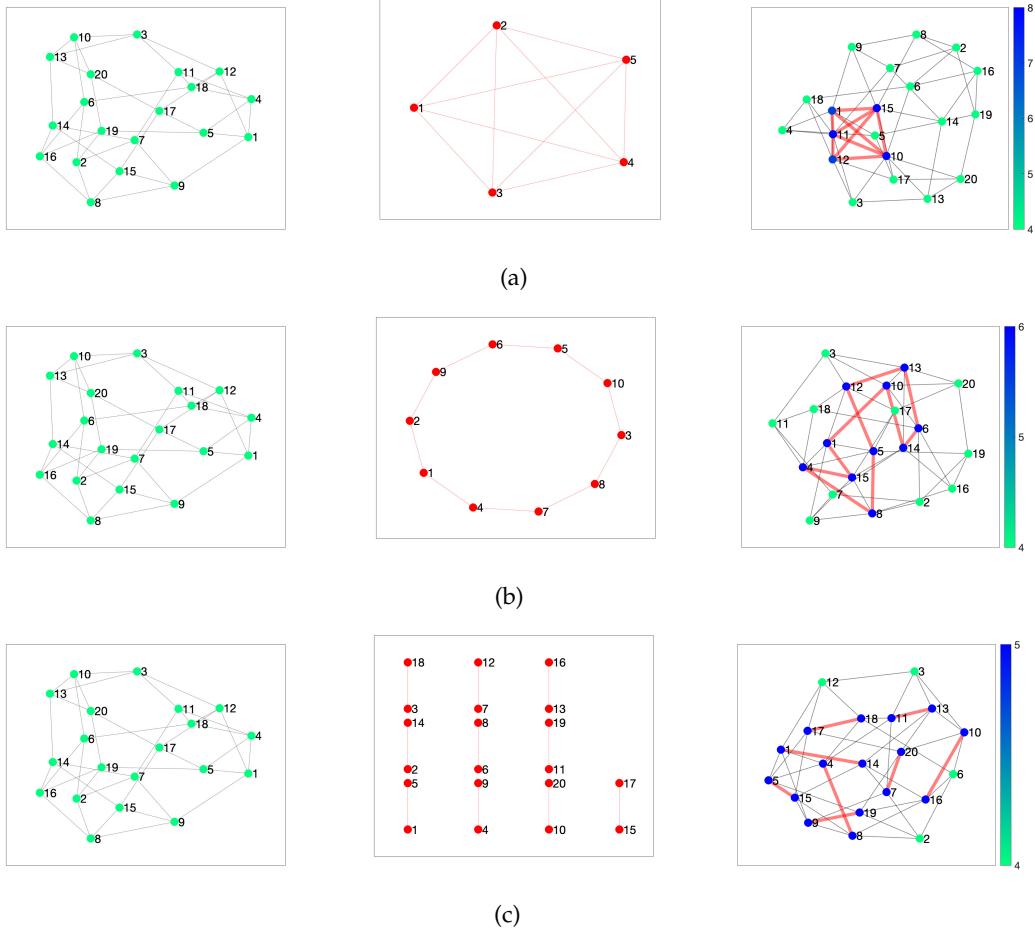
FIGURE 5.2: Figure showing generation of signed networks for different values of *p*. Figures (a), (b) and (c) show the generation of networks for (i) $p = 0.25$, (ii) $p = 0.5$ and (iii) $p = 1$, respectively. First a positive regular graph is created where $\langle k_a \rangle = 4$, shown in the left panel of each figure ((a), (b) and (c)). We use the same positive graph in each case to show how varying *p* changes the distribution and placement of negative edges.

nodes from the negative component are merged with high-degree nodes from the positive subgraph such that negative edges are preferentially added to high-degree nodes (CP-Reg-High). A similar approach is adopted when placing negative edges on low degree nodes (CP-Reg-Low). Finally, nodes from both subgraphs can be merged randomly for randomly placement of negative edges in the network (CP-Reg-Rand).

Reg-CP : Next, we have networks with a homogeneous positive component represented by a k_+-regular graph where $k_+ = \langle k_a \rangle = 16$, and a negative component represented using a core-periphery structure with $\langle k_b \rangle = 4$. Nodes in the negative subgraph are split equally into two classes: (i) high-degree nodes ($k_{b,high} = 2(k_b/p - 1)$) and (ii) low-degree nodes ($k_{b,low} = 2$). Here *p* ranges from $0.15 \leq p \leq 1$ as it is infeasible to form core-periphery structures for $p = 0.1$.

For each class of networks discussed above, we run simulations over 10 networks and
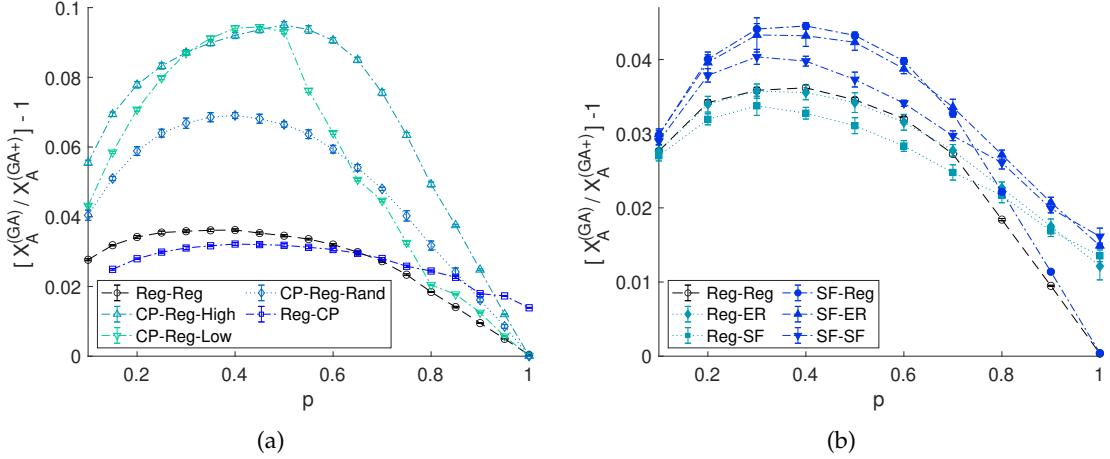
FIGURE 5.3: Figure showing relative gain in vote-shares as distribution of negative edges (p) is varied. Panel (a) shows the effect of heterogeneity of positive (CP-Reg) and negative edges (Reg-CP), benchmarked against networks where both positive and negative components are regular (Reg-Reg). Similarly, panel (b) examines role of heterogeneity in more generic networks. Heterogeneity of negative edges are varied as the negative graph is changed from regular (Reg-Reg) to random (Reg-ER), and finally, scale-free (Reg-SF) network. Heterogeneity of the positive component is achieved by replacing the regular positive graph (Reg-Reg) with a scale-free network (SF-Reg). For all simulations, networks of size $N = 1000$ nodes with $\langle k_a \rangle = 16$ and $\langle k_b \rangle = 4$ are used. Results are averaged over 10 instances and errorbars depict 95% confidence intervals.

summarise the results in Fig. 5.3a. We assume controller B is unaware of negative edges and passively targets each node with unit resource $p_{B,i} = 1$ where $\forall i \in \{1, 2, .., N\}$, and controller A has a budget $B_A = N$.

Scale-free and random networks To generalise our results further, we extend our study to more realistic Erdös-Rényi (ER) graphs (Erdős and Rényi (1960)) and Barabási-Albert (BA) networks (Barabási (2013)). Here the positive subgraph is either a homogeneous regular graph (Reg) or a heterogeneous scale-free network (SF), and the negative component is varied between a regular graph (Reg), a random graph (ER) or a scale-free network (SF). We assume heterogeneity is analogous to degree variance (Bell (1992)), and is gradually increased in negative components by replacing regular networks (Reg-Reg) with random graphs (Reg-ER) and then finally with scale-free networks (Reg-SF). In each case, the placement of negative edges is random. Finally, as before, results are compared against the homogeneous Reg-Reg graphs and present the results in Fig. 5.3b.

Fig. 5.3 highlights the effect of degree heterogeneity on gain in vote-shares as distribution of negative edges p is varied. In general, we find that relative gain in vote-shares first increases and then decreases as negative edges are distributed more evenly in the network. For networks with a homogeneous negative component the relative gain reduces to 0. An analytical explanation for this effect is provided in Appendix C.3.

In Fig. 5.3a, we find that networks with a positive core-periphery component and a regular negative subgraph (CP-Reg) yield the highest relative gain in vote-shares. Furthermore, relative gain is higher when placement of negative ties are degree-correlated (as in CP-Reg-High and CP-Reg-Low) compared to when randomly placed in the network (CP-Reg-Rand). A controller achieves maximum relative gain in vote-shares when negative edges are confined to nodes of a single degree-type ($p = 0.5$), high-degree (CP-Reg-High) or low-degree nodes (CP-Reg-Low). The maximum gain obtained in networks where negative edges are preferentially placed on high-degree nodes (CP-Reg-High) is $\approx 9.51\%$, as opposed to a maximum gain of $\approx 9.44\%$ when they are placed on low-degree nodes (CP-Reg-Low).

We observe similar trends in Fig. 5.3b where a maximum gain of $\approx 4.45\%$ is obtained in networks with a heterogeneous scale-free positive component and a homogeneous regular graph (SF-Reg). This result is close to the $\approx 6.91\%$ gain observed in CP-Reg-Rand networks where negative edges are randomly added to nodes (as in SF-Reg). We also find that substituting negative regular subgraphs with more heterogeneous components, such as a random graph (ER) or a scale-free network (SF), progressively decreases gain in vote-shares.

Once we have studied the effect of network topology on the efficiency of the GA algorithm, we explore the effect of resource conditions and competitor allocations on the gain in vote-shares.

5.5.2   Role of resource conditions and competitor allocations

We first examine how resource availability (B_A/B_B) impacts relative gain in vote-shares. Our preliminary results indicate insignificant gains ($\leq 1\%$) when the controller has more budget than its competitor $B_{A,i} > B_{B,i}$, implying that controllers with sufficient budgets can only marginally gain from exploiting their knowledge of negative ties. We therefore focus our analysis on lower budgets $B_{A,i} \leq B_{B,i}$, which yields a more interesting case for examination.

To determine if resource conditions can further exacerbate the difference in vote-shares between GA and $GA^{(+)}$, we limit our investigation to topologies that yield highest gains in the earlier experiments. To that effect, we choose networks with heterogeneous positive components and homogeneous negative subgraphs (CP-Reg). Simulations are performed on a single variant of these networks (CP-Reg-High) for conciseness.

The budget for controller B is fixed at $B_{B,i} = 1$ per node for the rest of the experiments in this section. Controller A has a per node budget $0.05 \leq B_{A,i} \leq B_{B,i}$. For simplicity, controller B continues to target the network passively. We assign three strategies to B where they either (i) avoid nodes with negative edges, (ii) strictly target

these nodes, or (iii) target all nodes uniformly, where B evenly distributes its resources over all the targeted nodes. We define these strategies specifically to examine how the gain in vote-shares depends on the competitor's knowledge of the network structure as well as on how they use this information.

Gain in vote-shares for each case is presented in Figs. 5.4a to 5.4c. Here Fig. 5.4a demonstrates the gain in vote-shares when controller B avoids nodes with negative edges. In Fig. 5.4b, controller B strictly targets nodes with negative ties and in Fig. 5.4c they target the network uniformly.

We find that controller A can lose considerable vote-shares by not employing an informed approach when controller B deliberately avoids targeting negative edges. On the other hand, as expected, we find that the difference in vote-shares is the least when controller B targets nodes with negative edges, as then the GA algorithm no longer has an advantage over the naïve approach $GA^{(+)}$.

Once again, we find that in all three cases, maximum gain is obtained when negative ties are restricted to high-degree nodes in the network ($p = 0.5$). We also observe that gain in vote-shares first increases with p, and then gradually decreases to 0. In Fig. 5.4a, a maximum gain of $\approx 17.85\%$ is observed when the budget ratio is $B_A / B_B = B_A = 0.3$. In Fig. 5.4b, gain is maximum ($\approx 5.75\%$) when controllers have equal budgets $B_A = B_B = 1$, and we note that difference in vote-shares is positively correlated with the budget B_A. Finally in Fig. 5.4c, we find that the gain in vote-shares reaches a maximum of $\approx 10.32\%$ at $B_A = 0.6$.

Next, we examine the effect of competitor allocations on relative gain in vote-shares. We fix the available resources for both controllers at $B_A = B_B = N$ to reduce any impact of uneven budgets on the results. Once again we consider CP-Reg-High networks.

We define allocations for controller B, using ϵ_B, which is the fraction of resources allocated to nodes with negative links. It is assumed that resources are always uniformly spread over targeted nodes, where $\epsilon_B \cdot B_B / p$ amount of influence is applied on every node that has a negative link, while $(1 - \epsilon_B) \cdot B_B / (1 - p)$ is given to the rest of the nodes (with strictly positive edges). We gradually vary $0 \leq \epsilon_B \leq 1$. Here boundary conditions correspond to instances where controller B either strictly targets nodes with negative ties ($\epsilon_B = 1$) or avoids them completely ($\epsilon_B = 0$). Results are shown in Fig. 5.4d.

As before, we observe that for all competitor allocations, the gain in vote-shares first increases and then decreases with p. The maximum gain of $\approx 10.13\%$ is obtained when controller B allocates a quarter of its budget $p_B = 0.25 \cdot B_B$ to all hub-nodes with negative edges where $p = 0.5$.
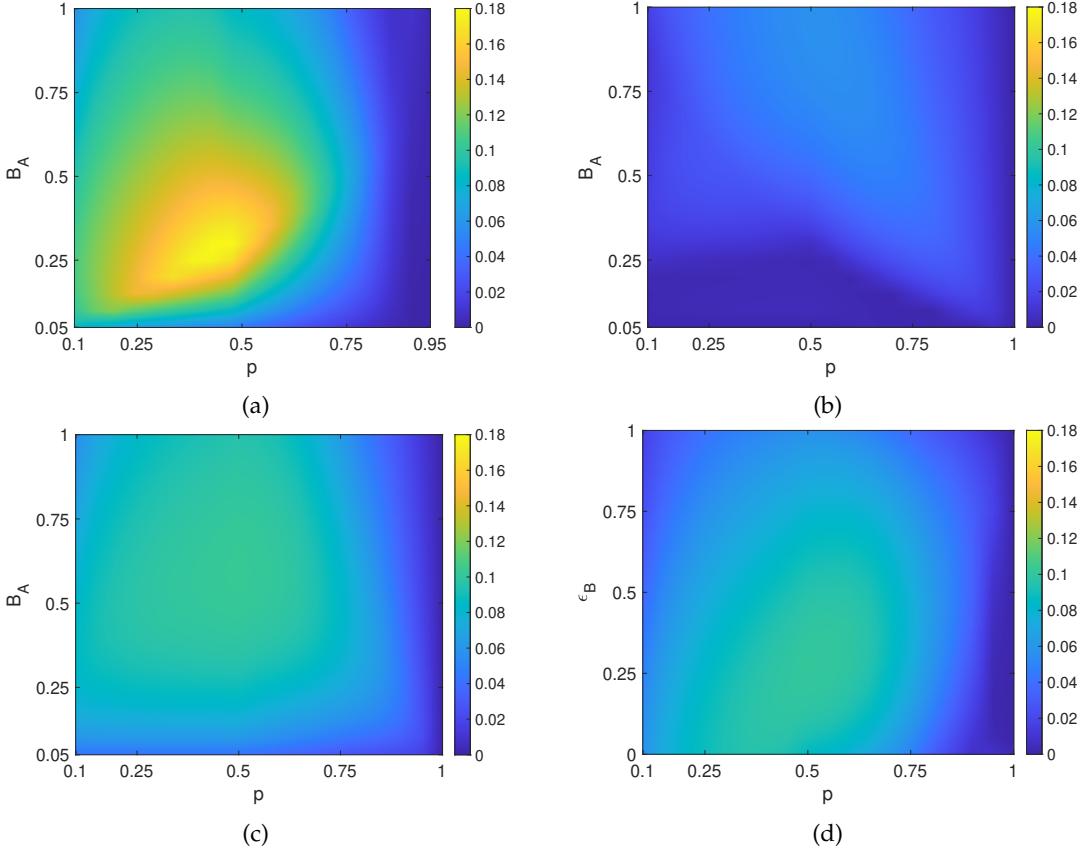
FIGURE 5.4: Panel showing relative gain in vote-share as p and budget ratios are varied against different competitor strategies. We examine three specific cases, controller B: (a) targets, or (b) avoids nodes with strictly positive edges and (c) targets all nodes uniformly. For all cases, controller B has a fixed budget of 1. Lastly, in (d) we quantify gain in vote-share as controller B changes its strategy by varying the fraction of budget ϵ_b used to target nodes with negative edges. Here both controllers have fixed budgets $B_A = B_B = 1$. Results are averaged over 10 CP-Reg-High networks of size $N = 1000$ nodes and $\langle k_a \rangle = 16$, $\langle k_b \rangle = 4$.

5.6   Analytical support

We now propose an analytical framework in support of our numerical results. Note that, obtaining closed-form analytical solutions for Eqs. (5.3) and (5.7) on networks with inherent complexities can be challenging. We therefore simplify the problem first by adopting a degree-based mean-field approach that approximates system dynamics and helps us obtain analytical expressions for optimal allocations. This approach works by grouping nodes with the same positive (k_a) and negative degrees (k_b) (where k_a is not correlated to k_b), and further assuming that their opinion state $x_{k_a k_b}$ is correlated to their overall degree ($\{k_a, k_b\}$). In doing so, we homogenise the effect of a heterogeneous neighbourhood on the state of a node which increases tractability of the problem. Additionally, we assume external controllers A and B uniformly target all nodes within a given class (with $a_{k_a k_b}$ and $b_{k_a k_b}$ allocations respectively) which further reduces the degrees of freedom in our problem.

We first derive the state of a node, with k_a positive and k_b negative degrees, from the steady-state equation Eq. (5.2) as

$$x_{k_a k_b} = \frac{a_{k_a k_b} + k_b + k_a \langle x_a \rangle - k_b \langle x_b \rangle}{k_a + k_b + a_{k_a k_b} + b_{k_a k_b}}. \tag{5.9}$$

Here $\langle x_a \rangle$ and $\langle x_b \rangle$ represent the expected state of a neighbouring node across a positive and a negative edge, respectively. Assuming that $P_a(k_a)$ and $P_b(k_b)$ describe the positive and negative degree distributions of a network, the expected behaviour of a neighbour at the end of a positive or a negative edge ($\langle x_a \rangle$ and $\langle x_b \rangle$) can be obtained as

$$\langle x_a \rangle = \sum_{k_b} P_b \sum_{k_a} P_a \frac{k_a}{\langle k_a \rangle} x_{k_a k_b} \tag{5.10}$$

$$\langle x_b \rangle = \sum_{k_a} P_a \sum_{k_b} P_b \frac{k_b}{\langle k_b \rangle} x_{k_a k_b}.$$

Note that here the term $k_a / \langle k_a \rangle$ (or $k_b / \langle k_b \rangle$) ensures that a node with higher positive (or negative) degree k_a (or k_b) has a higher chance of appearing as a neighbour exerting positive influence.

Now, using Eq. (5.9) and Eq. (5.10) we obtain the following self-consistency relations,

$$\langle x_a \rangle = \langle \frac{k_a}{\langle k_a \rangle} \frac{a_{k_a k_b} + k_b}{\Delta} \rangle + \langle \frac{k_a^2}{\langle k_a \rangle} \frac{1}{\Delta} \rangle \langle x_a \rangle - \langle \frac{k_a k_b}{\langle k_a \rangle} \frac{1}{\Delta} \rangle \langle x_b \rangle \tag{5.11}$$

$$\langle x_b \rangle = \langle \frac{k_b}{\langle k_b \rangle} \frac{a_{k_a k_b} + k_b}{\Delta} \rangle + \langle \frac{k_a k_b}{\langle k_b \rangle} \frac{1}{\Delta} \rangle \langle x_a \rangle - \langle \frac{k_b^2}{\langle k_b \rangle} \frac{1}{\Delta} \rangle \langle x_b \rangle.$$

These can be further solved to finally arrive at the mean-field expressions

$$\langle x_a \rangle = \left[\langle \frac{k_a}{\langle k_a \rangle} \frac{a_{ab} + k_b}{\Delta} \rangle - \frac{\langle \frac{k_a k_b}{\langle k_a \rangle} \frac{1}{\Delta} \rangle \langle \frac{k_b}{\langle k_b \rangle} \frac{a_{ab} + k_b}{\Delta} \rangle}{1 + \langle \frac{k_b^2}{\langle k_b \rangle} \frac{1}{\Delta} \rangle} \right] \left[1 - \langle \frac{k_a^2}{\langle k_a \rangle} \frac{1}{\Delta} \rangle + \frac{\langle \frac{k_a k_b}{\Delta} \rangle^2 \frac{1}{\langle k_a \rangle \langle k_b \rangle}}{1 + \langle \frac{k_b^2}{\langle k_b \rangle} \frac{1}{\Delta} \rangle} \right]^{-1},$$

$$\text{and,} \quad \langle x_b \rangle = \left[\langle \frac{k_b}{\langle k_b \rangle} \frac{a_{ab} + k_b}{\Delta} \rangle + \frac{\langle \frac{k_b}{\langle k_b \rangle} \frac{k_a}{\Delta} \rangle \langle \frac{k_a}{\langle k_a \rangle} \frac{a_{ab} + k_b}{\Delta} \rangle}{1 - \langle \frac{k_a^2}{\langle k_a \rangle} \frac{1}{\Delta} \rangle} \right] \left[1 + \langle \frac{k_b^2}{\langle k_b \rangle} \frac{1}{\Delta} \rangle + \frac{\langle \frac{k_a k_b}{\Delta} \rangle^2 \frac{1}{\langle k_a \rangle \langle k_b \rangle}}{1 - \langle \frac{k_a^2}{\langle k_a \rangle} \frac{1}{\Delta} \rangle} \right]^{-1},$$

$$\tag{5.12}$$

where $\Delta = a_{k_a k_b} + b_{k_a k_b} + k_a + k_b$.

Using the above, an approximation for the final configuration of the system at equilibrium can be obtained as

$$X_A = \langle x_{k_a,k_b} \rangle = \langle \frac{a_{ab} + k_b}{\Delta} \rangle + \langle \frac{k_a}{\Delta} \rangle \langle x_a \rangle - \langle \frac{k_b}{\Delta} \rangle \langle x_b \rangle, \tag{5.13}$$

where expressions for $\langle x_a \rangle$ and $\langle x_b \rangle$ come from Eq. (5.12).

Using Eq. (5.13) to derive closed-form analytical expressions for optimal allocations is still a complex task. Thus, we proceed with two alternative approaches. The first method uses a semi-analytical approach to obtain optimal allocations in simplified network structures using Eq. (5.13). Alternatively, we study the problem under limiting conditions which further streamlines the analytical approach by adding some additional assumptions to the model.

5.6.1   Semi-analytical approach

We approximate optimal allocations by solving Eq. (5.13) numerically in three types of networks, (i) Reg-Reg, (ii) CP-Reg-High and (iii) Reg-CP. We assume that networks are of size $N = 1000$ nodes $\langle k_a \rangle = 16$ and $\langle k_b \rangle = 4$, and p is varied between $0.075 \leq p \leq 0.975$. In all cases, controller B targets all nodes equally $b_{k_a,i k_b,i} = 1$, $\forall i \in [1, 2, \ldots, m]$ where there are m types (or groups) of nodes.

In Reg-Reg networks, we have two types of nodes: (i) $p \cdot N$ nodes with $k_{a,1} = 16$ positive edges and $k_{b,1} = 4/p$ negative edges and the remaining (ii) $(1-p) \cdot N$ nodes with $k_{a,2} = 16$ positive edges and no negative ties $k_{b,2} = 0$. We assume controller A distributes ϵ_A fraction of its budget equally over all nodes with negative edges, such that $a_{k_a,1 k_b,1} = \epsilon_A \cdot B_A / p$. The rest of the nodes receive $a_{k_a,2 k_b,2} = (1 - \epsilon_A) \cdot B_A / (1 - p)$ allocations.

In CP-Reg-High networks, nodes can be segregated in two or three classes depending on the value of p. The positive component in these networks have a core-periphery structure with high-degree nodes $k_{a,1} = 30$ and low-degree nodes $k_{a,2} = 2$. When $p = 0.5$, we have two groups of nodes, (i) high-degree nodes with negative links and (ii) low degree nodes with only positive edges. As described earlier, resources are split into ϵ_A which is allocated to the first group of nodes and $(1 - \epsilon_A)$ given to low-degree nodes. For $p < 0.5$, we have three groups of nodes, (i) high-degree nodes with negative links $(k_{a,1} + k_{b,1})$, (ii) high-degree nodes with only positive links $(k_{a,1})$ and (iii) low degree nodes with only positive links $(k_{a,2})$. The same can be derived for $p > 0.5$, where all high-degree nodes, and a fraction of low-degree nodes have negative links. In each case, the budget is split over three groups as $\epsilon_{A,1}$, $\epsilon_{A,2}$ and $[1 - (\epsilon_{A,1} + \epsilon_{A,2})]$. We simultaneously solve for $\epsilon_{A,1}$ and $\epsilon_{A,2}$ to determine the optimal allocation on the network. The amount of budget allocated to nodes with negative edges is given by $\epsilon_A = \epsilon_{A,1}$ when $p < 0.5$ and by $\epsilon_A = \epsilon_{A,1} + \epsilon_{A,2}$ when $p > 0.5$.
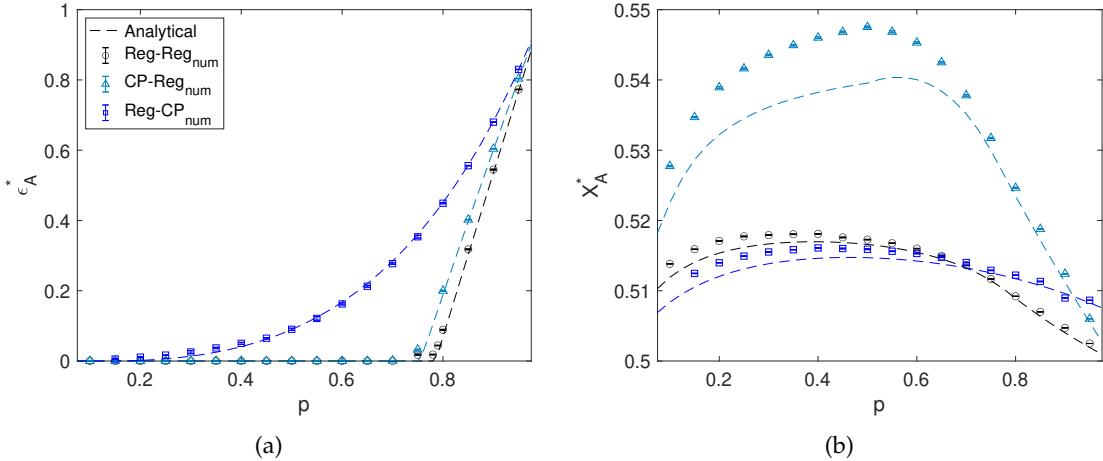
FIGURE 5.5: Figure showing optimal allocations and vote-shares at equilibrium for values of $p \in [0.075, 0.975]$. Results shown for three types of networks : (i) Reg-Reg, (ii) CP-Reg and (iii) Reg-CP. Analytical solutions are shown using dashed lines. Numerical results are obtained through simulations on networks of size N=1000 nodes and averaged over 10 networks. Error bars are shown for 95% confidence intervals.

Optimal allocations are derived in Reg-CP networks using a comparable approach. Nodes form three clusters for $p < 1$, (i) p_1 nodes with high-negative-degrees $k_{b,1} = 2(\langle k_b \rangle / p - 1)$ and k_a positive edges, (ii) p_2 nodes with low-negative-degree $k_{b,2} = 2$ and k_a positive edges, and finally (iii) nodes with only k_a positive edges. Here all classes of nodes have the same positive degree $k_a = 16$ and the fraction of nodes with negative edges are given as $p = p_1 + p_2$. Budget allocation for each class of nodes is given as $a_{k_a k_{b,1}} = \epsilon_{A,1}/p_1$, $a_{k_a k_{b,2}} = \epsilon_{A,2}/p_2$ and $a_{k_a k_{b,3}} = (1 - \epsilon_A)/(1 - (p_1 + p_2))$, where nodes with negative edges receive $\epsilon_A = (\epsilon_{A,1} + \epsilon_{A,2})$ fraction of the budget.

For each type of network, we obtain related expressions using Eq. (5.13), which we solve semi-analytically to obtain optimal allocations ϵ_A^* and maximum vote-shares X_A^*. Results are presented in Fig. 5.5 where they are compared against previously obtained numerical results.

From initial observations, we find that our numerical results are in good agreement with the theoretical solutions in both Figs. 5.5a and 5.5b. There is however, some discrepancy in vote-share results (Fig. 5.5b), which can be attributed to uniform allocations to each group of nodes in the analytical approach. The numerical method in contrast allows more sophisticated, flexible allocations to nodes of the same type, thereby realising higher vote-shares. In Fig. 5.5a, we note that when negative edges are more clustered (i.e. $p < 0.3$), nodes with negative edges are strictly avoided in Reg-Reg and CP-Reg-High networks. However, as p increases and negative edges are more dispersed over the network, allocations to these nodes increase. This behaviour contrasts Reg-CP graphs, where we find positive allocations to nodes with negative ties even at low values of p. We further observe that, these allocations are to nodes

with low-negative-degrees, hinting that allocations to a node are correlated to its negative degree.

5.6.2   In the limit of large $\langle k_a \rangle$

As an alternate approach, we apply a limiting condition to simplify the mean-field approximation in Eq. (5.13). Here, we study influence dynamics in the limit of large average positive degree, in comparison to controller budgets and average negative degree, i.e. $\langle k_a \rangle \gg \langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + \langle k_b \rangle$. We are motivated observations of real-world networks where negative edges are often sparingly present among densely connected positive edges (Leskovec et al. (2010)). We begin by performing a series expansion on Eq. (5.13) in the above limit to obtain,

$$
\begin{aligned}
X_A \approx {} & \langle \frac{a_{k_a k_b}}{k_a} \rangle + \langle \frac{k_b}{k_a} \rangle + \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle} - \frac{\langle \frac{(a_{k_a k_b} + k_b)(a_{k_a k_b} + b_{k_a k_b} + 2k_b)}{k_a} \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle} \\
& + \frac{a_{k_a k_b} + \langle k_b \rangle}{(a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle)^2} \left(\langle \frac{k_b^2}{k_a} \rangle + \langle \frac{(a_{k_a k_b} + b_{k_a k_b} + k_b)(a_{k_a k_b} + b_{k_a k_b} + 3k_b)}{k_a} \rangle \right. \\
& \left. - \langle \frac{a_{k_a k_b} + b_{k_a k_b} + k_b}{k_a} \rangle \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle} - \langle \frac{k_b}{k_a} \rangle \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + \langle k_b \rangle} \right),
\end{aligned}
\tag{5.14}
$$

and use it to derive optimal allocations $a^*_{k_a k_b}$, obtained as

$$
a^*_{k_a k_b} = \frac{1}{2} \left(\frac{\langle a_{k_a k_b} \rangle - \langle b_{k_a k_b} \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} b_{k_a k_b} + \frac{\langle a_{k_a k_b} \rangle - 3\langle b_{k_a k_b} \rangle - 2\langle k_b \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} k_b + \langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle \right).
\tag{5.15}
$$

Details of the derivation can be found in Appendix C.2.

From Eq. (5.15) we can see that optimal allocation $a^*_{k_a k_b}$ to a node characteristically depends on its negative degree k_b and the competitor allocations $b_{k_a k_b}$ to it. We verify these findings in the rest of the section using numerical simulations.

We first validate the dependence of optimal allocations $a^*_{k_a k_b}$ on competitor allocations $b_{k_a k_b}$, where controller B has a k_a-dependent strategy. SF-Reg networks are used to test this relation, as the scale-free positive component can then be exploited to generate multiple data points for comparison between numerical and analytical results. Note that there is a negative-degree dependent term in Eq. (5.15) that contributes to optimal allocations. We reduce the effect of this term in our analysis by first setting the per node budget for controller A to $\langle a_{k_a k_b} \rangle = 3\langle b_{k_a k_b} \rangle + 2\langle k_b \rangle$. However, this condition constrains the analysis to positive correlations between allocations from both
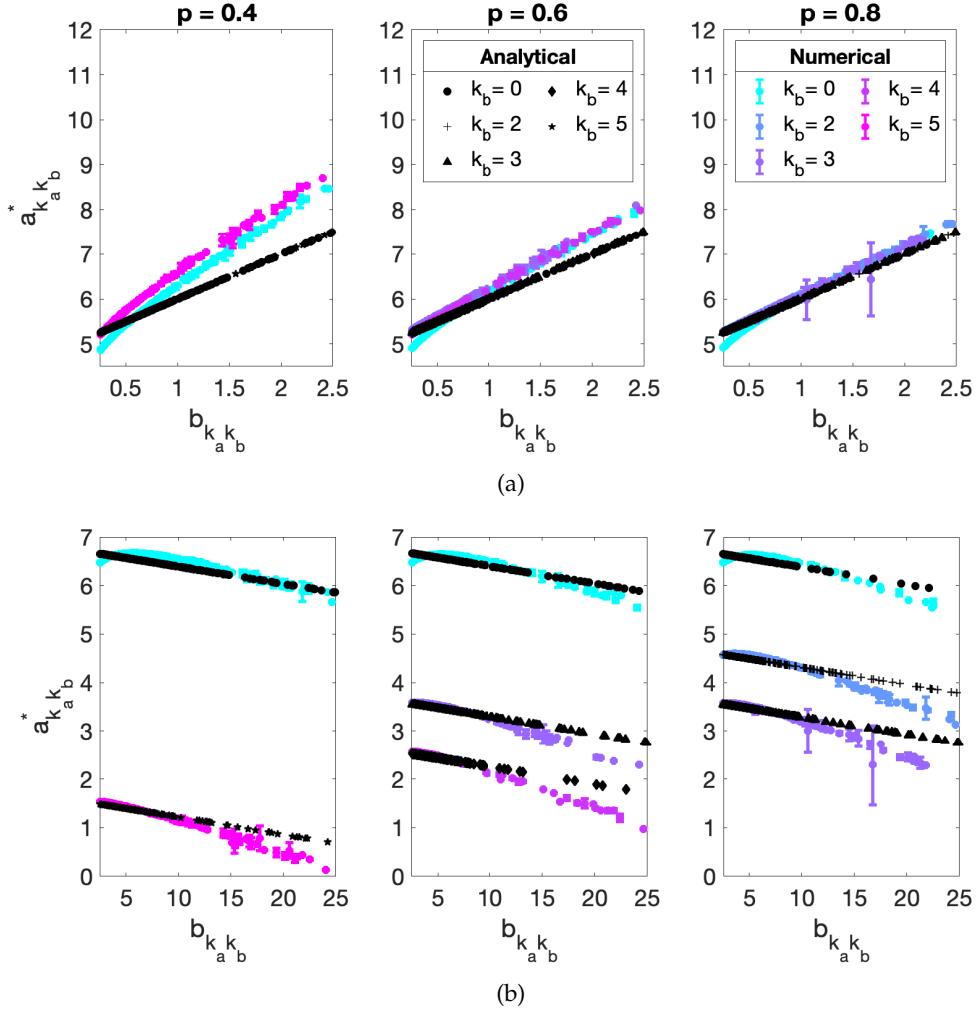
FIGURE 5.6: Figures showing mean allocations to nodes as a function of competitor allocations $b_{k_a k_b}$ in SF-Reg networks of size $N = 1000$ nodes where $\langle k_a \rangle = 50$ and $\langle k_b \rangle = 2$. We examine three instances of negative tie distributions $p \in [0.4, 0.6, 0.8]$ (left to right). We show (a) positive and (b) negative correlations between optimal allocations and competitor allocations. Controllers have the following budgets: (a) $\langle a_{k_a k_b} \rangle = 5.5$ and $\langle b_{k_a k_b} \rangle = 0.5$, and (b) $\langle a_{k_a k_b} \rangle = 4.5$ and $\langle b_{k_a k_b} \rangle = 5$. Controller B here follows a k_a−dependent strategy. Numerical results are averaged over 10 networks. Errorbars show a 95% confidence interval.

controllers, as now $\langle a_{k_a k_b} \rangle > \langle b_{k_a k_b} \rangle$ (see the first term in Eq. (5.15)). When analysing the negative relation between the optimal allocations and the competitor allocations, the budget is set such that $\langle a_{k_a k_b} \rangle < \langle b_{k_a k_b} \rangle$. In this case however, the negative-degree-dependent term contributes to the optimal allocation result, and the regular structure of the negative component in the networks is chosen to limit the variation in optimal allocation $a^*_{k_a k_b}$ caused by negative degrees.

For our simulations we consider networks where $\langle k_a \rangle = 50$, is sufficiently large compared to $\langle k_b \rangle = 2$. Controller B distributes its budget in proportion to the positive degree of nodes ($p_{B,i} \propto k_{a,i}$). When analysing the positive relation between allocations $a^*_{k_a k_b}$ and $b_{k_a k_b}$, the budget for controllers A and B are fixed as $\langle a_{k_a k_b} \rangle = 5.5$ and

(a)

(b)
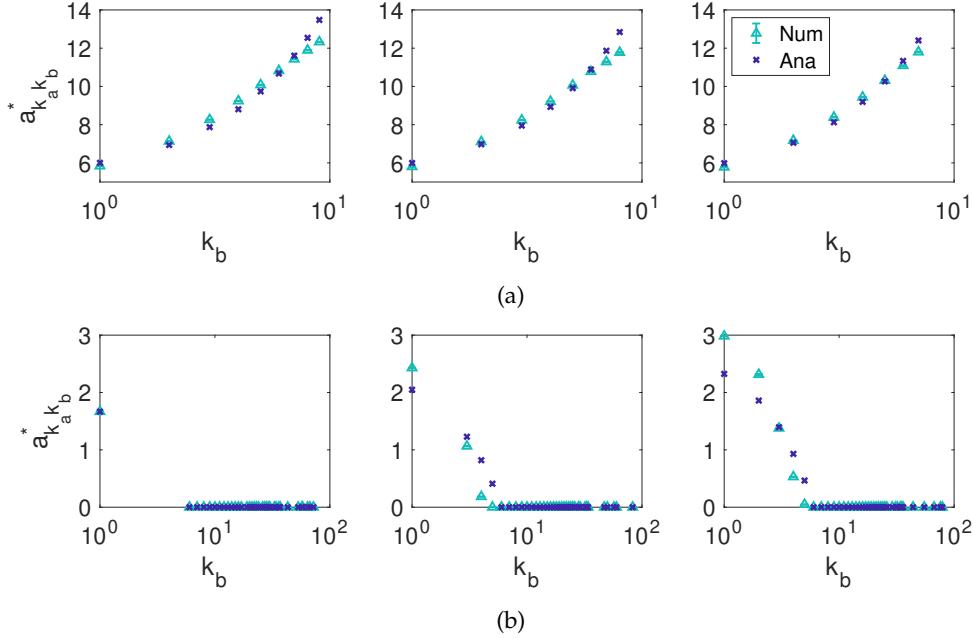
FIGURE 5.7: Figures showing mean allocations to nodes as a function of negative degrees k_b in networks of size $N = 1000$ nodes. We examine three instances of negative tie distributions $p \in [0.4, 0.6, 0.8]$ (left to right). Figure (a) demonstrates a positive correlations between optimal allocations and negative degrees k_b in Reg-ER networks where $\langle k_a \rangle = 50$ and $\langle k_b \rangle = 2$. Similarly, Figure (b) shows a negative relation between the two, in Reg-SF networks where $\langle k_a \rangle = 16$ and $\langle k_b \rangle = 4$. Controllers have the following budgets: (a) $\langle a_{k_a k_b} \rangle = 7$ and $\langle b_{k_a k_b} \rangle = 0.5$, and (b) $\langle a_{k_a k_b} \rangle = 1$ and $\langle b_{k_a k_b} \rangle = 1$. Controller B targets the network uniformly. Numerical results are averaged over 10 networks. Errorbars show a 95% confidence interval.

$\langle b_{k_a k_b} \rangle = 0.5$ respectively. For negative correlations we carefully choose $\langle a_{k_a k_b} \rangle = 4.5$ and $\langle b_{k_a k_b} \rangle = 5$ to avoid defying the positivity constraint on optimal allocations. We summarise results from the simulations in Figs. 5.6a and 5.6b, where for the sake of brevity, we demonstrate results in only three instances of $p \in [0.4, 0.6, 0.8]$.

Next, we explore the dependence of optimal allocations $a^*_{k_a k_b}$ on the negative degree k_b of nodes. We study positive regular graphs with negative heterogeneous components (to ensure multiple data-points of comparison between analytical and numerical results). Note that the correlation is positive when $\langle a_{k_a k_b} \rangle > 3 \langle b_{k_a k_b} \rangle + 2 \langle k_b \rangle$, and negative when $\langle a_{k_a k_b} \rangle < 3 \langle b_{k_a k_b} \rangle - 2 \langle k_b \rangle$. We conduct simulations to test each relation and present the results in Figs. 5.7a and 5.7b. To analyse the positive correlation, we run simulations in Reg-ER networks where $\langle k_a \rangle = 50$ and $\langle k_b \rangle = 2$. Here we assume that controller B has a budget of $B_B = \langle b_{k_a k_b} \rangle \cdot N = 0.5 \cdot N$, which they distribute uniformly across the network. Uniform allocation from the competitor limits the effect of the corresponding $b_{k_a k_b}$-dependent term on optimal allocations. We assume controller A has a budget of $\langle a_{k_a k_b} \rangle = 7$ which satisfies the necessary condition to study the positive correlation between optimal allocations and negative degree of nodes. In the second instance, we use a Reg-SF network with $\langle k_a \rangle = 16$ and $\langle k_b \rangle = 4$. Controllers have equal budgets $\langle a_{k_a k_b} \rangle = \langle b_{k_a k_b} \rangle = 1$, which eliminates the effect of the

$b_{k_a k_b}$-dependent term in Eq. (5.15). As before, controller B targets the network uniformly.

As shown in Figs. 5.6 and 5.7, we find that analytical results approximate numerical simulations reasonably well under the limiting condition. Fig. 5.6a shows a positive linear relation between optimal allocations $a^*_{k_a k_b}$ and competitor allocation $b_{k_a k_b}$ under conditions of excess budget. We also observe that agreeability between numerical and analytical results increases with homogeneity of negative edges in the network ($p = 0.8$). Fig. 5.6b verifies a negative relation between $a^*_{k_a k_b}$ and $b_{k_a k_b}$ when resources are scarce. However, in this instance, we find that optimal allocations are grouped, based on the negative degree of nodes, as the k_b-dependent term is not 0.

In Figs. 5.7a and 5.7b, we demonstrate correlations between optimal allocations and negative degree of nodes k_b. Fig. 5.7a shows a positive linear relation with negative degree k_b for larger budgets $\langle a_{k_a k_b}\rangle > 3\langle b_{k_a k_b}\rangle + 2\langle k_b\rangle$. Contrarily, an inverse relationship is observed between the optimal allocations and the k_b, in Fig. 5.7b, when budget $\langle a_{k_a k_b}\rangle < 3\langle b_{k_a k_b}\rangle - 2\langle k_b\rangle$.

The above results show that despite the additional assumptions, the proposed analytical framework provides a good estimate of the system dynamics in heterogeneous signed networks. We further show that in the limit of large average positive degree, optimal allocations depend on the negative degree of nodes and the competitor allocations on nodes —and not on the positive degree of nodes. The type of correlation, positive or negative, relies on the budget available to the focal controller. When resources are in excess, allocations increase linearly with competitor allocation and negative degree, whereas under low budget conditions this relation becomes inversely proportional.

5.7 Game-theoretic scenario

So far we optimise allocations for controller A against a passive competitor B. In this section, we explore the problem under game-theoretic settings where both controllers A and B actively optimise budget distribution over signed networks. In our game, controllers represent players who optimise their respective strategies (p_A and p_B), to maximise their utilities or vote-shares (X_A and X_B).

Our motive here, as before, is to determine if the controller A receives a more favourable outcome against an active controller B, when they are fully informed of the network structure. For this reason, we explore and compare two cases that emerge from varying awareness of negative ties in controller A. In the first case, we assume A has complete knowledge of the network structure and in the second, we assume that

they are unaware of polarised edges and can only observe the network as a strictly unsigned (positive) graph. Controller B in both cases is unaware of signed edges.

We first define the problem analytically and then compare it to numerical results in Fig. 5.8. To begin, we derive an analytical expression for the instance where controllers have imperfect knowledge of the network. We use the degree-based mean-field approach discussed in Section 5.6 to classify nodes based on their degrees, where node behaviour $\langle x_k \rangle$ strongly correlates with their degree k. Since controllers cannot discern between positive and negative edges, they consider a node i has degree k_i where $k_i = k_{a,i} + k_{b,i}$.

First, we use the steady-state equation (Eq. (5.2)) to obtain the state of a node with degree k,

$$x_k = \frac{k\langle x \rangle + a_k}{k + a_k + b_k},\tag{5.16}$$

where a_k and b_k are allocations from controllers A and B for the given class of nodes, and $\langle x \rangle$ is the average behaviour of a node in the network. Here too controllers are assumed to distribute allocations uniformly within each class.

Given a degree distribution P_k, the expected behaviour of a node in the network is,

$$\langle x \rangle = \frac{\sum_k P_k k x_k}{\langle k \rangle} = \frac{\sum_k P_k k x_k}{\sum_k P_k k}.\tag{5.17}$$

Using Eq. (5.16) and Eq. (5.17), we obtain

$$\langle x \rangle = \frac{\sum_k \frac{P_k k a_k}{k+a_k+b_k}}{\sum_k \frac{P_k k (a_k+b_k)}{k+a_k+b_k}}.\tag{5.18}$$

This leads us to an expression for the total vote-share at steady-state as,

$$X_A^{(+)} = \sum_k P_k x_k = \frac{\left(\sum_k \frac{P_k k}{k+a_k+b_k}\right)\left(\sum_k \frac{P_k k a_k}{k+a_k+b_k}\right)}{\left(\sum_k \frac{P_k k (a_k+b_k)}{k+a_k+b_k}\right)} + \left(\sum_k \frac{P_k k}{k+a_k+b_k}\right).\tag{5.19}$$

Although we show the above derivations for A, the same process can be repeated to determine the utility function for controller B,

$$X_B^{(+)} = \frac{\left(\sum_k \frac{P_k k}{k+a_k+b_k}\right)\left(\sum_k \frac{P_k k b_k}{k+a_k+b_k}\right)}{\left(\sum_k \frac{P_k k (a_k+b_k)}{k+a_k+b_k}\right)} + \left(\sum_k \frac{P_k k}{k+a_k+b_k}\right).\tag{5.20}$$

Note that, Eq. (5.19) is the utility function we use for controller A when they are unaware of negative edges in the network and hence they target the network indiscriminately, while Eq. (5.13) guides a negative-tie sensitive approach. To determine how much controller A gains from exploiting their knowledge of the network in the game-theoretic setting, we compare vote-shares obtained using both approaches.

We optimise the utility functions of both controllers simultaneously (as shown in Section 3.5) to obtain the game equilibrium in each case. We determine the pair of optimal strategies at the equilibrium for both games (one where A is aware of negative ties and another where they are not, both times against an imperceptive competitor B). In the instance where both controllers are unaware of negative ties in the network, the utility functions Eq. (5.19) and Eq. (5.20) are simultaneously solved to arrive at the equilibrium state $[X_A^*, X_B^*]$. The corresponding strategies at this point are given by $[p_A^*, p_B^*]$. Note that here the true vote-shares $[\widehat{X}_A^*, \widehat{X}_B^*]$ differ from their observed utilities $[X_A^*, X_B^*]$, and can be obtained by inserting p_A^* and p_B^* in Eq. (5.3). In the other instance, controller A (using Eq. (5.13) as their utility function) has access to their true utility at all times. Once an equilibrium state is reached, $[p_A^*, p_B^*]$ are used the determine the true vote-share for B, $\widehat{X}_B^*$. Here it is important to note that both controllers assume the network structure to be common knowledge at all times. Therefore, from both their perspectives, the game always appears to be a zero-sum game (Fudenberg and Tirole (1991)).

While solving the problem semi-analytically, we express strategies using ϵ_A (or ϵ_B), which represents the fraction of resources allocated by a controller to nodes with negative ties. Solving their vote-share functions simultaneously, we obtain the equilibrium state $[\epsilon_A^*, \epsilon_B^*]$, and finally measure the difference in vote-shares.

For our simulations, we explore three types of networks: (i) Reg-Reg, (ii) CP-Reg-High and (iii) Reg-CP, where $\langle k_a \rangle = 16$ and $\langle k_b \rangle = 4$. For brevity, we examine only one instance of negative tie distribution $p = 0.5$. Note that for Reg-Reg and CP-Reg-High networks, resources are split between two node classes as ϵ_A and $1 - \epsilon_A$. For Reg-Reg networks the *two* classes of nodes are (i) $\{k_{a,1}, k_{b,1}\} = \{16, 8\}$ and (ii) $\{k_{a,2}, k_{b,2}\} = \{16, 0\}$. For CP-Reg-High networks they are (i) $\{k_{a,1}, k_{b,1}\} = \{30, 8\}$ and (ii) $\{k_{a,2}, k_{b,2}\} = \{2, 0\}$. In Reg-CP networks, we have *three* classes of nodes, where p nodes with negative edges are split even between p_1 and p_2 such that (i) $p_1 = 0.25$, $\{k_{a,1}, k_{b,1}\} = \{16, 14\}$, (ii) $p_2 = 0.25$, $\{k_{a,2}, k_{b,2}\} = \{16, 2\}$ and finally the rest of the nodes (iii) $p_3 = 0.5$, $\{k_{a,3}, k_{b,3}\} = \{16, 0\}$. Consequently the budget is split three-ways into ϵ_1, ϵ_2 and $\epsilon_3 = 1 - (\epsilon_1 + \epsilon_2)$. Here $\epsilon = \epsilon_1 + \epsilon_2$. In all cases, the necessary boundary conditions are imposed. Fig. 5.8 demonstrates how equilibrium strategies and gain in vote-shares depend on budget ratios. Analytical results are complemented with numerical simulations. GA is used as an optimiser when controller A has complete knowledge of the network and conversely $GA^{(+)}$ is used when they cannot

discern the polarity of the edges. In both cases, controller B uses $GA^{(+)}$ to determine best response.

We find that simulations always converge to a pure equilibrium in all cases. When controllers cannot observe negative edges, the equilibrium strategy is identical for both controllers, and uniformly targets the network ($\epsilon_A = \epsilon_B = p = 0.5$) (shown in Chapter 3). On the other hand, we find that controller A avoids nodes with negative ties when they have full knowledge of the network under conditions of limited resources. This strategy influences the allocation patterns of competitor B, who now diverts less resources to nodes with negative edges. We see that by avoiding negative edges in the network, controller A inadvertently shares information about their own prioritisation of nodes in the network, which then results in a loss in vote-shares as shown in Fig. 5.8.

As the budget ratio increases, controller A starts redirecting some resources to nodes with negative edges. When this occurs, controller B increases their allocations to these nodes. As the budget ratio increases to $B_A/B_B \approx 5$, controller B stops competing over nodes with negative ties and begins to gradually reduce their allocation (ϵ_B) to them. Around $B_A/B_B = 10$, both controllers target the network uniformly ($\epsilon_A = \epsilon_B = p = 0.5$), and beyond this point, ϵ_B continues to decrease as ϵ_A increases steadily. Once again, we observe a region where controller A loses vote-shares from implementing their knowledge of negative edges in the network. This happens as they start pursuing nodes with negative edges, with more than half of their budget $\epsilon_A > 0.5$. However, as budget ratio increases to $B_A/B_B \approx 10^2$, controller A once again starts gaining from the knowledge of the network structure.

We also note that the results are similar across all three types of networks, and numerical simulations closely approximate analytical results in all cases. Gain in vote-shares is significant in CP-Reg networks compared to the other networks, and we find early allocations to nodes with negative edges in Reg-CP networks, which is consistent with results in Fig. 5.5.

5.8   Summary and conclusions

The study of opinion dynamics has been conventionally focused on networks with strictly positive edges. In the real-world however, networks often contain negative social connections, which can spread negative or opposing influence, thus creating a pressing need to understand how these edges affect *influence maximisation* efforts in networks.

To address this concern, we presented a model for competitive spread of opinions in signed networks under voter dynamics with continuous distribution of influence in
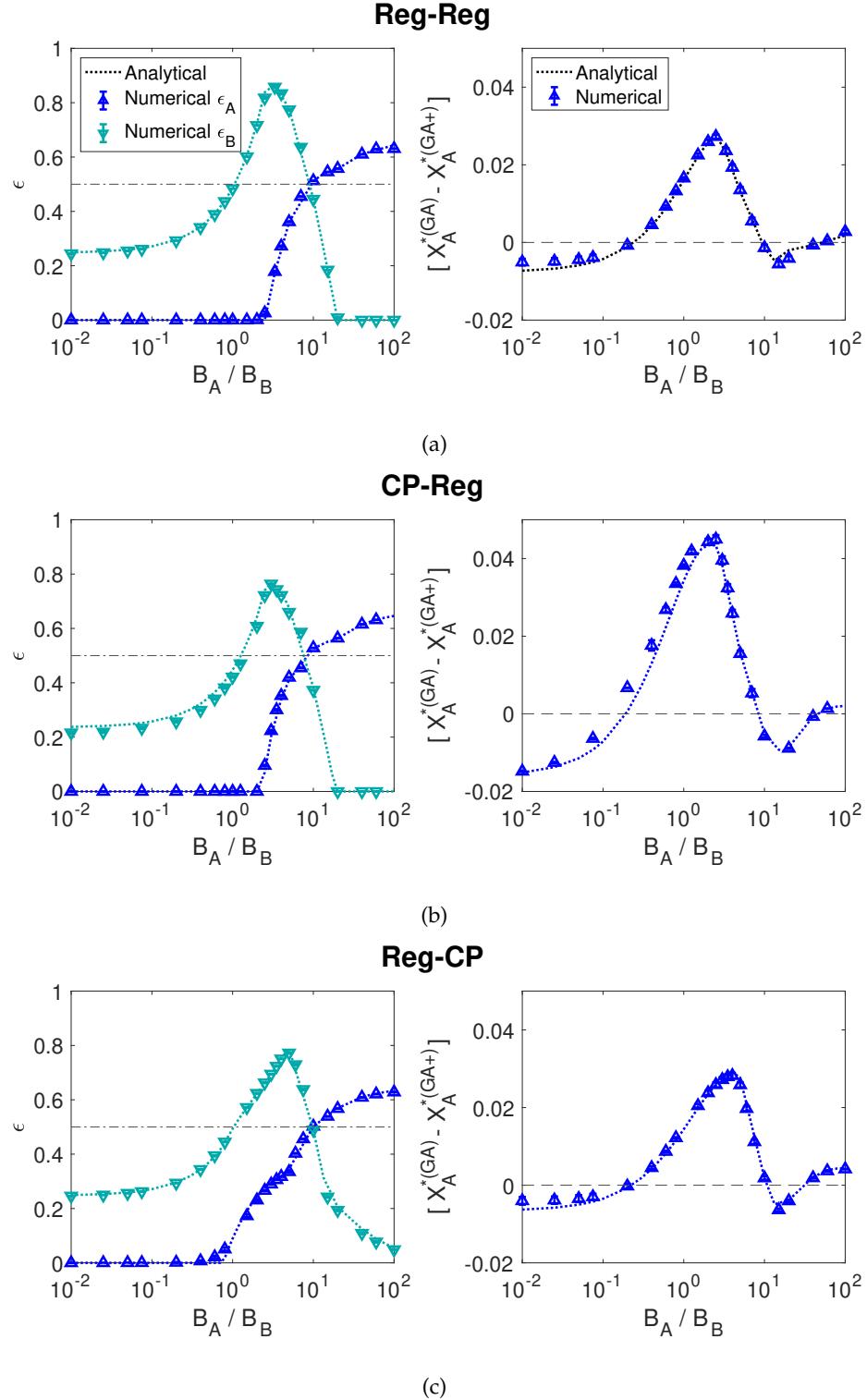
FIGURE 5.8: Figures compare analytical and numerical results of equilibrium states. Figures on the left depict change in strategies (ϵ) for both players as the budget ratio changes. Here ϵ is the fraction of resource given to nodes with negative edges. Figures on the right show gain in true vote-shares at equilibrium $[X_A^{*(GA)} - X_A^{*(GA+)}]$. Simulations are conducted in networks of size N=200, $\langle k_a \rangle = 16$, $\langle k_b \rangle = 4$ where negative ties are distributed only over half the network ($p = 0.5$). We consider three network types (a) Reg-Reg, (b) CP-Reg-High and (c) Reg-CP networks. In Reg-CP networks, $\epsilon = \epsilon_1 + \epsilon_2$, where ϵ_1 and ϵ_2 are fractions of resources given to high-negative degree and low-negative degree nodes in the negative core-periphery component respectively. Numerical simulations are run with a step-size $\eta = 5$ and terminated using a $\mu = 10^{-7}$ approximation factor. Results are averaged over 10 networks and show errorbars with 95% confidence intervals.

this chapter. For comparison, we propose a complementary approach where controllers observe only the absolute weights of all edges i.e. they consider all edges to be positive. In both instances we present gradient ascent algorithms to numerically solve the problem in large-scale arbitrary networks. We test the robustness of our results in networks of varied structures under diverse budget conditions and competitor allocations.

We find that in networks where 20% of edges are negative, controllers gain maximally (nearly 18%) from awareness of negative edges, under low budget conditions, and against competitors who deliberately avoid nodes with negative connections. We also propose a supporting analytical approach to verify the accuracy of our algorithms. Additionally, the analytical method yields closed-form solutions in simplified network structures that provide valuable insights to the problem. We draw key observations and intuitions about optimal allocations from these analytical solutions and show that in networks with highly concentrated positive links, allocations on nodes are driven by their negative degrees and the competitor's allocation on these nodes. Finally, we demonstrate the problem under game-theoretic settings, where we highlight conditions under which a controller could lose vote-shares by implementing strategies that exploit the knowledge of negative ties in the network. Specifically, we show that when controllers have considerably less resources (or in some cases, excess budget), the way they prioritise nodes to target, may inadvertently disclose additional information about the network structure to a competitor, thus compromising the controller's position of advantage.

The results in this chapter, some of which are unexpected and others more intuitive, present compelling evidence for heeding negative ties in any *influence maximisation* exercise, which fundamentally contributes to the literature on competitive opinion dynamics in signed networks. Possible extensions to this work would include studying the problem under different constraint functions. For instance, controllers may have to pay an additional cost to identify negative ties in the network, in which case the optimisation problem not only determines optimal allocations to the network but also determines how much of the budget should be optimally spent on determining negative edges in the network. Furthermore, going forward, this work could also serve as a foundation to guide empirical investigation on maximising opinion spread in the presence of negative edges in the real world.

Chapter 6

Conclusions and Future Directions

The work in this thesis explores the *influence maximisation* problem as a means to study network-based interventions in social networks. We focus on competitive settings where interventions aim to maximise the spread of a desirable behaviour by minimising the spread of competing (and undesirable) behaviours. In this work, we propose algorithms, that combine with *network theory* concepts, to determine the optimal distribution of resources that can successfully promote desired collective outcomes in populations in the presence of competition.

In the rest of the chapter, we summarise the work presented in this thesis by highlighting our results, and our contributions to the field. Finally, we close with a brief discussion on some open questions that can further research along this line of investigation.

6.1   Continuous allocation of resources

The *influence maximisation* problem in its traditional setup assumes that influence propagates in the network by means of a two-step process —a set of nodes are first "seeded" with a desired behaviour, which then spreads to the rest of the population through interpersonal interactions.

The traditional *influence maximisation* problem offers a black-box solution which only presents the most influential nodes in the network, without revealing any information about how the budget should be utilised to maximise the effect of the intervention. In this thesis, we extend this approach further by proposing a model where resources are continuously distributed. Nodes are targeted with varying intensities, based on their role in the influence spread process, therefore yielding a richer solution to the *influence maximisation* problem that explicitly shows how resources should be allocated for the maximum effect, which is a result that could be highly beneficial to policy-makers.

Additionally, the two-step spread process implicitly assumes that the network is always successfully seeded at the start of the *influence maximisation* process. The act of influencing people however, is significantly more complicated, and involves ample uncertainty. To address this concern, we modify the approach where instead of assuming that seed nodes are flipped to the desired state at the start of the process we let the dynamics evolve under a continuous application of influence on the network. The approach closely reflects many real-world settings where constant sources of information are used to inform, educate and influence people (Sela et al. (2018)).

In Chapter 3 we examine the competitive *influence maximisation* problem with continuous allocation of resources. The problem is first studied in a simplified star topology to obtain closed-form analytical expressions for optimal allocations in real-world settings. A key result obtained here shows that the optimal configurations of allocations are defined by the ratio of controller budgets. More specifically, we observe that when the focal controller has less budget in comparison to their competitor, it is optimal to avoid nodes targeted by the competitor. Whereas when the focal controller has more budget, more resources are diverted to the nodes targeted by the competitor. More importantly, we show that continuous allocation of resources consistently outperform discrete allocations to the network, irrespective of competitor allocations and budget conditions. Finally, we study the problem in a game-theoretic setting where both controllers actively target the network. We propose an algorithm to determine the optimal allocation strategy for both controllers, and we find that the equilibrium strategy is to target the network uniformly.

6.2   Nonlinear cost of influence

As highlighted earlier, the continuous distribution of resources in our model addresses the limitation of the traditional discrete setup where only the most influential nodes in the network are identified. However distributing resources heterogeneously yields a related concern —how do we quantify the effect of allocated resources to a node in terms of the strength of influence experienced by them?

As a first step (in Chapter 3), we assume that the amount of resources allocated is directly proportional to the strength of influence. However this assumption may not consistently apply to all real-world settings. For instance, some interventions take longer to be understood or adopted (e.g. adoption of green technologies). Therefore in Chapter 4 we further generalise our model by considering other nonlinear costs of allocations, i.e. we consider nonlinear relations between the cost of influence and its effect on the node.

In the first setting, we assume that a controller pays a fixed cost to prime any node in the network for influence. In the real world, this may reflect the removal of adoption

barriers in a population to facilitate the spread of a behaviour (Mobarak and Saldanha (2022); Gates (2019)). We explore this problem analytically in star networks to obtain preliminary insights about the optimal configurations of allocations and further employ these results to design heuristics for larger networks. Principally, our results show that the configuration of optimal allocations always shift from continuous to discrete allocations as the cost of access increases. Thus bridging the gap between past work, that typically considers discrete allocations, and our current model that assumes continuous distribution of resources. We further present a simple heuristic method that uses this result and show that it significantly outperforms many traditional centrality-based heuristics (including greedy methods) in real-world networks.

In addition to the above setting, we consider other instances where the effect of influence on a node increases nonlinearly with allocations. We focus on two particular instances, one which exhibits a *diminishing returns* effect on influence as allocations increase, and the other demonstrates a *delayed effect* on influence. We propose analytical and numerical methods to handle both settings in the competitive *influence maximisation* setup, and further show that our numerical results closely match our analytical results in synthetic core-periphery networks. We then employ the numerical approach to analyse a real-world network, where we illustrate the significance of employing optimal strategies as opposed to common heuristic approaches. Our main observation in both synthetic and real-world networks is that while the optimal strategy is driven by budget availability in the *delayed effect* setting, the availability of resources has limited effect on the optimal allocation pattern in the *diminishing returns* setting. The vote-share yields from optimal allocations also differ significantly across both regimes, i.e. optimal strategies offer significant increases in vote-shares in the *delayed effect* region when compared to degree-based or uniform targeting. While optimal strategies in the *diminishing returns* setting yield little to no improvement in vote-shares when compared to naïve approaches.

6.3   Presence of negative ties

In Chapter 5 we explore the competitive *influence maximisation* problem in networks with negative edges (known as signed networks). Where most prior work has focused on networks with strictly positive edges, we investigate the need to navigate antagonistic relationships in *influence maximisation* efforts under competitive conditions.

For this purpose, we propose a method that maximises influence spread in signed networks, and compare it to a traditional approach where all edges are assumed to be positive (naïve approach). We first explore the problem in a real-world signed network and show that by considering negative ties in the *influence maximisation*

process, a controller can gain nearly 9% in vote-shares. We further examine the impact of network topology, resource conditions and competitor strategies on the difference in vote-shares obtained by both methods (i.e. the negative-tie sensitive approach and the naïve approach). We show that the knowledge of negative ties in the network is highly beneficial when the competitor also has access to this information and deliberately avoids negative ties in their *influence maximisation* strategy. Conversely, we show that the controller does not gain any vote-shares when the competitor targets a network (with homogeneously distributed negative ties) uniformly. We further demonstrate that optimal allocations to nodes can be explained by their negative degrees and the competitor allocations to them, and we show that the correlation between them depends on the budget available to the controller. We finally explore the problem in a game-theoretic setting, and illustrate how the knowledge of negative ties affects outcomes in the game.

6.4   Future directions

We propose extensions to existing work in the field of competitive *influence maximisation* in this thesis. By doing so we also open up new avenues of research that can be explored beyond the scope of this thesis.

Our work here focuses on the voter model for reasons highlighted in Chapter 1. Our results therefore strictly apply to the voter dynamics and are not transferable to other dynamic models. An interesting way forward would be to apply the model proposed in this work, and its extensions to other models, such as majority dynamics, bounded-confidence models and so on, to study how results vary across other real-world instances.

Additionally, in Chapter 4, we introduced a cost of access parameter in our model to simulate the notion of adoption barriers in social intervention processes. As a first step, we assumed a fixed cost of access for all nodes in the network. While this setting gives us some interesting results, a way to take this work forward would be to consider other cost functions. For example, considering a heterogeneous cost where the cost of accessing a node is driven by its topological properties (for example, a leader node with high degree-centrality may not be open to external influence and therefore will be less accessible or have a high cost of access). It would also be interesting to study the models proposed in Chapter 4 under game-theoretic settings where apriori information about competitor allocations is not available.

In Chapter 5 we study the competitive *influence maximisation* problem in the presence of negative ties, where we show that prior knowledge of negative edges can significantly improve the outcome of network-based interventions in most settings. We assume that where controllers are aware of negative edges, they are given this

information apriori. However, it has been shown that negative ties are often difficult to observe in the real world, and often require additional exploration of the network. An interesting extension to this model would be to analyse settings where controllers have no prior information of the network structure and have to use a portion of their budget to determine the presence of negative ties. The question then becomes —how does a controller optimally divide their budget between exploration and *influence maximisation*?

Finally, the work presented in this thesis is theoretical. As pointed out in Chapter 1, research in social sciences require a balance between theory and empirical results, and thus we hope that the results in this thesis can be built upon further and used to advance empirical investigations in this field.

Appendix A

Continuous allocation of resources

A.1 Determining optimal allocations in star networks using analytical methods

A.1.1 Competitor B targets the hub

Assuming competitor B targets the hub with b, we employ Eq. (3.2) to obtain the total vote-share for controller A as

$$X_A = \frac{\frac{(n-k_A+1)((a-\alpha)\alpha+ak_A)}{(a-\alpha)\alpha+(a+b)k_A+\alpha b} + \frac{k_A\left(\frac{\alpha}{k_A}+\frac{(a-\alpha)\alpha+ak_A}{(a-\alpha)\alpha+(a+b)k_A+\alpha b}\right)}{1+\frac{\alpha}{k_A}}}{n+1}, \tag{A.1}$$

where a is the budget available to controller A. A fraction of this budget α is given to k_A peripheral nodes, i.e. each peripheral node is targeted with $\frac{\alpha}{k_A}$. The remaining budget $(a-\alpha)$ is given to the hub. Here n is the number of peripheral nodes.

Now, differentiating Eq. (A.1) wrt to k_A yields

$$\frac{\partial X_A}{\partial k_A} = \frac{(a-\alpha+b+n+1)\alpha^2 b}{(n+1)((a+b-\alpha)\alpha+(a+b)n)^2}. \tag{A.2}$$

Observe that $\frac{\partial X_A}{\partial k_A} > 0$ for any $0 < \alpha \le a$, implying that the optimal solution k_A^* lies on the boundary.

Replacing $k_A = n$ we solve $\frac{\partial X_A}{\partial \alpha} = 0$ to obtain

$$\alpha^* = (\pm\sqrt{(n+1)^2 + (a+b)} - (n+1))n. \tag{A.3}$$

A.1.2 Competitor B targets the periphery

Here competitor B targets the periphery with budget b, i.e. each peripheral node receives $\frac{b}{n}$ allocations. Assuming controller A targets k_A peripheral nodes with a fraction α of the total budget a (where $(a - \alpha)$ is allocated to the hub) in a star network with n peripheral nodes, Eq. (3.2) yields a system of N-equations that give the vote-share for each node in the network. Now, given the competitor's allocation on the network is fixed, and controller A's strategy is parameterised using (α, k_A) we obtain three groups of nodes: (i) the hub node with $p_A = a - \alpha$ and $p_B = 0$, (ii) k_A peripheral nodes targeted with $p_A = \frac{\alpha}{k_A}$ and $p_B = \frac{b}{n}$, and finally (iii) $(n - k_A)$ peripheral nodes targeted with $p_A = 0$ and $p_B = \frac{b}{n}$. The behaviour of all nodes within each group are identical and given by

$$x_{A,1} = \frac{(n+b)(((a-\alpha)\alpha + ak_A)n + bk_A(a-\alpha))}{((\alpha+k_A)b + ak_A + (a-\alpha)\alpha)n^2 + (bk_A + (a-\alpha)(2k_A+\alpha))bn + (a-\alpha)b^2k_A},$$

$$x_{A,2} = \frac{\frac{\alpha}{k} + x_{A,1}}{1 + \frac{\alpha}{k_A} + \frac{b}{n}},$$

$$x_{A,3} = \frac{x_{A,1}}{1 + \frac{b}{n}},$$

respectively.

The total vote-share is then given by

$$X_A = \frac{x_{A,1} + k_A x_{A,2} + (n - k_A)x_{A,3}}{n+1}. \tag{A.4}$$

Differentiating Eq. (A.4) wrt to k_A yields,

$$\frac{\partial X_A}{\partial k_A} = \frac{1}{n+1}\left(\frac{\partial x_{A,1}}{\partial k_A} + k_A\frac{\partial x_{A,2}}{\partial k_A} + x_{A,2} + (n-k_A)\frac{\partial x_{A,3}}{\partial k_A} - x_{A,3}\right).$$

We know that $x_{A,3} = \frac{x_{A,1}}{1+\frac{b}{n}}$. Therefore, replacing $\frac{\partial x_{A,3}}{\partial k_A} = \frac{1}{1+\frac{b}{n}}\frac{\partial x_{A,1}}{\partial k_A}$ in the above expression gives us

$$\frac{\partial X_A}{\partial k_A} = \frac{1}{n+1}\left(\left(\frac{n-k_A}{1+\frac{b}{n}}+1\right)\frac{\partial x_{A,1}}{\partial k_A} + k_A\frac{\partial x_{A,2}}{\partial k_A} + x_{A,2} - x_{A,3}\right). \tag{A.5}$$

We now look at each term individually to show that the overall expression $\frac{\partial X_A}{\partial k_A} \geq 0$. We evaluate the terms under the conditions that $a, b, n > 0$, $0 \leq \alpha \leq a$ and $0 \leq k_A \leq n$.

$$\frac{\partial x_{A,1}}{\partial k_A} = \frac{(b+n)(a-\alpha+n)\alpha^2 n^2 b}{\left(((\alpha+k_A)b + ak_A + (a-\alpha)\alpha)n^2 + (bk_A + (a-\alpha)(2k_A+\alpha))bn + (a-\alpha)b^2 k_A\right)^2}$$
$$\geq 0;$$

$$\frac{\partial x_{A,2}}{\partial k_A} = \frac{((a-\alpha+b)n + (a-\alpha)b)(a-\alpha+n)(b+n)}{\left(((\alpha+k_A)b + ak_A + (a-\alpha)\alpha)n^2 + (bk_A + (a-\alpha)(2k_A+\alpha))bn + (a-\alpha)b^2 k_A\right)^2}$$
$$> 0;$$

$$x_{A,2} - x_{A,3} = \frac{((1-x_{A,1})n + b)n\alpha}{(n+b)((\alpha+k_A)n + bk_A)}$$
$$\geq 0.$$

Therefore we show that X_A is monotonically increasing in k_A. Replacing $k_A = n$ we solve $\frac{\partial X_A}{\partial \alpha} = 0$ to obtain

$$\alpha^* = a + n + 1 \pm \sqrt{(a+b) + (n+1)^2}. \tag{A.6}$$

A.1.3 Competitor B targets uniformly

Following a similar approach as above (in Appendix A.1.2) and assuming B targets all nodes with $\frac{b}{n+1}$, we obtain three groups of nodes: (i) the hub node ($x_{A,1}$) with $p_A = a - \alpha$ and $p_B = \frac{b}{n+1}$ allocations, (ii) k_A peripheral nodes ($x_{A,2}$) targeted with $p_A = \frac{\alpha}{k_A}$ and $p_B = \frac{b}{n+1}$, and finally (iii) $(n - k_A)$ peripheral nodes ($x_{A,3}$) targeted with $p_A = 0$ and $p_B = \frac{b}{n+1}$.

The expression for total vote-share is then given by

$$X_A = \frac{x_{A,1} + k_A x_{A,2} + (n - k_A)x_{A,3}}{n+1}. \tag{A.7}$$

where

$$x_{A,1} =$$

$$\frac{(1+\beta)((a-\alpha)(a+k_A\beta)+ak_A)}{\beta^3 k_A + ((n+a-\alpha+2)k_A+\alpha)\beta^2 + ((a-\alpha)(2k_A+1)+(n+1)(k_A+1))\beta + ak_A + (a-\alpha)\alpha}$$

$$x_{A,2} =$$

$$\frac{\alpha\beta^2 + ((a-\alpha+n+1-k_A)\alpha + ak_A)\beta + ak_A + (a-\alpha)\alpha}{\beta^3 k_A + ((n+a-\alpha+2)k_A+\alpha)\beta^2 + ((a-\alpha)(2k_A+1)+(n+1)(k_A+1))\beta + ak_A + (a-\alpha)\alpha},$$

and finally $x_{A,3} = x_{A,1}$.

We assume $\beta = \frac{b}{n+1}$ and differentiate $\frac{\partial X_A}{\partial k_A}$. Note that as competitor strategy is independent of k_A it should have no effect on our results.

We obtain

$$\frac{\partial X_A}{\partial k_A} = \frac{1}{n+1}\left(\left(\frac{n-k_A}{1+\beta}+1\right)\frac{\partial x_{A,1}}{\partial k_A} + k_A\frac{\partial x_{A,2}}{\partial k_A} + x_{A,2} - x_{A,3}\right),$$

where

$$\frac{\partial x_{A,1}}{\partial k_A} =$$

$$\frac{(1+\beta)(a-\alpha+n+1+\beta)\alpha^2\beta}{\left(\beta^3 k_A + ((n+a-\alpha+2)k_A+\alpha)\beta^2 + ((a-\alpha)(2k_A+1)+(n+1)(k_A+1))\beta + ak_A + (a-\alpha)\alpha\right)^2}$$

$$\geq 0;$$

$$\frac{\partial x_{A,2}}{\partial k_A} =$$

$$\frac{(1+\beta)(a-\alpha+n+1+\beta)((a-\alpha+n+1)(\beta)+a-\alpha+(\beta)^2)(\beta)\alpha}{\left(\beta^3 k_A + ((n+a-\alpha+2)k_A+\alpha)\beta^2 + ((a-\alpha)(2k_A+1)+(n+1)(k_A+1))\beta + ak_A + (a-\alpha)\alpha\right)^2}$$

$$\geq 0;$$

$$x_{A,2} - x_{A,3} = \frac{(n+1)((1-x_{A,1})n+b+(1-x_{A,1}))\alpha}{(b+n+1)((n+1)\alpha + (b+n+1)k_A)}$$

$$\geq 0.$$

Given that $a, b, n > 0$, $0 \leq \alpha \leq a$ and $0 \leq k_A \leq n$, we obtain $\frac{\partial X_A}{\partial k_A} \geq 0$ from the above. Thus we replace $k_A = n$ and solve $\frac{\partial X_A}{\partial \alpha} = 0$ to obtain

$$\alpha^* = \frac{an}{n+1}. \tag{A.8}$$

A.2 Game-theoretic analysis

To analyse the game-theoretic scenario in a star network, we begin by parameterising the allocation vectors for both controllers A and B, such that their vote-shares are given by $X_A(\alpha, k_A)$ and $X_B(\beta, k_B)$ respectively, where α and β are the amounts of resources allocated to k_A and k_B peripheral nodes. When controller strategies are not fixed, and A targets k_A peripheral nodes and B targets k_B peripheral nodes, we get (i) $\frac{k_A k_B}{n}$ peripheral nodes targeted by both controllers, (ii) $k_A - \frac{k_A k_B}{n}$ peripheral nodes targeted only by A, (iii) $k_B - \frac{k_A k_B}{n}$ peripheral nodes targeted only by B, and (iv) $n - (k_A + kB - \frac{k_A k_B}{n})$ untargeted peripheral nodes. Based on this, we obtain the total vote-share function for A as

$$X_A = \frac{\begin{aligned} &((1+b_p)^2(a_h+\alpha)n^2 - (((k_B - b_h - 1)b_p - b_h - 1)\alpha + a_h(-1 + (k_B - 1)b_p))(1+b_p)n \\ &+ \alpha((-b_h - 1)b_p + a_h - b_h - 1)k_B b_p)k_A^2 + \alpha((1+b_p)(\alpha + a_h(b_p + 2))n^2 \\ &+ ((1+b_p)(b_h + 1)\alpha - (b_p + 2)a_h(-1 + (k_B - 1)b_p))n + k_B a_h b_p \alpha)k_A \\ &+ a_h(n(1+b_p) + 1 + (-k_B + 1)b_p)\alpha^2 n \end{aligned}}{\begin{aligned} &(((((1+b_p)\alpha + (k_B + a_h + b_h)b_p + a_h + b_h)k_A + ((k_B + a_h + b_h)b_p + a_h + b_h)\alpha) \\ &((1+b_p)k_A + \alpha)n - \alpha k_B b_p k_A((b_p + 2)k_A + \alpha))(n+1) \end{aligned}}$$

where $a_h = a - \alpha$, $b_h = b - \beta$ and $b_p = \frac{\beta}{k_B}$ are replaced to make the expressions more presentable. Note that a similar expression can be obtained for the vote-share X_B for controller B.

We obtain the partial differential as $\frac{\partial X_A}{\partial k_A}$ follows,

$$\frac{\partial X_A}{\partial k_A} = \frac{\begin{aligned} &(n - k_B)(b_p^3 k_A^2 n(b_h + k_B) + b_p^2 k_A n(2_h + 3b_h k_A + 2k_A k_B) + \alpha b_p^2 k_A k_B(2n - k_A)) \\ &+ \alpha^2 n(b_p k_B(a_h + n) + b_h n(b_p + 1)) + a_h b_p k_A k_B n(2\alpha + k_A) + 2\alpha b_p k_A n^2(2b_h + k_B) \\ &+ b_p k_A^2 n^2(3b_h + k_B) + b_h k_A n^2(2\alpha + k_A) - 2b_h b_p k_A k_B n(\alpha + k_A) \end{aligned}}{\begin{aligned} &(n+1)(((((1+b_p)\alpha + (k_B + a_h + b_h)b_p + a_h + b_h)k_A + ((k_B + a_h + b_h)b_p + a_h + b_h)\alpha) \\ &((1+b_p)k_A + \alpha)n - \alpha k_B b_p k_A((b_p + 2)k_A + \alpha))^2. \end{aligned}}$$

Observe that here $\frac{\partial X_A}{\partial k_A} \geq 0$, given that $0 < k_A \leq n$; $0 < k_B \leq n$; $\{a_h, b_h, b_p\} \geq 0$ and $n > 0$. A similar approach can be used to show that $\frac{\partial X_B}{\partial k_B} \geq 0$ (as players are inter-changeable in this game).

Replacing $k_A = n$ and $k_B = n$ in X_A and X_B, and finally solving $\{\nabla_\alpha X_A, \nabla_\beta X_B\} = 0$ simultaneously, gives us

$$\left\{\alpha^*, \beta^*\right\} = \left\{\frac{an}{n+1}, \frac{bn}{n+1}\right\}. \tag{A.9}$$

A.3   Convexity proof

Here we prove that the vote-share function is convex in competitor allocations and concave in controller allocations. To begin, we consider the vote-share function of, say competitor B, which is given by,

$$X_B = [L + diag(p_A + p_B)]^{-1} p_B. \tag{A.10}$$

Considering Eq. (A.10) is a function of p_A, we restrict the vote-share function to an arbitrary line $g(t) = X_B(p_A + tV)$ such that

$$g(t) = [L + diag(p_A + tV + p_B)]^{-1} p_B. \tag{A.11}$$

where $(p_A + tV) \in \mathbb{R}^N_+$; $t \in \mathbb{R}$ and $V \in \mathbb{R}^N$.

To check convexity, we need to obtain the second derivative, $\frac{\partial^2 g(t)}{\partial t^2}$. We begin by differentiating Eq. (A.11) wrt t to obtain,

$$[L + diag(p_A + tV + pB)]\frac{\partial g(t)}{\partial t} + diag(V)g(t) = 0. \tag{A.12}$$

Now, replacing Eq. (A.11) in Eq. (A.12), and differentiating the resulting expression wrt t, we get

$$[L + diag(p_A + tV + pB)]\frac{\partial^2 g(t)}{\partial t^2} + diag(V)\frac{\partial g(t)}{\partial t}$$
$$- [L + diag(p_A + tV + pB)]^{-1} diag(V)^2 [L + diag(p_A + tV + pB)]^{-1} p_B = 0. \tag{A.13}$$

Replacing Eqs. (A.11) and (A.12) in Eq. (A.13) and solving for $\frac{\partial^2 g(t)}{\partial t^2}$ we get,

$$\frac{\partial^2 g(t)}{\partial t^2} = 2[L + diag(p_A + tV + p_B)]^{-3} diag(V^2) p_B. \tag{A.14}$$

Note that the expression on the right yields a $N \times N$ Hessian matrix with strictly positive elements[1], thus proving that $g(t)$ is a convex function of t, and in extension X_B is a convex function of p_A (Boyd et al. (2004)). Moreover, as $X_B = (1 - X_A)$, we can conclude that X_A is a concave function of p_A, with a global maximum.

---

[1]Note that, $[L + diag(p_A + tV + p_B)]$ here is a class of M-matrices (Plemmons (1977)). The inverse of such a matrix is entity-wise nonnegative (Johnson (1982)).

Appendix B

Nonlinear costs of allocation

B.1 Analysis of the star topology

Here we illustrate results of the competitive *influence maximisation* problem —under a fixed cost of access —in the star topology for unequal budgets.
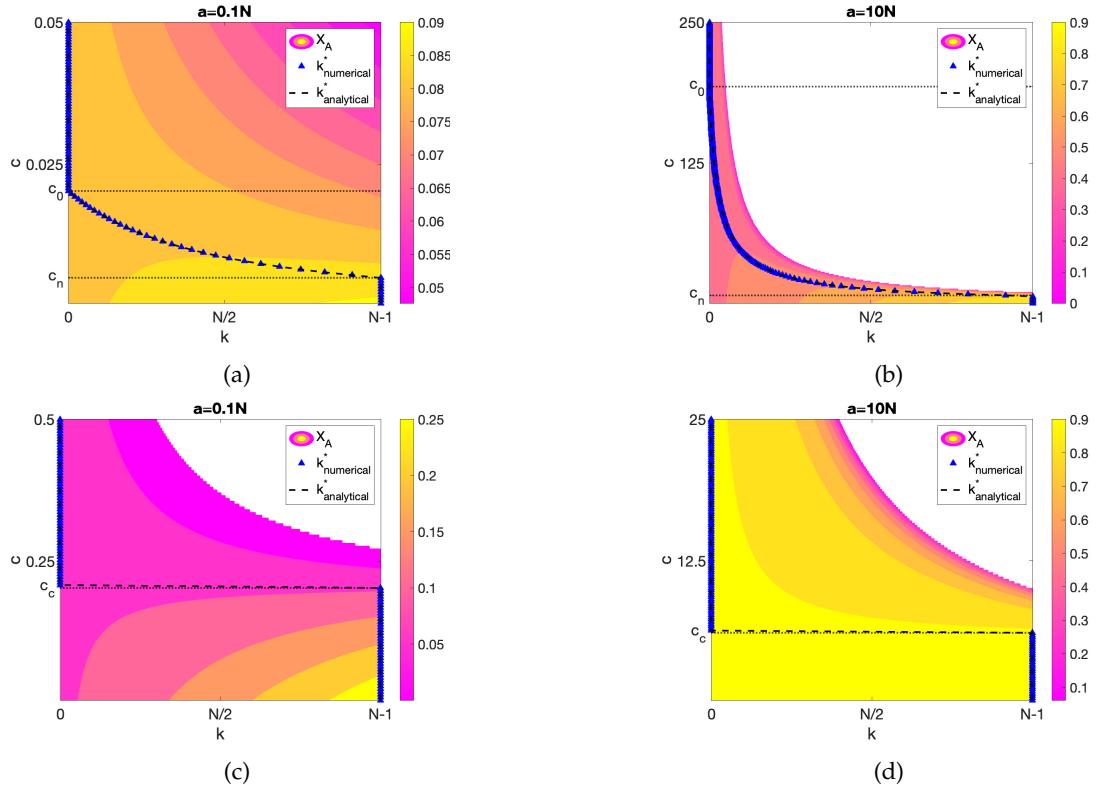
FIGURE B.1: Figure showing how optimal allocations vary with cost of access c, in a star network of size $N = 1000$. Figure (a) and (b) illustrate the instance where competitor B targets the network uniformly. Controller budget in Figure (a) is $B_A = 0.1B_B$, while in Figure (b) $B_A = 10B_B$. Competitor B targets the hub node in Figures (c) and (d). Controller has less budget $B_A = 0.1B_B$ in Figure (c), and $B_A = 10B_B$ in Figure (d). The blank regions corresponds to values that do not meet the budget constraint.

B.2   Patterns of optimal allocations in a real-world collaboration network for varying costs of access
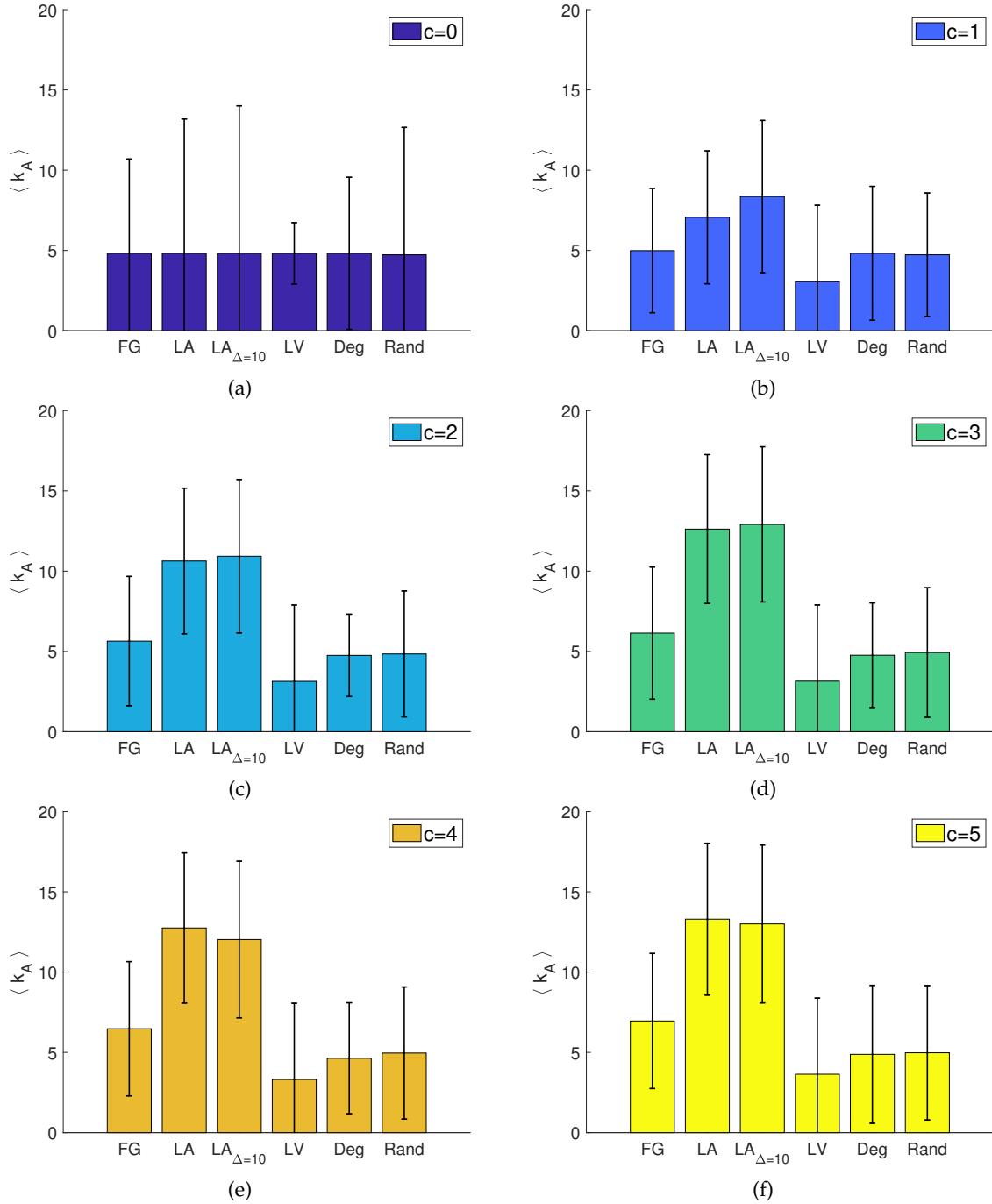
FIGURE B.2: Figure showing the optimal allocation patterns of heuristics —in terms of the average degree of targeted nodes $\langle k_A \rangle$ —for increasing costs of access, in a real-world collaboration network (Rossi and Ahmed (2015); Guimera et al. (2003)). Error-bars represent standard errors.

B.3 Algorithm for the *diminishing returns* setting

Assuming that allocation vector is incrementally changed with ρ_A, the second-order Taylor expansion of the constraint function yields,

$$\sum_i (p_{A,i} + \rho_{A,i})^\gamma \approx \sum_i^N p_{A,i}^\gamma + \gamma \sum_i p_{A,i}^{(\gamma-1)} \rho_{A,i} + \frac{\gamma(\gamma-1)}{2} p_{A,i}^{\gamma-2} \rho_{A,i}^2. \tag{B.1}$$

As $\rho_{A,i} \ll p_{A,i}$, from Eq. (B.1) we get

$$\gamma \sum_i p_{A,i}^{(\gamma-1)} \rho_{A,i} + \frac{\gamma(\gamma-1)}{2} p_{A,i}^{\gamma-2} \rho_{A,i}^2 = 0, \tag{B.2}$$

which yields the dual problem as,

$$U_A = X_A + \eta\gamma \sum_i p_{A,i}^{(\gamma-1)} \rho_{A,i} + \frac{\gamma(\gamma-1)}{2} p_{A,i}^{\gamma-2} \rho_{A,i}^2, \tag{B.3}$$

where η is the step-size.

Differentiating U_A with respect to $\rho_{A,i}$ we get,

$$\rho_{A,i} + \eta\gamma \left[p_{A,i}^{\gamma-1} + (\gamma-1) p_{A,i}^{\gamma-2} \rho_{A,i} \right] = 0, \tag{B.4}$$

which can be rearranged to obtain the optimal direction $\Delta p_{A,i}$,

$$\Delta p_{A,i} = \rho_{A,i} = -\frac{1}{\gamma-1} \left(\frac{\rho_{A,i}}{\eta\gamma p_{A,i}^{\gamma-2}} + p_{A,i} \right) \tag{B.5}$$

As the second term in Eq. (B.5) does not affect the direction, we obtain

$$\Delta p_{A,i} = \frac{\rho_{A,i}}{\eta\gamma(\gamma-1) p_{A,i}^{\gamma-2}}, \tag{B.6}$$

and therefore we update the allocation vector $p_{A,i}$ as

$$p_{A,i} \leftarrow p_{A,i} + \frac{\rho_{A,i}}{\eta p_{A,i}^{\gamma-2}}, \tag{B.7}$$

normalised at every time step to satisfy the budget constraint.

Appendix C

In the presence of negative ties

C.1 A comparative approach: removing negative ties

When controllers are unable to detect or observe negative edges in the network, i.e.
$w_{ij} = max(0, w_{ij})$, the optimisation problem reduces to

$$p_A^* = \text{argmax}_{p_A} X_A^{*(\phi)}(L^{(\phi)}, p_B, B_A). \tag{C.1}$$

where $L^{(\phi)}$ is the updated Laplacian. Following the same process as before we use the
gradient $\nabla_{p_A} X_A^{(\phi)} = 1/N \vec{1}^T [L^{(\phi)} + diag(p_A + p_B)]^{-1}(I - diag(x_A^{(\phi)})$ to optimise
allocations p_A^* in a gradient ascent algorithm $GA^{(\phi)}$. Here
$x_{A,i}^{*(\phi)} = (p_{A,i} + \sum\limits_{j}^{k_a} w_{ji} x_{A,j}^{(\phi)}) / (\sum\limits_{j}^{k_a} w_{ji} + p_{A,i} + p_{B,i})$.

We then run GA, $GA^{(+)}$ and $GA^{(\phi)}$ on the Bitcoin network and present the respective
gains in vote-shares in Fig. C.1.

We find that the method assuming negative edges in the network to be positive $GA^{(+)}$
consistently outperforms $GA^{(\phi)}$, where negative edges are not considered at all. To
further show that the vote-shares obtained through both methods are identical in
undirected networks (and that comparing our results to only $GA^{(+)}$ is sufficient), we
look at the allocation expression in Eq. (C.9). Here we find that the optimal allocation,
in the absence of any knowledge of negative edges, depends solely on the individual
budgets of the controllers and the adversarial allocations on the nodes, but not on the
degrees of the nodes.
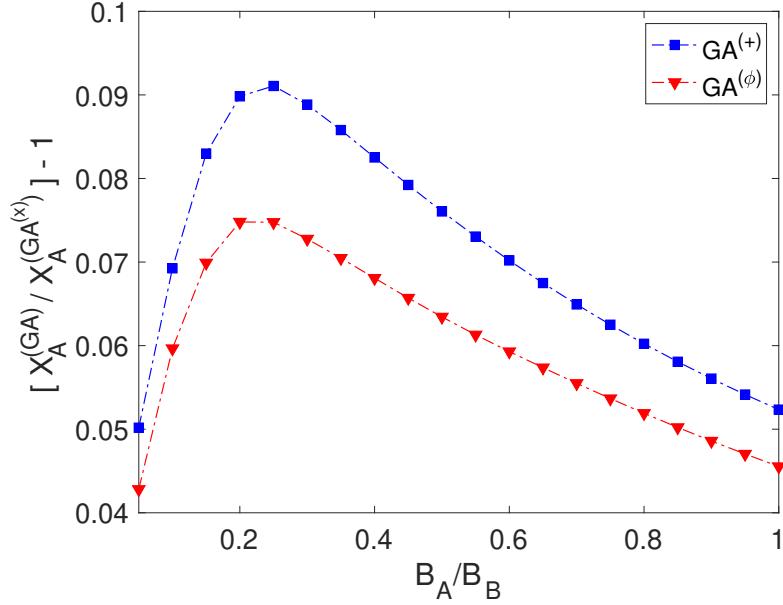
FIGURE C.1: Figure showing gain in vote-shares when comparing the negative-tie sensitive optimisation approach GA, to traditional approaches, $GA^{(+)}$ and $GA^{(\phi)}$ for budget ratios $B_A/B_B \in [0.05, 1]$. Controller B here targets the network uniformly.

C.2   Limiting case

We begin with the series expansion of the steady-state equation in Eq. (5.13) to obtain,

$$X_A = \langle \frac{a_{k_a k_b}}{k_a} \rangle + \langle \frac{k_b}{k_a} \rangle + \langle 1 - \frac{a_{k_a k_b} + b_{k_a k_b} + k_b}{k_a} \rangle \langle x_a \rangle - \langle \frac{k_b}{k_a} \rangle \langle x_b \rangle, \qquad \text{(C.2)}$$

where a second-order expansion for $\langle x_a \rangle$ gives us,

$$\langle x_a \rangle = \frac{a + \langle k_b \rangle - \langle \frac{(a+k_b)(a+b+2k_b)}{k_a} \rangle}{a + b + 2\langle k_b \rangle} + \frac{a + \langle k_b \rangle}{(a + b + 2k_b)^2} \left(\langle \frac{k_b^2}{k_a} \rangle + \langle \frac{(a + b + k_b)(a + b + 3k_b)}{k_a} \rangle \right),$$

$$\text{(C.3)}$$

and a zero-order expansion for $\langle x_b \rangle$ gives us

$$\langle x_b \rangle = \frac{a + \langle k_b \rangle}{a + b + \langle k_b \rangle}. \qquad \text{(C.4)}$$

Finally, replacing Eqs. (C.3) and (C.4) in Eq. (C.2) and ignoring higher-order terms, we obtain

$$
X_A \approx \langle \frac{a_{k_a k_b}}{k_a} \rangle + \langle \frac{k_b}{k_a} \rangle + \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle} - \frac{\langle \frac{(a_{k_a k_b} + k_b)(a_{k_a k_b} + b_{k_a k_b} + 2k_b)}{k_a} \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle}
$$

$$
+ \frac{a_{k_a k_b} + \langle k_b \rangle}{(a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle)^2} \left(\langle \frac{k_b^2}{k_a} \rangle + \langle \frac{(a_{k_a k_b} + b_{k_a k_b} + k_b)(a_{k_a k_b} + b_{k_a k_b} + 3k_b)}{k_a} \rangle \right. \tag{C.5}
$$

$$
\left. - \langle \frac{a_{k_a k_b} + b_{k_a k_b} + k_b}{k_a} \rangle \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + 2\langle k_b \rangle} - \langle \frac{k_b}{k_a} \rangle \frac{a_{k_a k_b} + \langle k_b \rangle}{a_{k_a k_b} + b_{k_a k_b} + \langle k_b \rangle} \right).
$$

Note that all terms in Eq. (C.5) are averaged over the joint positive and negative degree distribution $P_{k_a k_b}$.

We can now apply the Lagrange method to maximise vote-shares X_A against a passive controller B. The Lagrangian is derived as $\mathcal{L} = X_A + \lambda(\sum_{k_a k b} P_{k_a k_b} a_{k_a k_b} - \langle a_{k_a k_b} \rangle N)$, where λ is the Lagrangian multiplier. Differentiating $\mathcal{L}$ wrt allocations $a_{k_a k_b}$ we obtain

$$
\frac{\partial \mathcal{L}}{\partial a_{k_a k_b}} = \frac{1}{k_a} - \frac{2a_{k_a k_b} + b_{k_a k_b} + 3k_b}{\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle} \frac{1}{k_a} + 2\frac{\langle a_{k_a k_b} \rangle + \langle k_b \rangle}{(\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle)^2} \frac{a_{k_a k_b} + b_{k_a k_b} + 2k_b}{k_a}
$$

$$
- \frac{\langle a_{k_a k_b} \rangle + 2\langle k_b \rangle}{\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle} \frac{1}{k_a} + \lambda = 0. \tag{C.6}
$$

Solving for $a_{k_a k_b}$ gives us,

$$
\implies a_{k_a k_b} = \frac{1}{2} \left(\frac{\langle a_{k_a k_b} \rangle - \langle b_{k_a k_b} \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} b_{k_a k_b} + \frac{\langle a_{k_a k_b} \rangle - 3\langle b_{k_a k_b} \rangle - 2\langle k_b \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} k_b \right.
$$

$$
\left. + \langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle + \langle k_a \rangle \frac{\lambda(\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle)^2}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} \right), \tag{C.7}
$$

which still contains the Lagrangian multiplier λ. To appropriately deal with λ we average over Eq. (C.7) and assume the budget per node $\langle a_{k_a k_b} \rangle$ is sufficiently large. Therefore $\frac{\lambda(\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle)^2}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} \to 0$, which finally gives us the expression for the optimal allocation,

$$
a_{k_a k_b}^* = \frac{1}{2} \left(\frac{\langle a_{k_a k_b} \rangle - \langle b_{k_a k_b} \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} b_{k_a k_b} + \frac{\langle a_{k_a k_b} \rangle - 3\langle b_{k_a k_b} \rangle - 2\langle k_b \rangle}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle} k_b + \langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle \right). \tag{C.8}
$$

C.3   Uniformly distributed negative edges and uniform competitor allocations

When a controller cannot observe negative edges, the expression for optimal allocation is given as,

$$a_k = \frac{1}{2}\left(\frac{\langle a_k \rangle - \langle b_k \rangle}{\langle b_k \rangle}b_k + \langle a_k \rangle + \langle b_k \rangle\right). \tag{C.9}$$

where $k = k_a + k_b$.

The final vote-share in this case $X_A^{(+)}$ is obtained by replacing Eq. (C.9) in Eq. (C.5). Gain in vote-shares can therefore be quantified as,

$$
\begin{aligned}
X_A - X_A^{(+)} = {} & \frac{1}{\langle a_{k_a k_b} \rangle + \langle b_{k_a k_b} \rangle + 2\langle k_b \rangle}\left((\langle b_{k_a k_b} \rangle + \langle k_b \rangle)\langle\frac{a_{k_a k_b} - a_k}{k_a}\rangle\right. \\
& + \frac{\langle a_{k_a k_b} \rangle + \langle k_b \rangle}{\langle a_{k_a k_b} \rangle + \langle a_{k_a k_b} \rangle + 2\langle k_b \rangle}\langle\frac{(a_{k_a k_b} - a_k)(a_{k_a k_b} + a_k + 2\langle b_{k_a k_b} \rangle + 4k_b)}{k_a}\rangle \\
& \left. - \langle\frac{(a_{k_a k_b} - a_k)(a_{k_a k_b} + a_k + \langle b_{k_a k_b} \rangle + 3k_b)}{k_a}\rangle\right).
\end{aligned}
$$

Furthermore, the term $a_{k_a k_b} - a_k$ in the above expression can be derived using Eqs. (C.8) and (C.9) as,

$$a_{k_a k_b} - a_k = \left(1 - \frac{\langle a_{k_a k_b} \rangle - \langle b_{k_a k_b} \rangle}{2\langle b_{k_a k_b} \rangle}\frac{b_{k_a kb}}{\langle b_{k_a k_b} \rangle + \langle k_b \rangle}\right)\langle k_b \rangle + \frac{\langle a_{k_a k_b} \rangle - 3\langle b_{k_a k_b} \rangle - 2\langle k_b \rangle}{2(\langle b_{k_a k_b} \rangle + \langle k_b \rangle)}k_b. \tag{C.10}$$

We consider networks with regular negative graphs, $k_b = \langle k_b \rangle$ where an adversary uniformly targets the network, $b_{k_a k_b} = \langle b_{k_a k_b} \rangle$. The above relations further simplify Eq. (C.10) as,

$$a_{k_a k_b} - a_k = \left(\left(1 - \frac{\langle a_{k_a k_b} \rangle - \langle b_{k_a k_b} \rangle}{2(\langle b_{k_a k_b} \rangle + \langle k_b \rangle)}\right) + \frac{\langle a_{k_a k_b} \rangle - 3\langle b_{k_a k_b} \rangle - 2\langle k_b \rangle}{2(\langle b_{k_a k_b} \rangle + \langle k_b \rangle)}\right)\langle k_b \rangle = 0.$$

Therefore, it follows that gain $X_A - X_A^{(+)} = 0$, against an adversary targeting all nodes uniformly, in networks with regular negative components.

References

Jason Abaluck, Laura H Kwong, Ashley Styczynski, Ashraful Haque, Md Alamgir Kabir, Ellen Bates-Jefferys, Emily Crawford, Jade Benjamin-Chung, Shabib Raihan, Shadman Rahman, et al. Impact of community masking on covid-19: A cluster-randomized trial in bangladesh. *Science*, page eabi9069, 2021.

Daron Acemoğlu, Giacomo Como, Fabio Fagnani, and Asuman Ozdaglar. Opinion fluctuations and disagreement in social networks. *Mathematics of Operations Research*, 38(1):1–27, 2013.

Martin Aigner and S Axler. *A course in enumeration*, volume 1. Springer, 2007.

Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–36, 2017.

Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

Robert Axelrod. The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2):203–226, 1997.

Mehdi Azaouzi, Wassim Mnasri, and Lotfi Ben Romdhane. New trends in influence maximization models. *Computer Science Review*, 40:100393, 2021.

Younggue Bae and Hongchul Lee. Sentiment analysis of twitter audiences: Measuring the positive or negative influence of popular twitterers. *Journal of the American Society for Information Science and technology*, 63(12):2521–2535, 2012.

Abhijit Banerjee, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. The diffusion of microfinance. *Science*, 341(6144):1236498, 2013.

Suman Banerjee, Mamata Jenamani, and Dilip Kumar Pratihar. A survey on influence maximization in a social network. *Knowledge and Information Systems*, 62(9): 3417–3455, 2020.

Albert Barabasi. Twenty years of network science: From structure to control. In *APS March Meeting Abstracts*, volume 2019, pages S53–001, 2019.

Albert-László Barabási. Linked: The new science of networks, 2003.

Albert-László Barabási. Scale-free networks: a decade and beyond. *science*, 325(5939): 412–413, 2009.

Albert-László Barabási. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge university press, 2016.

Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272 (1-2):173–187, 1999.

Alvaro Barbero and Suvrit Sra. Modular proximal optimization for multidimensional total-variation regularization. *arXiv preprint arXiv:1411.0589*, 2014.

Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. *Dynamical processes on complex networks*. Cambridge university press, 2008.

Francis K Bell. A note on the irregularity of graphs. *Linear Algebra and its Applications*, 161:45–54, 1992.

Annika Bergström and Maria Jervelycke Belfrage. News in social media: Incidental consumption and the role of opinion leaders. *Digital journalism*, 6(5):583–598, 2018.

Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In *International workshop on web and internet economics*, pages 306–311. Springer, 2007.

Pablo Boczkowski, Eugenia Mitchelstein, and Mora Matassi. Incidental news: How young people consume news on social media. In *Proceedings of the 50th Hawaii international conference on system sciences*, 2017.

Pawel Boguslawski. *Modelling and analysing 3d building interiors with the dual half-edge data structure*. University of South Wales (United Kingdom), 2011.

Christian Borgs, Jennifer Chayes, Adam Tauman Kalai, Azarakhsh Malekian, and Moshe Tennenholtz. A novel approach to propagating distrust. In *International Workshop on Internet and Network Economics*, pages 87–105. Springer, 2010.

Allan Borodin, Yuval Filmus, and Joel Oren. Threshold models for competitive influence in social networks. In *International workshop on internet and network economics*, pages 539–550. Springer, 2010.

Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

Dan Braha and Marcus A M de Aguiar. Voting contagion: Modeling and analysis of a century of u.s. presidential elections. *PloS one*, 12(5):e0177970, 2017. ISSN 1932-6203.

Markus Brede, Valerio Restocchi, and Sebastian Stein. Resisting influence: How the strength of predispositions to resist control can change strategies for optimal opinion control in the voter model. *Frontiers in Robotics and AI*, 5(34), 2018.

Markus Brede, Valerio Restocchi, and Sebastian Stein. Transmission errors and influence maximization in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(3):033401, 2019.

David Bromell. Challenges in regulating online content. In *Regulating Free Speech in a Digital Age*, pages 29–53. Springer, 2022.

Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th international conference on World wide web*, pages 665–674, 2011.

Darrell L Butler and Martin Sellbom. Barriers to adopting technology. *Educause quarterly*, 2(1):22–28, 2002.

Guido Caldarelli. *Scale-free networks: complex webs in nature and technology*. Oxford University Press, 2007.

Tim Carnes, Chandrashekhar Nagarajan, Stefan M Wild, and Anke Van Zuylen. Maximizing influence in a competitive social network: a follower's perspective. In *Proceedings of the ninth international conference on Electronic commerce*, pages 351–360. ACM, 2007.

Claudio Castellano. Effect of network topology on the ordering dynamics of voter models. In *AIP Conference Proceedings*, volume 779, pages 114–120. AIP, 2005.

Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of modern physics*, 81(2):591, 2009a.

Claudio Castellano, Miguel A Muñoz, and Romualdo Pastor-Satorras. Nonlinear q-voter model. *Physical Review E*, 80(4):041129, 2009b.

Luca Catarinucci, Danilo De Donno, Luca Mainetti, Luca Palano, Luigi Patrono, Maria Laura Stefanizzi, and Luciano Tarricone. An iot-aware architecture for smart healthcare systems. *IEEE internet of things journal*, 2(6):515–526, 2015.

Damon Centola. The spread of behavior in an online social network experiment. *science*, 329(5996):1194–1197, 2010.

Sukankana Chakraborty, Sebastian Stein, Markus Brede, Ananthram Swami, Geeth de Mel, and Valerio Restocchi. Competitive influence maximisation using voting dynamics. In *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 978–985, 2019.

Georgios C Chasparis and Jeff S Shamma. Control of preferences in social networks. In *49th IEEE Conference on Decision and Control (CDC)*, pages 6651–6656. IEEE, 2010.

Shubo Chen and Kejing He. Influence maximization on signed social networks with integrated pagerank. In *2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity)*, pages 289–292. IEEE, 2015.

Wei Chen, Alex Collins, Rachel Cummings, Te Ke, Zhenming Liu, David Rincon, Xiaorui Sun, Yajun Wang, Wei Wei, and Yifei Yuan. Influence maximization in social networks when negative opinions may emerge and propagate. In *Proceedings of the 2011 siam international conference on data mining*, pages 379–390. SIAM, 2011.

Yunmei Chen and Xiaojing Ye. Projection onto a simplex. *arXiv preprint arXiv:1101.6081*, 2011.

Suqi Cheng, Huawei Shen, Junming Huang, Guoqing Zhang, and Xueqi Cheng. Staticgreedy: solving the scalability-accuracy dilemma in influence maximization. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pages 509–518. ACM, 2013.

Nicholas A Christakis and James H Fowler. The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370–379, 2007.

Nicholas A Christakis and James H Fowler. The collective dynamics of smoking in a large social network. *New England journal of medicine*, 358(21):2249–2258, 2008.

Christina Christodoulakis, Azin Asgarian, and Steve Easterbrook. Barriers to adoption of information technology in healthcare. In *Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering*, pages 66–75, 2017.

Andrew Clark and Radha Poovendran. Maximizing influence in competitive environments: A game-theoretic approach. In *International Conference on Decision and Game Theory for Security*, pages 151–162. Springer, 2011.

Peter Clifford and Aidan Sudbury. A model for spatial conflict. *Biometrika*, 60(3): 581–588, 1973.

Adam Coutts, David Stuckler, Rajaie Batniji, Sharif Ismail, Wasim Maziak, and Martin McKee. The arab spring and health: two years on. *International journal of health services*, 43(1):49–60, 2013.

Andrew T Crooks and Alison J Heppenstall. Introduction to agent-based modelling. In *Agent-based models of geographical systems*, pages 85–105. Springer, 2012.

Arnaud De Bruyn and Gary L Lilien. A multi-stage model of word-of-mouth influence through viral marketing. *International journal of research in marketing*, 25(3): 151–163, 2008.

Domenico De Stefano, Maria Prosperina Vitale, and Susanna Zaccarin. The scientific collaboration network of italian academic statisticians. In *45th Scientific Meeting of the Italian Statistical Society, Padua, Italy*, pages 16–18, 2010.

Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04):87–98, 2000.

Niveditha Devasenapathy, Suparna Ghosh Jerath, Saket Sharma, Elizabeth Allen, Anuraj H Shankar, and Sanjay Zodpey. Determinants of childhood immunisation coverage in urban poor settlements of delhi, india: a cross-sectional study. *BMJ Open*, 6(8), 2016. ISSN 2044-6055. . URL https://bmjopen.bmj.com/content/6/8/e013015.

Pedro Domingos and Matt Richardson. Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66, 2001.

Eve Dubé, Maryline Vivion, and Noni E MacDonald. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. *Expert review of vaccines*, 14(1):99–117, 2015.

David Easley, Jon Kleinberg, et al. *Networks, crowds, and markets*, volume 8. Cambridge university press Cambridge, 2010.

Sergey Edunov, Carlos Diuk, Ismail Onur Filiz, Smriti Bhagat, and Moira Burke. Three and a half degrees of separation. *Research at Facebook*, 694, 2016.

Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

Paul Erdos and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

Susanna Esposito, Nicola Principi, and Giuseppe Cornaglia. Barriers to the vaccination of children and adolescents and possible solutions. *Clinical Microbiology and Infection*, 20:25–31, 2014.

Arastoo Fazeli and Ali Jadbabaie. Game theoretic analysis of a strategic model of competitive contagion and product adoption in social networks. In *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pages 74–79. IEEE, 2012.

Arastoo Fazeli, Amir Ajorlou, and Ali Jadbabaie. Competitive diffusion in social networks: Quality or seeding? *IEEE Transactions on Control of Network Systems*, 4(3): 665–675, 2016.

Rick Ferguson. Word of mouth and viral marketing: taking the temperature of the hottest trends in marketing. *Journal of consumer marketing*, 2008.

Martin Flintham, Christian Karner, Khaled Bachour, Helen Creswick, Neha Gupta, and Stuart Moran. Falling for fake news: investigating the consumption of news via social media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2018.

David R Fortin and Ruby Roy Dholakia. Interactivity and vividness effects on social presence and involvement with a web-based advertisement. *Journal of business research*, 58(3):387–396, 2005.

Drew Fudenberg and Jean Tirole. *Game theory*. MIT press, 1991.

Satoru Fujishige. *Submodular functions and optimization*. Elsevier, 2005.

Serge Galam. Application of statistical physics to politics. *Physica A: Statistical mechanics and its applications*, 274(1-2):132–139, 1999.

Serge Galam. Minority opinion spreading in random geometry. *The European Physical Journal B-Condensed Matter and Complex Systems*, 25(4):403–406, 2002.

Serge Galam. Sociophysics: a personal testimony. *Physica A: Statistical Mechanics and its Applications*, 336(1-2):49–55, 2004.

Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 world wide web conference*, pages 913–922, 2018.

Melinda Gates. *The moment of lift: How empowering women changes the world*. Flatiron Books, 2019.

Nancy Girdhar and KK Bharadwaj. Signed social networks: a survey. In *International Conference on Advances in Computing and Data Sciences*, pages 326–335. Springer, 2016.

Daniel G Goldstein, R Preston McAfee, and Siddharth Suri. The effects of exposure time on memory of display advertisements. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 49–58. ACM, 2011.

Sanjeev Goyal, Hoda Heidari, and Michael Kearns. Competitive contagion in networks. *Games and Economic Behavior*, 2014.

Mark Granovetter. Threshold models of collective behavior. *American journal of sociology*, 83(6):1420–1443, 1978.

Boris L Granovsky and Neal Madras. The noisy voter model. *Stochastic Processes and their applications*, 55(1):23–43, 1995.

Rosanna E Guadagno, Adam Lankford, Nicole L Muscanell, Bradley M Okdie, and Debra M McCallum. Social influence in the online recruitment of terrorists and terrorist sympathizers: Implications for social psychology research. *Revue internationale de psychologie sociale*, 23(1):25–56, 2010.

Ramanthan Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web*, pages 403–412, 2004.

Roger Guimera, Leon Danon, Albert Diaz-Guilera, Francesc Giralt, and Alex Arenas. Self-similar community structure in a network of human interactions. *Physical review E*, 68(6):065103, 2003.

Michael Haenlein and Barak Libai. Targeting revenue leaders for a new product. *Journal of Marketing*, 77(3):65–80, 2013.

Nicholas Harrigan and Janice Yap. Avoidance in negative ties: Inhibiting closure, reciprocity, and homophily. *Social Networks*, 48:126–141, 2017.

Xiaochen He, Haifeng Du, Marcus W Feldman, and Guangyu Li. Information diffusion in signed networks. *PloS one*, 14(10):e0224177, 2019.

Xinran He, Guojie Song, Wei Chen, and Qingye Jiang. Influence blocking maximization in social networks under the competitive linear threshold model. In *Proceedings of the 2012 siam international conference on data mining*, pages 463–474. SIAM, 2012.

Rainer Hegselmann, Ulrich Krause, et al. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation*, 5 (3), 2002.

Sven Heidenreich and Tobias Kraemer. Innovations—doomed to fail? investigating strategies to overcome passive innovation resistance. *Journal of Product Innovation Management*, 33(3):277–297, 2016.

Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4): 599–653, 2000.

Laurie Collier Hillstrom. *The# metoo movement*. ABC-CLIO, 2018.

Pru Hobson-West. Understanding vaccination resistance: moving beyond risk. *Health, risk & society*, 5(3):273–283, 2003.

Richard A Holley and Thomas M Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability*, pages 643–663, 1975.

Peter J Hotez. Texas and its measles epidemics. *PLoS medicine*, 13(10):e1002153, 2016.

Rui-Ting Huang, David M Deggs, M Khata Jabor, and Krisanna Machtmes. Faculty online technology adoption: The role of management support and organizational climate. *Online Journal of Distance Learning Administration*, 14(2):1–11, 2011.

Matthew O Jackson. *Social and economic networks*. Princeton university press, 2010.

Prateek Jain, Purushottam Kar, et al. Non-convex optimization for machine learning. *Foundations and Trends® in Machine Learning*, 10(3-4):142–363, 2017.

Kathy Shadle James. People who were obese tried diets but felt they needed ongoing support to empower them to make lifestyle changescommentary. *Evidence-based nursing*, 12(3):92–92, 2009.

Marco A Janssen. Agent-based modelling. *Modelling in ecological economics*, 155(1): 172–181, 2005.

Charles R Johnson. Inverse m-matrices. *Linear Algebra and its Applications*, 47:195–216, 1982.

Weijia Ju, Ling Chen, Bin Li, Wei Liu, Jun Sheng, and Yuwei Wang. A new algorithm for positive influence maximization in signed networks. *Information Sciences*, 512: 1571–1591, 2020.

Elihu Katz. The two-step flow of communication: An up-to-date report on an hypothesis. *Public opinion quarterly*, 21(1):61–78, 1957.

David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146, 2003.

William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

Tamara Kharroub and Ozen Bas. Social media and protests: An examination of twitter images of the 2011 egyptian revolution. *New Media & Society*, 18(9):1973–1992, 2016.

István Z Kiss, Joel C Miller, Péter L Simon, et al. Mathematics of epidemics on networks. *Cham: Springer*, 598, 2017.

Martin Knapp, Michelle Funk, Claire Curran, Martin Prince, Margaret Grigg, and David McDaid. Economic barriers to better mental health practice and policy. *Health policy and planning*, 21(3):157–170, 2006.

Paul L Krapivsky and Sidney Redner. Dynamics of majority rule in two-state interacting spin systems. *Physical Review Letters*, 90(23):238701, 2003.

Andreas Krause and Daniel Golovin. Submodular function maximization. *Tractability*, 3:71–104, 2014.

Chris J Kuhlman, VS Anil Kumar, Madhav V Marathe, SS Ravi, and Daniel J Rosenkrantz. Finding critical nodes for inhibiting diffusion of complex contagions in social networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 111–127. Springer, 2010.

Chris J Kuhlman, VS Anil Kumar, and SS Ravi. Controlling opinion propagation in online networks. *Computer Networks*, 57(10):2121–2132, 2013.

Srijan Kumar, Francesca Spezzano, VS Subrahmanian, and Christos Faloutsos. Edge weight prediction in weighted signed networks. In *Data Mining (ICDM), 2016 IEEE 16th International Conference on*, pages 221–230. IEEE, 2016.

Srijan Kumar, Bryan Hooi, Disha Makhija, Mohit Kumar, Christos Faloutsos, and VS Subrahmanian. Rev2: Fraudulent user prediction in rating platforms. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 333–341. ACM, 2018.

Kenneth Lange. *MM optimization algorithms*. SIAM, 2016.

Laurent M Lapierre and Nicholas Bremner. Reversing the lens: How can followers influence their leader's behavior? 2010.

Bibb Latané. The psychology of social impact. *American psychologist*, 36(4):343, 1981.

Harold J Leavitt. Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, 46(1):38, 1951.

Joy Leopold and Myrtle P Bell. News media and the racialization of protest: An analysis of black lives matter articles. *Equality, Diversity and Inclusion: An International Journal*, 2017.

Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 420–429, 2007.

Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Signed networks in social media. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1361–1370, 2010.

Dong Li, Zhi-Ming Xu, Nilanjan Chakraborty, Anika Gupta, Katia Sycara, and Sheng Li. Polarity related influence maximization in signed social networks. *PloS one*, 9(7): e102199, 2014.

Dong Li, Cuihua Wang, Shengping Zhang, Guanglu Zhou, Dianhui Chu, and Chong Wu. Positive influence maximization in signed social networks based on simulated annealing. *Neurocomputing*, 260:69–78, 2017.

Yanhua Li, Wei Chen, Yajun Wang, and Zhi-Li Zhang. Influence diffusion dynamics and influence maximization in social networks with friend and foe relationships. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 657–666, 2013.

Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. Influence maximization on social graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 30(10): 1852–1872, 2018.

Wenxin Liang, Chengguang Shen, Xiao Li, Ryo Nishide, Ian Piumarta, and Hideyuki Takada. Influence maximization in signed social networks with opinion formation. *IEEE Access*, 7:68837–68852, 2019.

Wei Liu, Xin Chen, Byeungwoo Jeon, Ling Chen, and Bolun Chen. Influence maximization on signed networks under independent cascade model. *Applied Intelligence*, 49(3):912–928, 2019.

László Lovász. Submodular functions and convexity. In *Mathematical programming the state of the art*, pages 235–257. Springer, 1983.

Christopher Lynn and Daniel D Lee. Maximizing influence in an ising network: A mean-field optimal solution. In *Advances in Neural Information Processing Systems*, pages 2495–2503, 2016.

Axel Maireder and Christian Schwarzenegger. A movement of connected individuals: Social media in the austrian student protests 2009. *Information, Communication & Society*, 15(2):171–195, 2012.

Philip M Massey, Matthew D Kearney, Michael K Hauer, Preethi Selvan, Emmanuel Koku, and Amy E Leader. Dimensions of misinformation about the hpv vaccine on instagram: Content and network analysis of social media characteristics. *Journal of medical Internet research*, 22(12):e21451, 2020.

Antonia Maria Masucci and Alonso Silva. Strategic resource allocation for competitive influence in social networks. In *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 951–958. IEEE, 2014.

Naoki Masuda. Opinion control in complex networks. *New Journal of Physics*, 17(3): 033031, 2015.

Naoki Masuda, Nicolas Gibert, and Sidney Redner. Heterogeneous voter models. *Physical Review E*, 82(1):010103, 2010.

Kimberly A Maxwell. Friends: The role of peer influence across adolescent risk behaviors. *Journal of Youth and adolescence*, 31(4):267–277, 2002.

Edward Mills, Alejandro R Jadad, Cory Ross, and Kumanan Wilson. Systematic review of qualitative studies exploring parental beliefs and attitudes toward childhood vaccination identifies common barriers to vaccination. *Journal of clinical epidemiology*, 58(11):1081–1088, 2005.

Ahmed Mushfiq Mobarak and Neela A Saldanha. Remove barriers to technology adoption for people in poverty. *Nature Human Behaviour*, 6(4):480–482, 2022.

Mauro Mobilia. Does a single zealot affect an infinite group of voters? *Physical review letters*, 91(2):028701, 2003.

Mauro Mobilia. Nonlinear q-voter model with inflexible zealots. *Physical Review E*, 92 (1):012803, 2015.

Mauro Mobilia, Anna Petersen, and Sidney Redner. On the role of zealotry in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08029, 2007.

Sue Moon. Analysis of twitter unfollow: how often do people unfollow in twitter and why? In *Proceedings of the Third international conference on Social informatics*, pages 7–7. Springer-Verlag, 2011.

Jacob Levy Moreno. Who shall survive?: A new approach to the problem of human interrelations. 1934.

Megan A Moreno and Jennifer M Whitehill. Influence of social media on alcohol use in adolescents and young adults. *Alcohol research: current reviews*, 36(1):91, 2014.

Mark Newman. *Networks*. Oxford university press, 2018.

Mark Newman, Albert-László Barabási, and Duncan J Watts. *The structure and dynamics of networks*. Princeton university press, 2011.

Mark EJ Newman. Models of the small world. *Journal of Statistical Physics*, 101(3): 819–841, 2000.

Mark EJ Newman. The structure and function of complex networks. *SIAM review*, 45 (2):167–256, 2003.

Marcella Nicolini and Massimo Tavoni. Are renewable energy subsidies effective? evidence from europe. *Renewable and Sustainable Energy Reviews*, 74:412–423, 2017.

Shira Offer. Negative social ties: Prevalence and consequences. *Annual Review of Sociology*, 47, 2021.

Clare Oliver-Williams, Elizabeth Brown, Sara Devereux, Cassandra Fairhead, Isaac
    Holeman, et al. Using mobile phones to improve vaccination uptake in 21 low-and
    middle-income countries: systematic review. *JMIR mHealth and uHealth*, 5(10):e7792,
    2017.

Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in
    scale-free networks. *Physical review letters*, 86(14):3200, 2001.

Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro
    Vespignani. Epidemic processes in complex networks. *Reviews of modern physics*, 87
    (3):925, 2015.

Gordon Pennycook, Tyrone Cannon, and David G Rand. Prior exposure increases
    perceived accuracy of fake news. 2018.

Antonio F Peralta, Adrián Carro, M San Miguel, and R Toral. Analytical and
    numerical study of the non-linear noisy voter model on complex networks. *Chaos:
    An Interdisciplinary Journal of Nonlinear Science*, 28(7):075516, 2018.

Jürgen Pfeffer, Thomas Zorbach, and Kathleen M Carley. Understanding online
    firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of
    Marketing Communications*, 20(1-2):117–128, 2014.

Robert J Plemmons. M-matrix characterizations. i—nonsingular m-matrices. *Linear
    Algebra and its applications*, 18(2):175–188, 1977.

B Aditya Prakash, Alex Beutel, Roni Rosenfeld, and Christos Faloutsos. Winner takes
    all: competing viruses or ideas on fair-play networks. In *Proceedings of the 21st
    international conference on World Wide Web*, pages 1037–1046, 2012.

Hamish Pringle and Les Binet. How marketers can use celebrities to sell more
    effectively. *Journal of Consumer Behaviour: An International Research Review*, 4(3):
    201–214, 2005.

Suhas Ranganath, Xia Hu, Jiliang Tang, and Huan Liu. Understanding and identifying
    advocates for political campaigns on social media. In *Proceedings of the Ninth ACM
    International Conference on Web Search and Data Mining*, pages 43–52, 2016.

Chamil Rathnayake and Daniel D Suthers. Networked solidarity: An exploratory
    network perspective on twitter activity related to# illridewithyou. In *2016 49th
    Hawaii International Conference on System Sciences (HICSS)*, pages 2058–2067. IEEE,
    2016.

Sidney Redner. Reality-inspired voter models: A mini-review. *Comptes Rendus
    Physique*, 20(4):275–292, 2019.

Lindsay Redpath. Confronting the bias against on-line learning in management
    education. *Academy of Management Learning & Education*, 11(1):125–140, 2012.

Julie Ricard and Juliano Medeiros. Using misinformation as a political weapon: Covid-19 and bolsonaro in brazil. *Harvard Kennedy School Misinformation Review*, 1 (3), 2020.

George Ritzer et al. *The Blackwell encyclopedia of sociology*, volume 1479. Blackwell Publishing New York, NY, USA, 2007.

Everett M Rogers. *Diffusion of innovations*. Simon and Schuster, 2010.

Puck Rombach, Mason A Porter, James H Fowler, and Peter J Mucha. Core-periphery structure in networks (revisited). *SIAM review*, 59(3):619–646, 2017.

Daniel M Romero, Christopher M Kribs-Zaleta, Anuj Mubayi, and Clara Orbe. An epidemiological approach to the spread of political third parties. 2009.

Guillermo Romero Moreno, Sukankana Chakraborty, and Markus Brede. Shadowing and shielding: Effective heuristics for continuous influence maximisation in the voting dynamics. *Plos one*, 16(6):e0252515, 2021.

Jon Roozenbeek, Claudia R Schneider, Sarah Dryhurst, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. Susceptibility to misinformation about covid-19 around the world. *Royal Society open science*, 7(10):201199, 2020.

J Ben Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

J Niels Rosenquist, Joanne Murabito, James H Fowler, and Nicholas A Christakis. The spread of alcohol consumption behavior in a large social network. *Annals of internal medicine*, 152(7):426–433, 2010.

Ryan A. Rossi and Nesreen K. Ahmed. The network data repository with interactive graph analytics and visualization. In *AAAI*, 2015. URL http://networkrepository.com.

Eleni Sardianou and P Genoudi. Which factors affect the willingness of consumers to adopt renewable energies? *Renewable energy*, 57:1–4, 2013.

Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.

Andrés Scherman, Arturo Arriagada, and Sebastián Valenzuela. Student and environmental protests in chile: The role of social media. *Politics*, 35(2):151–171, 2015.

Frank Schweitzer and Laxmidhar Behera. Neighborhood approximations for non-linear voter models. *Entropy*, 17(11):7658–7679, 2015.

Alon Sela, Dmitri Goldenberg, Irad Ben-Gal, and Erez Shmueli. Active viral marketing: Incorporating continuous active seeding efforts into the diffusion model. *Expert Systems with Applications*, 107:45–60, 2018.

Parongama Sen and Bikas K Chakrabarti. *Sociophysics: an introduction*. Oxford University Press, 2014.

Chengguang Shen, Ryo Nishide, Ian Piumarta, Hideyuki Takada, and Wenxin Liang. Influence maximization in signed social networks. In *International Conference on Web Information Systems Engineering*, pages 399–414. Springer, 2015.

Alessandro Siani. Measles outbreaks in italy: A paradigm of the re-emergence of vaccine-preventable diseases in developed countries. *Preventive medicine*, 121: 99–104, 2019.

Leslie B Snyder. Health communication campaigns and their impact on behavior. *Journal of nutrition education and behavior*, 39(2):S32–S40, 2007.

Adriana Solovei and Bas van den Putte. The effects of five public information campaigns: The role of interpersonal communication. *Communications*, 45(s1): 586–602, 2020.

Vishal Sood and Sidney Redner. Voter model on heterogeneous graphs. *Physical review letters*, 94(17):178701, 2005.

Ajitesh Srivastava, Charalampos Chelmis, and Viktor K Prasanna. Social influence computation and maximization in signed networks with competing cascades. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, pages 41–48, 2015.

Dietrich Stauffer and Muhammad Sahimi. Discrete simulation of the dynamics of spread of extreme opinions in a society. *Physica A: Statistical Mechanics and its Applications*, 364:537–543, 2006.

Thanos G Stavropoulos, Asterios Papastergiou, Lampros Mpaltadoros, Spiros Nikolopoulos, and Ioannis Kompatsiaris. Iot wearable sensors and devices in elderly care: a literature review. *Sensors*, 20(10):2826, 2020.

Luc Steels. A self-organizing spatial vocabulary. *Artificial life*, 2(3):319–332, 1995.

Zachary C Steinert-Threlkeld et al. Spontaneous collective action: Peripheral mobilization during the arab spring. *American Political Science Review*, 111(2): 379–403, 2017.

Pieter H Streefland. Public doubts about vaccination safety and resistance against vaccination. *Health policy*, 55(3):159–172, 2001.

Katarzyna Sznajd-Weron. Sznajd model and its applications. *arXiv preprint physics/0503239*, 2005.

Shoma Tanabe and Naoki Masuda. Complex dynamics of a nonlinear voter model with contrarian agents. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 23(4): 043136, 2013.

Takuma Tanaka and Toshio Aoyagi. Optimal weighted networks of phase oscillators for synchronization. *Physical Review E*, 78(4):046210, 2008.

Youze Tang, Xiaokui Xiao, and Yanchen Shi. Influence maximization: Near-optimal time complexity meets practical efficiency. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 75–86. ACM, 2014.

Youze Tang, Yanchen Shi, and Xiaokui Xiao. Influence maximization in near-linear time: A martingale approach. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pages 1539–1554. ACM, 2015.

Samia Tasnim, Md Mahbub Hossain, and Hoimonty Mazumder. Impact of rumors and misinformation on covid-19 in social media. *Journal of preventive medicine and public health*, 53(3):171–174, 2020.

Ralph Tench and Brian Jones. Social media: the wild west of csr communications. *Social Responsibility Journal*, 2015.

Noel M Tichy, Michael L Tushman, and Charles Fombrun. Social network analysis for organizations. *Academy of management review*, 4(4):507–519, 1979.

Jeffrey Travers and Stanley Milgram. The small world problem. *Phychology Today*, 1(1): 61–67, 1967.

Joshua A Tucker, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal, and Brendan Nyhan. Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018)*, 2018.

Thomas W Valente. Network interventions. *science*, 337(6090):49–53, 2012.

Sander van der Linden. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, 28(3):460–467, 2022.

Frank Vanclay. Barriers to adoption: a general overview of the issues. *Rural Society*, 2 (2):10–12, 1992.

Federico Vazquez, Paul L Krapivsky, and Sidney Redner. Constrained opinion dynamics: Freezing and slow evolution. *Journal of Physics A: Mathematical and General*, 36(3):L61, 2003.

Alessandro Vespignani. Twenty years of network science, 2018.

Arun Vishwanath. Habitual facebook use and its impact on getting deceived on social media. *Journal of Computer-Mediated Communication*, 20(1):83–98, 2015.

Nicholas J Watkins, Cameron Nowzari, Victor M Preciado, and George J Pappas. Optimal resource allocation for competitive spreading processes on bilayer networks. *IEEE Transactions on Control of Network Systems*, 2016.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440, 1998.

Duncan J Watts, Jonah Peretti, and Michael Frumin. *Viral marketing for the real world*. Harvard Business School Pub. Boston, 2007.

Philip J Weiser. The future of internet regulation. *UC Davis L. Rev.*, 43:529, 2009.

Klaus-Peter Wiedmann, Nadine Hennigs, Lars Pankalla, Martin Kassubek, and Barbara Seegebarth. Adoption barriers and resistance to sustainable solutions in the automotive sector. *Journal of Business Research*, 64(11):1201–1206, 2011.

Bryan Wilder and Yevgeniy Vorobeychik. Controlling elections through social influence. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 265–273. International Foundation for Autonomous Agents and Multiagent Systems, 2018.

Bryan Wilder, Laura Onasch-Vera, Juliana Hudson, Jose Luna, Nicole Wilson, Robin Petering, Darlene Woo, Milind Tambe, and Eric Rice. End-to-end influence maximization in the field. In *AAMAS*, volume 18, pages 1414–1422, 2018a.

Bryan Wilder, Sze-Chuan Suen, and Milind Tambe. Preventing infectious disease in dynamic populations under uncertainty. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018b.

Jiyoung Woo, Jaebong Son, and Hsinchun Chen. An sir model for violent topic diffusion in social media. In *Intelligence and Security Informatics (ISI), 2011 IEEE International Conference on*, pages 15–19. IEEE, 2011.

Jen-Her Wu and Shu-Ching Wang. What drives mobile commerce?: An empirical evaluation of the revised technology acceptance model. *Information & management*, 42(5):719–729, 2005.

Amulya Yadav, Leandro Soriano Marcolino, Eric Rice, Robin Petering, Hailey Winetrobe, Harmony Rhoades, Milind Tambe, and Heather Carmichael. Preventing hiv spread in homeless populations using psinet. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

Amulya Yadav, Hau Chan, Albert Xin Jiang, Haifeng Xu, Eric Rice, and Milind Tambe. Using social networks to aid homeless shelters: Dynamic influence maximization under uncertainty. In *AAMAS*, volume 16, pages 740–748, 2016.

Amulya Yadav, Ritesh Noothigattu, Eric Rice, Laura Onasch-Vera, Leandro Soriano Marcolino, and Milind Tambe. Please be an influencer?: Contingency-aware influence maximization. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pages 1423–1431. International Foundation for Autonomous Agents and Multiagent Systems, 2018a.

Amulya Yadav, Bryan Wilder, Eric Rice, Robin Petering, Jaih Craddock, Amanda Yoshioka-Maxwell, Mary Hemler, Laura Onasch-Vera, Milind Tambe, and Darlene Woo. Bridging the gap between theory and practice in influence maximization: Raising awareness about hiv among homeless youth. In *IJCAI*, pages 5399–5403, 2018b.

Xiangyu Yang, Jiashan Wang, and Hao Wang. Towards an efficient approach for the nonconvex lp ball projection: algorithm and analysis. *arXiv preprint arXiv:2101.01350*, 2021.

Ercan Yildiz, Asuman Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation (TEAC)*, 1(4):1–30, 2013.

Lindsay E Young, Emily Sidnam-Mauch, Marlon Twyman, Liyuan Wang, Jackie Jingyi Xu, Matthew Sargent, Thomas W Valente, Emilio Ferrara, Janet Fulk, and Peter Monge. Disrupting the covid-19 misinfodemic with network interventions: network solutions for network problems. *American journal of public health*, 111(3):514–519, 2021.

John Zarocostas. How to fight an infodemic. *The lancet*, 395(10225):676, 2020.

Haifeng Zhang, Ariel D Procaccia, and Yevgeniy Vorobeychik. Dynamic influence maximization under increasing returns to scale. In *AAMAS*, pages 949–957. Citeseer, 2015.

Dawei Zhu, Jian Wang, and Knut Reidar Wangen. Hepatitis b vaccination coverage rates among adults in rural china: are economic barriers relevant? *Vaccine*, 32(49): 6705–6710, 2014.