

From Hostility to Hyperlinks: Mining Social Networks with Heterogenous Ties



Shazia'Ayn Virji Babul
New College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

2025

Acknowledgements

“They [human lives] are composed like music. Guided by his sense of beauty, an individual transforms a fortuitous occurrence (Beethoven’s music, death under a train) into a motif, which then assumes a permanent place in the composition of the individual’s life. Anna could have chosen another way to take her life. But the motif of death and the railway station, unforgettably bound to the birth of love, enticed her in her hour of despair with its dark beauty. Without realizing it, the individual composes his life according to the laws of beauty even in times of greatest distress. It is wrong, then, to chide the novel for being fascinated by mysterious coincidences (like the meeting of Anna, Vronsky, the railway station, and death or the meeting of Beethoven, Tomas, Tereza, and the cognac), but it is right to chide man for being blind to such coincidences in his daily life. For he thereby deprives his life a dimension of beauty.”

— Milan Kundera, *The Unbearable Lightness of Being* [119]

When I first read those lines in the summer of 2019, I had been struck for reasons I could not quite name yet. Over the course of these four years, largely due to the many people who guided me along the way, I think I have finally found the words. Like most other people who find themselves in the field of Complex Systems, I have always loved patterns. The idea that there are patterns and motifs that emerge in real life, and that we must be responsible for noticing them, else to deprive ourselves of a “dimension of beauty”; Kundera has put into words, romantically, the reason why we secretly study Network Science.

This journey began when I first met Renaud Lambiotte later that same year in 2019. I was in my final year at Princeton, and, as often happens in times of great transition, a panic had begun to set in as I had realized that I was completely and utterly lost. I had come across his research, and it had alighted my curiosity in a way that I had not experienced before. We had our first meeting online, some two years before we ever met in person, a meeting for which I am sure I was woefully unprepared. Renaud very kindly spent that hour explaining to me the basics of Network Science. That first meeting was indicative of his incredible generosity of time and generosity of knowledge, which has marked each and every one of our encounters since.

I am incredibly grateful for his careful mentorship, for his support in allowing me to explore beyond Oxford, from the Alan Turing Institute to Spotify, and for his remarkable sense of knowing when I needed guidance, of which he was always happy to offer, and when to leave me to my own explorations. While this proved daunting at times, I could not have grown as a scientist without that space, which allowed my own curiosity to flourish. It has been a privilege to learn under him.

It is no small feat, in a mathematics department, that Renaud also fostered an encouraging and welcoming environment within the our Network Science research group, equally due to the leadership

of Nicola Pedreschi and Tim LaRock. This environment can often be a hostile one, especially for someone like me, but I am acutely aware that if I suffered very little for being an outsider it was very much due to both Renaud and the two of you – thank you for your mentorship and guidance, but more so for being friends, for one too many Campari spritz, and for being excellent guides into the debauchery of Network Science conferences.

The works that Renaud and I produced together, which comprise this thesis, are the combined efforts of many other collaborators we worked with over the years; I would like to thank all of my fellow researchers in the Network Science group at Oxford, particularly Karel Devriendt and Yu Tian, with whom we authored some of the works included in this thesis. I would also like to express my gratitude to Desislava Hristova, Antonio Lima, and Mariano Beguerisse-Díaz, for their mentorship during my time at Spotify Research. Finally, thank you to Javier Garcia-Bernardo, who I met following my first conference talk, and who offered me the chance to work alongside him, Elena Candellone, Alexandre Bovet and Özgür Togay, on the project which would become the final chapter of this thesis.

* * *

I would be remiss if I did not mention that over the four years devoted to understanding the diversity of social ties, I indeed learned a great deal about the topic outside the pages of this thesis.

As I have walked along this path, which has taken me far from the world I knew best, and at times has been solitary, it has never been lonely, thanks to my friends. In particular, I would mention E Jeremijenko-Conley, Miriam Moukafih, Zoé Robin, Layla Varkey, Sanam Yar, Anna Peran and Hamza Hashem, to name but a few – who were there to catch me when I stumbled, who gifted me with their compassion, and who were my confidants and my sources of laughter during the light times, and during the dark ones, my home away from home.

To Sofia Medina, from the first moment I heard your cackle ringing throughout the atrium of the Mathematical Institute, I knew we would either be absolute enemies or fast friends. I find myself infinitely grateful that it was the latter; these four years would have been unbearable without you by my side through both the joys and perils we experienced both due to the dissertation, and the growing pains of adulthood. If it was a certainty that I would meet Sofia, it was by pure chance that we would find Valentina Antonaccio that summer day in Berlin, who fit so easily into the fabric of our lives, balancing our giddy hysteria with calmness, and our impulsiveness with wisdom. To both of you, thank you for four years of steady friendship, the silliness and sincerity, for the tears wiped away and the laughter shared.

The sticky core of the examined life is that, look too closely, and serendipity can lose all meaning, dimensions of beauty can become impossible to spot, and it can become easy to slip into cynicism.

So I am grateful to have had Iman Sheriff, Oathman Maqari, and Cal Berer for noticing it for me when I could not, for their hands on my shoulders when I slipped, and for picking up the pieces on that night of a thousand and one nights, when I was certain everything would fall apart.

Thank you to John Pougué-Biyong, for when the world seemed small from where I was standing, you made me see it as much bigger, the possibilities wider, than I could have seen by myself. From you, I take many learnings, and what a lovely thing it was to be seen.

To my sister Aliya-Nur, you think I am a maverick because you cannot see that I am merely following in your footsteps; my journey made easier, following the paths first carved out by you. I find it most difficult to put words to my gratitude so instead I recall a memory; when I was a child, scared of the dark, I would crawl into your room every night and find a mattress laid out in preparation for me to sleep on. Love is feeling around in the dark and finding a mattress laid out in preparation. This always brought me great comfort because if I was certain of one thing, it was that nothing could hurt me as long as you were around. This certainty has grown, unwaveringly strong, every day of our lives since. Thank you for your guidance, gracefully delivered though rarely received well, your courage, which I have spent my life trying to emulate, and your faith in me, such that when mine failed, yours could suffice.

Lastly, I would like to thank my parents; everything I am is because of both of you, the architects of the first motifs of my life. By the sacrifices that you made such that I have never lived a single day in lack, even if it meant you went without yourselves. First, to my father, from whom I first learned about that dimension of beauty in the patterns of the universe; *look*: when he showed me the caustics playing across our feet under the waves at the beach, or *listen*: tapping out the raag of the ghazal on his Sunday morning mug of chai. Though we followed different scientific paths for that search, he has never allowed me to forget that whatever one finds in that search to live a meaningful life, be it beauty or chaos, pain or joy, struggle or peace, that one could not exist without the other, and that they together, bring meaning. Second, to my mother at whose knee I learned to become an idealist; it is easy to mistake idealism with naivety. In reality, it takes great creativity to envision a world more just than the one we know, and great strength to work towards bringing that world into reality. Thank you for being my hope, and for reminding me of it, every time my own beliefs are shaken.

Finally, to my grandparents who left everything they knew behind to build anew in a country that was at best unfamiliar, and at worst, hostile, armed with the faith that their sacrifice would bring untold opportunity for the next generations: this is your legacy, much more than mine.

Abstract

Social networks are a fundamental tool for understanding emergent behaviour in human society, providing a mathematical framework that emphasizes the importance of interactions between the individuals in the network. While traditional social network models consider all social ties as uniform, either an edge exists or it does not, human nature is more complex and individuals can be linked by relationships that differ in nature, intensity, or sentiment. This tie-level complexity can be represented using more complex network models, including signed, weighted and multiplex networks, where edge-level attributes delineate between the types of interactions. A growing body of literature is devoted to developing methods for extracting information from such *heterogenous* networks, from probing the latent structure to investigating dynamical processes occurring overtop of them.

In this thesis, we focus on ties that vary in sentiment, using signed networks in which edges carry positive or negative weights, representing cooperative or antagonistic relationships, and ties that vary in nature, using weighted and multiplex network models. We present models and empirical studies that adapt traditional network science methods to extract information, detect multi-scale structure and characterize dynamical processes, to the heterogeneous network context. Our contributions here are four-fold. First, we present a signed network embedding algorithm capable of recovering continuous node attributes. Next, we develop a signed network random walk, which effectively captures macro-scale graph structure. Third, we conduct an empirical study of peer-to-peer music recommendations, using a model of social contagion that leverages the information contained in heterogeneous ties. Our final work draws upon the learnings from the previous chapters, presenting a study of political polarization on a social media platform, represented as a signed, weighted and multiplex network. Overall, this thesis presents methodological and empirical advances, which demonstrate the advantage of maintaining tie-level complexity in mining social networks.

Contents

1	Introduction	2
1.1	Motivation	2
1.2	Organization and Contributions	4
2	Preliminaries	6
2.1	Graphs	6
3	A proximity-based embedding for signed networks	15
3.1	Introduction	15
3.2	Related Literature	16
3.3	The SHEEP Hamiltonian	19
3.3.1	Relation to Spectral Methods	19
3.3.2	Spherical Constraint	20
3.3.3	Ground State Energy and Strong Balance	21
3.4	Weak Balance and Extension to Higher Dimensions	24
3.5	Applications	26
3.5.1	Proximity Based Node Attributes	26
3.5.2	Node Extremism Measure	31
3.6	Conclusions	35
4	Modelling random walks over signed networks	36
4.1	Introduction	36
4.2	Related Literature	38
4.3	Strong walks	40
4.3.1	Introduction and basic properties	40
4.3.2	Walk lengths	41
4.4	Strong walk based similarity matrix	42
4.5	Weak walks	43
4.5.1	Definition and basic properties	43

4.5.2	Walk length	45
4.6	Applications	46
4.6.1	Computational Design	47
4.6.2	Clustering on Empirical Networks	48
4.6.3	Semi-supervised Clustering on SSBMs	50
4.7	Conclusions	54
5	Social contagion of music taste over heterogeneous social ties	57
5.1	Introduction	57
5.2	Data and Methodology	60
5.2.1	Spotify Social Network	61
5.2.2	Link Sharing Events	61
5.2.3	Music Vector Embeddings	63
5.2.4	User-Artist Engagement	64
5.3	Results and Discussion	66
5.3.1	Music Taste Similarity	67
5.3.2	Tie Strength Effects	69
5.3.3	Social Cohesion	72
5.3.4	Predicting Post-share Engagement	74
5.4	Conclusions	77
6	Detecting ideological extremes in social media networks	80
6.1	Introduction	80
6.2	Data and Methodology	83
6.2.1	Menéame Dataset	83
6.2.2	Topic Modelling	85
6.2.3	Network Construction	87
6.2.4	Network Embeddings	88
6.2.5	News Outlet Ideology	89
6.2.6	Characterizing User Voting Behaviours	90
6.3	Results and Discussion	91
6.3.1	Mapping the Ideological Dimensions of News Outlets	91
6.3.2	User-Level Structural Polarization	93
6.4	Conclusion	103
7	Learnings and Future Directions	105

A	SHEEP Analysis and Proofs	109
A.1	Repelling Laplacian Spectra Proofs	109
A.2	Bi-Polarization Measure	113
A.3	Weak Balance and Extension to Higher Dimensions	115
A.3.1	Embedding dimension and number of clusters	117
A.4	Rankings on Generated Synthetic Graphs: Further Experiments	117
A.5	Node Extremism Measure and Degree	119
A.5.1	Congress 114 Correlations With Degree	119
B	Menéame Data and Methods	122
B.1	Network Reciprocity	122
B.2	Topic Modelling	122
B.2.1	BERTopic algorithm	122
B.2.2	hSBM Topic Model steps	123
B.2.3	Topic Modelling Comparison	123
B.3	Ideology from Twitter	124
	Bibliography	127

Statement of Originality

The works presented in this thesis are my own, but were produced under the supervision of my advisor Renaud Lambiotte, and in collaboration with other researchers, as described below.

Chapter 3 Renaud Lambiotte supervised this work, and helped me to conceptualize the proposed model and the theoretical experiments. The algorithm design, mathematical analysis, applications, and synthetic experiments were conducted by me. This work is published and available in [22].

Chapter 4 Renaud Lambiotte supervised this work, and helped me to conceptualize the proposed model. The work was also produced in collaboration with Yu Tian, who made some contributions to the theoretical development and the experimental sections. The remainder of the work, including the algorithm design, applications, and synthetic experiments were conducted by me. This work is published and available in [21].

Chapter 5 This chapter was produced during my internship with the Spotify Research team, where I was supervised by Desislava Hristova, Antonio Lima, and Mariano Beguerisse-Díaz. They helped me conceptualize the project, and, with the input of Renaud Lambiotte, refine the analysis and framing of the work. I was the lead on the project, running the models and the experiments. After completing my internship at Spotify, as I no longer had access to the Spotify data, Desislava Hristova used my codes to reproduce some of the analysis, and to anonymize the dataset for publication. This work is published and available in [106].

Chapter 6 This work was produced primarily as a collaboration between myself and Elena Candellone, and supervised by Javier Garcia-Bernardo and Alexandre Bovet. The project was conceptualized by all four of us, using a dataset curated by Javier Garcia-Bernardo and Elena Candellone. They worked on validating the topics found in the textual data, given their command of Spanish. Özgür Togay, provided context about political polarization from a social science perspective. The formal analysis was conducted as a joint effort between me and Elena Candellone. I implemented the analysis of the dataset using an algorithm I had designed (see Chapter 3) and worked on adapting the work to the empirical setting, as well as the interpretation of the results. This work is published and available in [39].

Chapter 1

Introduction

1.1 Motivation

Complex systems science concerns itself with systems composed of many interacting parts, and the patterns that *emerge* from the collective behaviour of the components [153]. Systems from a wide range of disciplines that exhibit this aptly named *emergent behaviour*, fall into the purview of complex systems science, from macro-scale phenomena, like the flocking behaviour of birds, to micro-scale interactions, such as amongst neurons in the brain. To gain insight into the dynamics of such systems, they are first expressed as mathematical models which reduce the complexity of the underlying system, while preserving essential properties. In this thesis, we consider the *network* as the base representation of the systems considered. A network, or a graph, is simply a collection of points, or *nodes*, with pairwise connections, called *edges*. Networks are fundamental to the study of complex systems, providing a mathematical framework that emphasizes the interactions between the components [152, 122].

This thesis investigates social network systems, in which nodes correspond to individuals and edges encode their pairwise interactions. Social network representations have been used to study a range of emergent phenomena in human society, from patterns encoded in the network structure, to the dynamics that occur on top of the network. Community detection, for example, probes the macro-scale network structure, detecting groups of individuals (communities) which are more closely connected to each other, than to individuals in the rest of the network [152, 178]. Dynamical processes, such as epidemic models of disease transmission, unfold on top of the social network; its structure is key to determining the speed and pattern of the spread [155].

Fundamentally, human interactions are non-uniform, and individuals comprising a social network are connected by relationships that may differ in nature, intensity, or sentiment. In its simplest form, a network representation obscures this complexity; an edge indicates the existence of a relationship, irrespective of its type [98]. Such complexity reduction is often a necessary methodological choice to probe the network structure using tools designed for traditional networks. Nevertheless, this

inevitably results in the loss of a crucial layer of information. Retaining the information describing diverse social ties lies at the core of this thesis; we present a series of methodologies and studies that extend traditional network science approaches to social networks with heterogeneous tie information, with the aim of extracting latent information, detecting multi-scale structure, or characterizing the diffusion dynamics occurring on top of the network.

This thesis focuses on ties that differ along two dimensions; *sentiment* and *nature*, where sentiment introduces theoretical constraints distinct from other forms of edge heterogeneity. While both dimensions represent forms of network heterogeneity, we organize our methodological contributions by addressing them separately, developing methods for sentiment-based heterogeneity in Chapters 3 and 4, and for heterogeneity in interaction nature in Chapter 5, before integrating both approaches in Chapter 6. As different models are appropriate for encoding this tie information, we consider a combination of signed, multiplex and weighted networks. Signed networks, in which edges carry either positive or negative weight, are used to differentiate between interactions that are cooperative or antagonistic in sentiment [99]. These networks pose a challenge to traditional methods for social networks, which are often predicated on the assumption that a connection is evidence of similarity. Both Chapter 3 and Chapter 4 introduce novel methods for social networks with antagonistic ties, with further applications presented in Chapter 6.

Ties that differ in *nature* can be represented using multiplex or weighted graphs, where the type or intensity of the interaction is encoded by an attribute on the edge [98]. In what follows, we use the term *heterogenous* to denote networks with diverse interaction types, rather than to networks that have heterogenous degree, as the term is occasionally employed in the literature [203]. Preserving this level of complexity is particularly crucial in modelling diffusive processes, from epidemic spread to information diffusion, as such processes can spread more effectively over distinct types of social ties [7]. Moreover, representing heterogeneous interactions is similarly essential for detecting latent structure in social networks, from the micro- to the macro-scale [98]. Chapter 5 presents a model of diffusion over a social network with heterogeneous interactions, examined through the lens of an empirical study. Chapter 6 integrates the contributions from previous chapters in a final empirical study of a network which is simultaneously signed, weighted and multiplex.

Through a series of methodological contributions, this thesis advances the network science literature on social networks with diverse ties. These contributions are bolstered by analysis on empirical social networks, which collectively demonstrate the value of maintaining a higher level of complexity in the representation of social interactions, an approach crucial for uncovering latent interaction patterns, and elevating our understanding of social processes.

1.2 Organization and Contributions

Chapter 2: Preliminaries

This chapter provides an overview of network science, establishing the mathematical foundation on which the subsequent research is based. We introduce basic graph properties, and the graph notation that we maintain throughout the thesis. The chapter also provides an introduction to signed networks, their properties and uses, graph embeddings, random walks on graphs, and the role of spreading processes in networked systems.

Chapter 3: A proximity-based embedding for signed networks

In this chapter, we present the first of two methods that focus on heterogeneity associated to the presence of negative ties, presenting a spectral embedding method for signed networks. The algorithm, SHEEP, is constructed as the minimum energy configuration of a Hamiltonian which depends on both the multi-scale structure of the network, and the edge signs. Unlike other signed network embedding methods, which primarily focus on partitioning the graph into clusters of nodes, SHEEP preserves proximal relationships, and is thus capable of recovering continuous node attributes across a range of synthetic and empirical networks. In Chapter 6, we apply the SHEEP algorithm to detect political polarization and partisanship on an empirical signed network constructed from user interactions on a social media platform [22].

Chapter 4: Modelling random walks over signed networks

This chapter presents our second method devoted to extracting information from signed networks, extending the concept of a random walk over a traditional network, to the signed context. Random walks can be used to detect community structure in graphs, measure node importance, or capture the similarity between nodes. Adapting the random walk to signed networks is non-trivial, especially for the purpose of community detection, where the goal is to partition the graph into communities with predominantly positive internal edges, placing negative edges between communities. Prior works on signed network random walks focus on the case where there are only two such communities, which is uncommon in empirical settings. In this chapter, we propose a signed network random walk capable of capturing the structure of a network with more than two such communities. We demonstrate that the walk generates a similarity matrix between nodes, which can be used downstream for unsupervised and semi-supervised clustering, often outperforming other walk-based methods [21].

Chapter 5: Social contagion of music taste over heterogeneous social ties

As with random walks, other diffusive processes can also be adapted to networks with heterogeneous edges. This chapter focuses on different types of heterogeneity, examining ties that differ in intensity, modality and reciprocity, and their influence on diffusion dynamics. The analysis is grounded in an

empirical study that characterizes how music taste spreads via person-to-person recommendations over the Spotify social network. To that end, we investigate the various social and contextual factors that influence the outcome of these recommendations, including the strength of the sender-receiver relationship, the user’s role in the Spotify social network, the social cohesion apparent in their network, and how similar the new artist is to the receiver’s taste. Our findings suggest that the success of the spread is affected by the strength of the social ties, and the social cohesion in the user’s surrounding network, features which are preserved by maintaining the heterogenous nature of the network. Finally, we leverage the findings to construct a predictive model of whether a shared music track will result in the receiver’s successful ongoing engagement with the shared artist, providing new insights into the multifaceted mechanisms underpinning the interplay between music discovery and social processes. This analysis was conducted in collaboration with the Spotify Research team [106].

Chapter 6: Detecting ideological extremes in social media networks

This final chapter weaves together many of the results from the preceding chapters, employing methodologies developed for both signed, weighted and multiplex networks, to quantify the political polarization on Menéame, an online social network that facilitates political exchange and information diffusion. We adopt a dual-method approach, analysing both the signed and unsigned network representations, using the SHEEP algorithm presented in Chapter 3 for the signed network. By comparing outcomes from both networks, we quantify how negative ties enhance the understanding of polarization across diverse conversation topics on the platform. While the unsigned Menéame network effectively delineates ideological communities, only by incorporating negative ties can we identify ideologically extreme users who exhibit extremity on specific topics, or engage in antagonistic behaviours. These findings reinforce the central premise of this thesis; the critical importance of preserving tie-level complexity in heterogeneous networks [39].

Chapter 2

Preliminaries

2.1 Graphs

This chapter introduces the mathematical background underpinning the research presented in the following four chapters. We begin with the fundamental unit of network science; a graph. A graph, or a network, comprises a collection of nodes, which are the entities in the graph, joined by edges that represent interactions between the entities.

A graph is defined as $\mathcal{G} = (V, E)$, where V is the node set, and E is the edge set. A node $i \in V$ and node $j \in V$ are connected by an edge if $(i, j) \in E$. Thus, the graph $\mathcal{G}$ has $|V| = n$ nodes and $|E| = m$ edges. A graph can be expressed as a mathematical object in a number of different ways, most frequently in this work by its *adjacency matrix* defined $A \in \mathbb{R}^{n \times n}$ where A_{ij} gives the value of the edge between nodes i and j . If no edge exists, $A_{ij} = 0$ [152].

If a graph is *undirected*, an edge from node i to node j is equivalent to an edge from node j to node i , and the resulting adjacency matrix will be symmetric. In the case of a *directed graph*, each edge has a direction, representing interactions that are not inherently reciprocal in nature. In this case, A_{ij} gives the value of the edge from node i to node j . In its simplest form, the graph $\mathcal{G}$ is *unweighted* and A_{ij} takes values in $\{0, 1\}$; an edge either exists between two nodes, or it does not. This thesis focuses on complex systems with *diverse interactions*, which are often represented by networks with weighted edges. In a *weighted graph*, the value of A_{ij} can be any real number value, which denotes the varied nature and intensity of the interaction [152]. Though the weights are typically take only positive values, this is not restrictive, as in the following section.

The degree of a node in a graph is defined by the sum of the weights of its incident edges. For a node i , the degree can be defined using the adjacency matrix as $k_i = \sum_j A_{ij}$. In what follows, we will often refer to the diagonal matrix of degrees defined as $D_{ij} = k_i \delta_{ij}$ [152].

Signed Networks

In *signed networks*, edges carry a positive or negative weight to represent the polarity of the interaction between the nodes, representing the dichotomy between interactions that are trustful or mistrustful, activatory or inhibitory, cooperative or antagonistic, depending on the context [8, 198, 189]. A fundamental example is that of a social network, where the sign of the edge indicates the sentiment of the interaction between individuals; a positive edge represents a friendly relation, while a negative edge represents animosity or antagonism.

Depending on the context, signed network representations can be constructed either via explicit or implicit signed interaction signals between entities in the network. In the case of online social interactions, which is the focus of the primary empirical studies in this thesis, signed networks are constructed by assigning a positive or negative weight to interactions between individuals. Some platforms naturally produce signed representations, as users can explicitly delineate the valence of their interaction [175, 133]. In other cases, signed signals can be inferred from surrounding textual context [169, 168].

Another widely used method is to construct signed social networks from bipartite interaction data, where connections only exist between nodes of two different types. This process, often referred to as *backboning*, extracts a signed network between entities in one group, where the weight of the edge indicates if the relationship is weaker or stronger than expected by chance [62, 18]. For other applications, signed networks are constructed from correlation or casualty matrices between time series, often used to study financial systems, or functional connectivity in the brain [62]. Similar correlation methods are applied to gene expression patterns to construct signed co-expression networks [138]. As with other network constructions, obtaining a signed networks representation necessarily involves modelling choices, including thresholding correlation values, aggregating explicit interactions, or inference through *backboning*, which can influence the observed structure.

As in the unsigned case, the signed graph $\mathcal{G} = (V, E)$ is the graph with node set V and edges E , where E^+ is the set of positive edges, and E^- is the set of negative edges, such that $E^+ \cup E^- = E$ and $E^+ \cap E^- = \emptyset$. As with unsigned networks, a signed network is characterized by its adjacency matrix $A_{s;ij}$, where $A_{s;ij} = 1$ indicates the presence of a positive edge $(i, j) \in E^+$, and $A_{s;ij} = -1$ indicates the presence of a negative edge $(i, j) \in E^-$. The signed adjacency matrix can be decomposed into two matrices, encoding the positive and negative edges separately, where the positive adjacency matrix A_+ is defined by $A_{+;ij} = 1$ if $A_{s;ij} = 1$, and 0 elsewhere. Similarly, the negative adjacency matrix A_- is defined by $A_{-;ij} = 1$ if $A_{s;ij} = -1$, and 0 elsewhere, such that $A_{s;ij} = A_{+;ij} - A_{-;ij}$.

The positive diagonal degree matrix is defined by $D_{+;ij} = k_{+;i} \delta_{ij}$, where $k_{+;i} = \sum_j A_{+;ij}$, the degree matrix of the positive sub-graph, and $D_{-;ij} = k_{-;i} \delta_{ij}$, where $k_{-;i} = \sum_j A_{-;ij}$, the degree matrix of the negative sub-graph, where in particular we note that $k_{-;i} \geq 0$. A node's total number of connections, positive or negative, is given by $k_i = \sum_j |A_{s;ij}| = k_{+;i} + k_{-;i}$ [152].

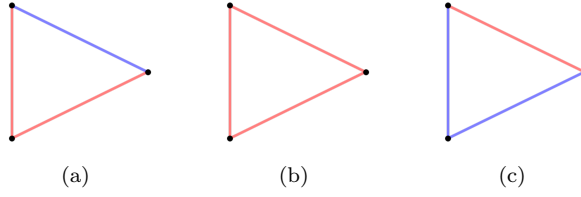


Figure 2.1: **Examples of Balanced and Unbalanced Cycles.** Strongly balanced cycles (a) have even numbers of negative edges. Weakly balanced cycles (b) have i negative edges, where $i \neq 1$. Unbalanced cycles (c) have a single negative edge. Throughout this thesis, red indicates an negative edge, blue indicates positive.

Structural Balance

In practice, social network analysis is heavily grounded in sociological theories used both to predict and explain the behaviour observed in the resulting system. The principle of *homophily*, the idea that individuals tend to form connections with those most similar to themselves, underpins many mathematical techniques used to measure individual similarity in networked systems, or to perform *community detection*, detecting groups of individuals more closely connected to each other than the rest of the network [152, 178].

However, negative links are naturally at odds with the principle of homophily, and techniques developed to extract information from unsigned networks often fail to capture dynamics present in signed networks, where connectivity is no longer an indication of similarity. Thus, signed networks require new techniques, inspired by alternative social theories. Signed social networks are most commonly understood through Heider’s theory of *structural balance*, a psychological theory which posits that interpersonal relationships tend to form structures that obey a notion of “balance”. The theory follows directly from the adage 1) *an enemy of my enemy is my friend*, and its natural extensions: 2) *a friend of a friend is also a friend* and 3) *an enemy of my friend is also my enemy* [102]. These rules can be translated directly into the language of network science, describing which triad structures are most favourable in a signed social network. Figure 2.1 (a) shows a balanced triad following the first adage, while Figure 2.1 (c) shows an unfavourable state or an unbalanced triad, which violates all three rules. The convention taken in this thesis is that red indicates a negative edge, while blue indicates a positive edge.

The idea of structural balance has been extended beyond the triad, to cycles of any length. Cartwright and Harary introduce the term *strong balance* to refer to a signed network where all cycles in the graph contain an even number of negative edges [42]. Fascinatingly, this local property has wider implications for the global structure of the graph. The authors show that strong balance is equivalent to *bi-polarization*, when a graph can be partitioned into two node clusters with positive edges within each cluster and negative edges between them [99, 42]. The communities in a signed social network are thus defined as groups of cohesive individuals with antagonistic relationships to the other groups.

A slightly weaker condition was introduced by Davis in [54], who defined a *weakly balanced* graph as one where no cycle has only a single negative edge. Figure 2.1 (b) shows a weakly balanced cycle. Davis proved that such graphs can be partitioned into k clusters of nodes with positive edges inside, and negative edges between them. When the graph is not strictly balanced, identifying the optimal partition of nodes into k clusters, where k is unknown, has been shown to be NP hard. Various methods have been proposed to solve this problem numerically [18, 65]. A popular approach takes inspiration from spin-glass literature, using simulated annealing to locate the optimal cluster assignment [200, 101]. Analogously to the term used for negative plaquettes in the Ising spin-glass model, unbalanced k -cycles on signed graphs are also called *frustration* [40].

Generative Models

Generative models provide mathematical frameworks for producing families of networks which share certain statistical properties, which often mirror those found in empirical networks. A prominent example is the presence of *community structure* in social networks, where nodes can be partitioned into groups more densely connected to each other, than to nodes in other communities. This property finds its dual in the canonical stochastic block model (SBM). Given k communities of pre-determined sizes $\{C_1, C_2, \dots, C_k\}$ and a $k \times k$ probability matrix P , the SBM generates a synthetic network where P_{ij} gives the probability that an edge exists between a node in cluster i and another node in cluster j [2]. To induce community structure, the intra-community edge probabilities must be higher than the inter-community probabilities. Refer to Figure 2.2 for an informative illustration of a three-community SBM.

In Chapter 3 and 4 of this thesis, we employ the signed counterpart, the signed stochastic block model (SSBM). Like the standard SBM, the SSBM is similarly a random graph model with a cluster structure that typically either exhibits strong or weak balance. As in the unsigned case, the SSBM has k clusters of nodes, and edge probabilities between each pair of nodes are independent, defined by the probability matrix P . In keeping with the paradigm of structural balance, all edges connecting nodes in the same cluster are by default assigned positive weight, while edges connecting nodes in different clusters are weighted negative. Unlike in the unsigned case, the edge existence probabilities for inter- and intra-community edges need not differ, as the structure of the network is determined by the signs, rather than the density of links [53]. This model can be perturbed by introducing a sign-flip parameter $\nu \in (0, 1)$, controlling the probability that the sign of an edge is flipped from the correct sign (positive inside communities and negative between communities).

The signed stochastic block model is the standard generative model for producing signed networks which exhibit structural balance, and are widely used in the literature to benchmark methods for detecting signed clusters [53]. In this thesis, we use SSBMs for two purposes: in Chapter 3, to

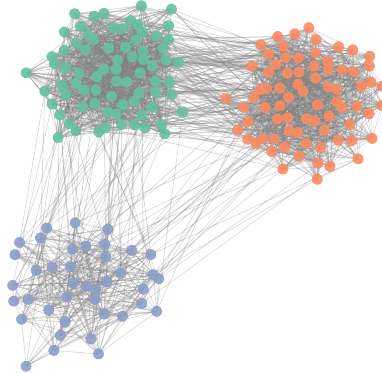


Figure 2.2: **Three-community stochastic block model (SBM)**

evaluate the properties of the signed network embedding method we propose; and in Chapter 4, to study the behaviour of our signed random walk over networks with community structure.

Spectral Properties

The spectral properties of the graph's matrix representation provide insights into the structural and dynamical characteristics of the network. In addition to the aforementioned adjacency matrix, the graph Laplacian is another common matrix representation of the network, which renders itself particularly useful in community detection problems.

For a matrix $M \in \mathbb{R}^{n \times n}$, we denote its eigenvalues $\lambda_i(M)$ and associated eigenvectors $\boldsymbol{\nu}_i(M)$. In this thesis, we focus primarily on undirected networks, for which the associated matrices are symmetric. A symmetric matrix M has a set of orthonormal eigenvectors and associated real eigenvalues which can be ordered as $\lambda_1(M) \geq \lambda_2(M) \dots \geq \lambda_n(M)$.

The (unsigned, undirected) graph Laplacian is defined via the adjacency matrix A and the degree matrix D as:

$$L = D - A \quad (2.1)$$

such L is symmetric and has a 0 eigenvalue corresponding to the eigenvector $\mathbf{1}$, the all ones vector. The Laplacian is positive-semidefinite such that $0 = \lambda_n \leq \lambda_{n-1} \dots \leq \lambda_1$ [28]. For a *connected* network, one such that every node is reachable from every other node in the graph, the second smallest eigenvalue of the Laplacian is strictly greater than zero, $\lambda_{n-1} > 0$, and in general the number of connected components of the graph is given by the multiplicity of the zero eigenvalue in the spectrum of the Laplacian [122].

Defining the Laplacian in the case of a signed graph is less straightforward, and two main variations have been proposed in the literature [185]. While the Laplacian of the positive part of the graph follows the same form as the unsigned Laplacian as $L_+ = D_+ - A_+$, there are two possibilities

for the negative graph Laplacian, corresponding to two possibilities for the signed Laplacian. The opposing Laplacian:

$$L_o = D_+ + D_- - (A_+ - A_-), \quad (2.2)$$

and the repelling Laplacian:

$$L_r = D_+ - D_- - (A_+ - A_-). \quad (2.3)$$

Both the repelling and opposing Laplacians are symmetric and real, thus their eigenvectors can be chosen to be orthonormal. Note that the repelling Laplacian follows precisely the same form as the standard graph Laplacian since $\sum_j A_{s;j} = D_{i;+} - D_{i;-}$, but since the net degree of a node can be negative, the matrix is indefinite. The opposing Laplacian is constructed using the diagonal absolute-value degree matrix defined by $\bar{D}_{ii} = \sum_j |A_{ij}|$, consequently, it is positive semi-definite, and its spectral properties are closely associated with the notion of strong balance [120].

Note that the literature does not have a consistent naming system for the different Laplacians we refer to here. This thesis adopts the term “opposing Laplacian” after [185], but this matrix is sometimes called the signed Laplacian in [120]. The repelling Laplacian appears under various names in other works, including the physics Laplacian in [79], the net Laplacian in [160], or simply the unsigned Laplacian as it shares an identical construction to the Laplacian on unsigned graphs.

Network Embeddings

Networks are high-dimensional objects, and their structure is often investigated through dimensionality-reduction techniques called network embedding. Network embedding encompasses a range of methods for obtaining a lower-dimension vector representation of the nodes, taking into account multi-scale graph topology, while still preserving certain structures. Embeddings have a range of downstream applications from clustering and community detection, to semi-supervised node classification, and are used in recommender systems to make recommendations based on proximity in a latent (embedding) space [209, 78].

The full embedding of a graph associates each node i to a position in k dimensional space, which is encoded by a $n \times k$ matrix X with entries $X_{i\alpha}$ or, equivalently, the set of n vectors $\mathbf{x}_i$. For further clarity in this work, we use Latin indices for nodes, and Greek indices for dimensions.

Classically, network embedding methods for unsigned graphs fall into two categories; spectral methods and physical methods, which have different downstream usages. Spectral methods produce k -dimensional vector representations of the nodes by taking the first k eigenvalues of a matrix associated with the graph. For example, the smallest non-zero eigenvalues of the Laplacian matrix, and their associated eigenvectors, are shown to produce an embedding of an unsigned graph which preserves local distances between nodes [28].

Another class of spectral techniques relies on singular value decomposition (SVD) rather than eigendecomposition, a series of techniques that are reminiscent of Principal Component Analysis. Correspondence Analysis (CA) is one such method that is broadly applicable to both networks, and to categorical data that can be represented as contingency tables [103, 91, 29, 90]. In particular, given an interaction matrix $M \in \mathbb{R}^{n \times m}$, the correspondence matrix is defined by $P = \frac{1}{\tau} M$, where $\tau = \sum_i \sum_j M_{ij}$ is the sum of all the entries of the matrix. The matrix of standardized residuals, S is computed:

$$S = D_r^{-\frac{1}{2}}(P - \mathbf{r}\mathbf{c}^T)D_c^{-\frac{1}{2}}, \quad (2.4)$$

where $\mathbf{r} = P \mathbf{1}$ and $\mathbf{c} = P^T \mathbf{1}$, which define the row and column masses, and $D_x = \text{diag}(\mathbf{x})$ denotes the vectors transformed into diagonal matrices. Finally, the SVD of the matrix is computed via $S = U\Lambda V^T$ such that $U^T U = V^T V = I$. The diagonal matrix Λ contains the singular values of S and is used to determine the embeddings of the entities represented in the interaction matrix. These embeddings, called *principal coordinates* are defined for the rows of the matrix M as $D_r^{-\frac{1}{2}} U \Lambda$, and for the columns as $D_c^{-\frac{1}{2}} V \Lambda$. This method is used extensively in Chapter 6.

Beyond purely matrix-based techniques, another prominent family of embeddings are energy-based physical models typically used to produce visually-informative two-dimensional depictions of an unsigned graph. An energy function is chosen by assigning attractive and repulsive forces to the edges, and an optimization algorithm employed to locate a minimum energy configuration. The Spring Embedder method was one of the earliest force directed models, which was further refined by Fruchterman and Reingold [68, 82]. Similarly, Kamada and Kawai introduced an algorithm which treats edges as springs, with rest length equal to the graph theoretical distance [111].

The advantage of spectral and physical methods lie in their interpretability, which becomes increasingly difficult to maintain as embedding algorithms increase in complexity. For example, random-walk based embedding methods use random walks to sample a node’s “context” or neighbourhood of connected nodes, capturing both local and global structure [47]. The sampled paths are used to generate node embeddings via the skip-gram or bag-of-words model [95, 210]. More recently, neural-network based methods have been increasing in popularity, particularly the Node2Vec algorithm which similarly uses random walks, but optimizes the representation using stochastic gradient descent [93, 47]. While such methods are perhaps less interpretable, they can out-perform spectral methods at downstream predictive tasks.

For signed networks, embedding methods typically seek to place nodes close together in latent space if they are connected by positive edges and paths, while simultaneously separating nodes connected by negative edges. Chapter 3 provides a comprehensive overview of embedding methods for signed graphs.

Random Walks on Networks

A vast area of network science is devoted to examining processes taking place *on top* of a network. Random walks are used to model diffusive processes over networks, finding applications in epidemic spread, or dissemination of information or behaviours over social networks [155]. They are extremely powerful for probing the structure of large-scale graphs, feeding into measures of node centrality, informing community detection algorithms, or describing other structural patterns [46].

A random walk begins at an initial node, and at each step randomly selects one of the incident edges and proceeds to an adjacent node, as in Figure 2.3. The walk itself is the path produced by a repeated series of these random choices, up to a defined stopping point. In its simplest form, then, the random walk is modelled as a Markov chain, such that the state of the walker at time t only depends on the state of the walker at time $t - 1$ [122]. On an unsigned graph, a discrete-time random walk evolves according to

$$n_j(t + 1) = \sum_i n_i(t) \frac{1}{k_i} A_{ij} \quad (2.5)$$

where $n_i(t)$ is the density of walkers on node i at time t and A_{ij} is the adjacency matrix of the graph and k_i is the degree of the node, such that $T = D^{-1}A$ is the transition matrix of the associated Markov chain. This equation describes the situation where an ensemble of walkers are processing on the graph simultaneously at discrete times.

As the transition matrix T is row-normalized with eigenvalue 1, the process described here has a stationary state $\mathbf{n}$ defined by the eigenvalue equation [140]:

$$\mathbf{n}(I - T) = 0. \quad (2.6)$$

The spectral properties of the transition matrix, and how they relate to the associated graph structure, are well known [139]. In particular, the eigenvalues are in the interval $[-1, 1]$, and the multiplicity of the dominant eigenvalue, $\lambda_1 = 1$ is equal to the number of connected components in the graph. Moreover, the transition matrix is related to the graph Laplacian via the relation $T = I - D^{-1}L$, such that the dynamics of the random walk are closely related to the spectral properties of the graph's Laplacian [141]. Chapter 4 provides a more comprehensive overview of the spectral properties of this matrix, and introduces a random walk process over signed networks.

Spreading Processes on Networks

Dynamical processes on networks, which encapsulate both random walk and spreading process, are used to model different phenomena. Spreading processes describe how states propagate between nodes, though unlike the random walk process, the number of diffusing agents is not conserved. Some of the most well-studied examples come from epidemiology, where infectious disease dynamics

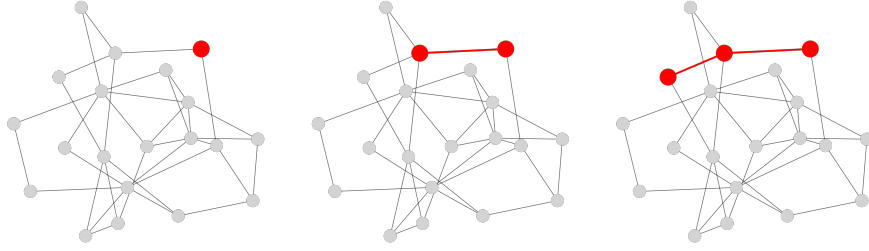


Figure 2.3: **First three steps in a random walk process.** Selected nodes are coloured in red.

are modelled using various generalizations of the SI model [152]. More recently, similar models have been applied to understand the spread of information, behaviours or trends between individuals in a social network [94]. A substantial body of literature explores the relationship between the network structure, and the speed of the spread [155]. Unlike the spread of infectious disease, voluntary behaviours tend to require social reinforcement to spread effectively, as in Granovetter’s threshold model, where the probability of a successful transmission event increases as a function of the number of previously activated connections [89]. The successful adoption of behaviours requiring social reinforcement, called complex contagion, is impacted by the network topology, for example, spreading more easily between densely connected nodes [44]. Chapter 5 draws on the complex contagion literature, presenting an empirical study of how music taste spreads over a social network with heterogenous edge types.

Chapter 3

A proximity-based embedding for signed networks

This chapter is based on the article *Babul, Shazia'Ayn, and Renaud Lambiotte. "SHEEP, a Signed Hamiltonian Eigenvector Embedding for Proximity." Communications Physics 7.1 (2024). 8 [22]*.

3.1 Introduction

Networks provide a powerful representation of complex systems, comprising of a collection of nodes joined by edges that represent interactions [152]. Typical examples stem from an array of disciplines, including sociology (social contact networks) urban infrastructures (public transportation networks), or biological interactions (gene interaction networks). As described in Chapter 2, edges can represent diverse interaction types, which can be distinguished by the edge weights, which identify the nature and intensity of the relationship. In this chapter, we consider signed networks, where positive or negative edge weights indicate the valence of the interaction between the nodes.

Despite their classical origin in the sociological literature [102], signed networks represent a relatively new field of exploration. However, research into signed networks has been increasing in popularity due to a number of examples from online social networks such as Epinions [175] and Slashdot [133] where users can mark each other as friend or foe. More recently, signed social networks have been constructed by inferring the signed signals using sentiment analysis, using data from social media platforms like Reddit [169] and Twitter [168]. We give an overview of the related literature on signed networks in the following sections, and direct the reader to [191] for a more detailed explanation of their properties and principles.

As described in Chapter 2, the structure of signed (and unsigned) networks are often investigated through dimensionality-reduction network embedding techniques. For signed networks, a successful embedding method will place nodes close together in latent space if they are connected by positive edges and paths, while separating nodes connected by negative edges [120, 53, 117, 79, 113].

In this chapter, we present a physically-inspired signed network embedding method called SHEEP (Signed Hamiltonian Eigenvector Embedding for Proximity), which incorporates local, intermediate and global information into a continuous proximity measure between nodes. Modelling the nodes as a system of particles interacting with attractive forces (positive edges) and repulsive forces (negative edges), we construct a Hamiltonian which is dependent on the distance between nodes, such that the relative embedding distance results in a similarity metric. Unlike most signed network embedding methods which are grounded in the principle of structural balance and focus on partitioning the graph into antagonistic clusters for binary classification tasks, our measure is intended to retain proximal information between nodes, and may be preferable in the case where a graph does not exhibit perfect structural balance.

We propose two applications of the signed network embedding method; (1) a way to recover continuous node attributes based on their proximity in the embedding, and (2) using the node’s distance to the origin as a measure of extremism or conflictuality, evaluating the method on both synthetic generated signed networks and empirical networks. In addition, the Hamiltonian admits a global minimum energy configuration, which can be re-configured as an eigenvector problem, and therefore is computationally efficient to compute. We also present an energy-based approach to finding the optimal embedding dimension for proximity. At the global level, the embedding generated by the minimum energy configuration is intrinsically related to the notion of structural balance, and the ground state energy can be used as a statistical test for global-level strong balance, or bi-polarization.

3.2 Related Literature

We provide here a brief review of the related methodologies developed to extract information from signed networks, and how they differ from the work presented in this chapter.

As first discussed in Chapter 2, signed networks are most commonly understood through the psychological theory of structural balance, developed to explain why “balanced” structures are more favourable in interpersonal relationships [102]. When a graph is structurally balanced, the nodes can be partitioned into k clusters with positive edges inside and negative edges between them. Figure 3.1 depicts a strongly balanced (2 faction) graph and a weakly balanced (3 faction) graph.

While structural balance provides a valuable framework for understanding signed networks, it has been argued in the literature that empirical networks are often not perfectly balanced, and the tendency for signed social networks to become balanced may not even hold over time [64]. In particular, signed networks may not have a natural cluster structure, and as proposed in [65], their formation may be influenced by a combination of other processes, resulting in different patterns than structure balance alone. It is precisely when ground truth clusters do not exist that it may be

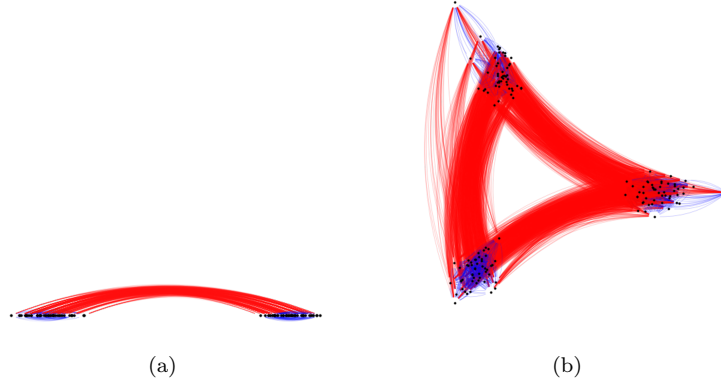


Figure 3.1: **Signed Network Global Structure.** Strongly balanced signed networks (a) have two opposing factions with positive edges inside and negative edges between. Weakly balanced networks (b) can have 3 or more factions.

more useful to consider methods that capture more nuanced relationships between nodes, such as proximity measures, which avoid making assumptions about faction numbers and memberships.

SHEEP is one such method; a spectral embedding method for signed graphs, which emphasizes representing proximal relationships between nodes over clustering assignments. SHEEP is situated within a broader class of spectral techniques for signed networks. While the majority of such methods on signed networks are primarily designed to cluster the graph into cohesive communities separated by antagonistic relationships, which is not our objective here, it is nevertheless a rich field and we give an overview of the relevant literature as follows.

Spectral clustering on signed graphs has been approached in various ways; in [9], the authors seek to optimize modularity, in a similar spirit to the modularity maximization methods for community detection on unsigned graphs [156]. In [120], Kunegis et al. propose a spectral method, embedding the signed graphs in two dimensions, using the first two eigenvectors of the opposing Laplacian. Extending the opposing Laplacian method to higher dimensions proved to have significant limitations in [49], where the authors instead minimize an objective called the Balanced Normalized Cut. Xiao et al. [206] introduce a method that uses a normalized version of the opposing Laplacian to identify locally optimal polarized communities, given a set of input nodes. In [53], Curcuringu et al. construct a regularized spectral algorithm that performs particularly well at recovering antagonistic clusters, even on sparse graphs.

Most similar to our work here is that of [117] in which Knyazev takes a mechanical perspective, arguing that signed graphs can be clustered using the repelling Laplacian, modelling clusters as eigenmodes of a mass-spring system with negative springs. The graph is clustered using the k smallest eigenvectors, where k is selected based on the largest observed eigen-gap. While SHEEP also employs the repelling Laplacian, our method adopts an energy-based approach, and we seek to

capture proximal relations rather than clusters. We further demonstrate that our method naturally outputs a statistical test for polarization.

In [36, 37], the authors present a proximity-based two-dimensional representation of a signed network onto a "conflict space" using the smallest eigenvalues of the adjacency matrix. Unlike this method which is designed to be two-dimensional for visual comprehension, SHEEP generalizes to higher dimensions. Other approaches similarly depart from the strict clustering paradigm necessitated by structural balance, for example by allowing for "neutral" actors to exist outside the clusters [33, 201].

In [79], Fox et al. apply QR factorization on the eigenvalues of Gremban's expansion, a matrix related to a random walk process on a signed network [193]. As pointed out in [53], most spectral clustering methods fix *a priori* the number of clusters $k + 1$, embed the nodes using the k smallest eigenvectors, and finally apply a k-means analysis to identify clusters. We direct the reader to [80] for a more detailed review of related spectral clustering methods.

SHEEP is a spectral embedding method, inspired by a physical Hamiltonian, thus it also falls into the energy-based family of graph embedding methods discussed in Chapter 2. Other physical approaches for signed networks have been proposed; in [113], Kermarrec et al. introduce an energy-based approach for visualizing signed graphs in two dimensions to detect clusters.

Unless the chosen energy function is strictly convex, the optimization algorithms employed for energy-based methods locate a local minima, which is acceptable for the purposes of graph drawing. SHEEP instead relies on a Hamiltonian which admits a global minimum. We draw inspiration from SpringRank, a physically inspired model that generates hierarchical rankings of nodes in directed, unsigned networks along a one dimensional line, by modelling directed edges as orientated springs [55]. This conceptualization inspires the spring potential based SpringRank Hamiltonian, defined:

$$H(\mathbf{s}) = \frac{1}{2} \sum_{i,j} A_{ij} (s_i - s_j - 1)^2, \quad (3.1)$$

where each node i is embedded at a real-valued position given by s_i . Each spring has a rest length set to 1, such that the energy is minimized when $s_i - s_j = 1$, preventing nodes from collapsing to the same position. The SpringRank Hamiltonian has a global minimum which allows for statistical comparisons of the ground state energy, to determine significance of the hierarchical structure. In this chapter, we generalize the SpringRank method to undirected signed networks, embedding the nodes in a k -dimensional space by defining a Hamiltonian that is the sum of repulsive and attractive forces corresponding to negative and positive edges respectively, and then defining the embedding by locating the global minimum energy configuration.

3.3 The SHEEP Hamiltonian

The generalization of the spring-force Hamiltonian to a signed graph setting is constructed as follows. Suppose each node i has an associated position vector $\mathbf{x}_i$ in metric space $\mathbb{R}^k$. Positive edges are modelled as spring attractive forces, and negative edges are associated with an anti-spring repulsive force, which is similarly quadratic in distance, resulting in:

$$H = \sum_{i,j} A_{+,ij} |\mathbf{x}_i - \mathbf{x}_j|^2 - \sum_{i,j} A_{-,ij} |\mathbf{x}_i - \mathbf{x}_j|^2. \quad (3.2)$$

Recall that $A_{-,ij} \geq 0$. The first term of the Hamiltonian is minimized when positively connected nodes are separated by a minimal distance, while the second term requires negatively connected nodes to be separated by the largest possible distance. We seek to find the set of position vectors $\{\mathbf{x}_i\}$ that describe the minimum energy configuration. When a negative edge exists between two nodes, their interaction term reaches its minimum when the two nodes are pushed apart to infinity. Consequently, we introduce a constraint on Eq. 3.2 to control for the explosion of distance.

3.3.1 Relation to Spectral Methods

Before introducing the constraint, we first illustrate the relationship between the proposed energy-based embedding and spectral approaches. As a graph exhibiting strong balance can be separated into two antagonistic factions (bi-polarization), we constrain ourselves to a one dimensional embedding for the remainder of this section. The vector $\boldsymbol{\pi}$ denotes the n -dimensional vector describing the positions π_i of the n nodes along a line.

Following from Chapter 2, a signed graph has two possible Laplacians; the opposing Laplacian defined $L_o = D_+ + D_- - A_s$ and the repelling Laplacian defined $L_r = D_+ - D_- - A_s$. These two matrices induce quadratic forms on the vector $\boldsymbol{\pi} \in \mathbb{R}^n$. Note that π_i is a scalar value associated with the node i . The induced quadratic form of the opposing Laplacian is:

$$\boldsymbol{\pi}^T L_o \boldsymbol{\pi} = \sum_{i,j} A_{+,ij} |\pi_i - \pi_j|^2 + \sum_{i,j} A_{-,ij} |\pi_i + \pi_j|^2, \quad (3.3)$$

and of the repelling Laplacian:

$$\boldsymbol{\pi}^T L_r \boldsymbol{\pi} = \sum_{i,j} A_{+,ij} |\pi_i - \pi_j|^2 - \sum_{i,j} A_{-,ij} |\pi_i - \pi_j|^2. \quad (3.4)$$

Notice that the quadratic form resulting from the repelling Laplacian is precisely the spring-inspired Hamiltonian proposed in Eq. 3.2, for one-dimensional position vectors.

Unlike the function arising from the inner product of the opposing Laplacian (Eq. 3.3), the inner product of the repelling Laplacian is a function of distance for pairs of nodes connected by both positive or negative edges. Consequently, in the embedding generated by the opposing Laplacian, the presence of many or few negative links between positively connected factions does not affect

the final positions, unlike the physically inspired method that we propose here. In particular, the opposing Laplacian identifies whether the graph is strongly balanced or not, but does not give an indication of the intensity of the negative interactions between the nodes, which is precisely the advantage of our proximity-based embedding. As we will show, the embedding defined by repelling Laplacian does not only depend on the number of positive and negative connections, but also on their specific arrangement between the nodes and on the constraints induced by the dimension k of the embedding.

Both signed Laplacians are symmetric and real, so their eigenvectors can be chosen to be orthonormal. Unlike the opposing Laplacian, which has been previously used for spectral embedding as in [120] because it is positive semi-definite, the repelling Laplacian is indefinite. From a physical perspective, its inner product describes a Hamiltonian that permits explosions of distances due to the quadratically increasing energy of repulsive forces.

3.3.2 Spherical Constraint

Let us first define the one-dimensional form where each node i is associated with a position π_i , such that $\boldsymbol{\pi} \in \mathbb{R}^n$. Then, our objective function to minimize is:

$$H(\boldsymbol{\pi}) = \sum_{i,j} A_{+,ij} |\pi_i - \pi_j|^2 - \sum_{i,j} A_{-,ij} |\pi_i - \pi_j|^2, \quad (3.5)$$

As the minimization of Eq. 3.5 is computationally difficult, we follow a method proposed in [112] for ordering nodes in graphs, and impose the spherical constraint condition that $\sum \pi_i^2 = R$. This constraint, which is equivalent to fixing the L^2 norm, relaxes the problem to admit a tractable global solution. The choice of this constraint follows the existing literature [112], however, we note that alternative constraints are possible and would, in general, lead to different solutions. This constraint can be written using a Lagrange multiplier, as:

$$\mathcal{L}(\boldsymbol{\pi}, \lambda) = H(\boldsymbol{\pi}) - \beta(\sum \pi_i^2 - R), \quad (3.6)$$

where β is the Lagrange multiplier. Taking the gradient and setting it to zero gives the following eigenvalue equation for the extrema:

$$L_r \boldsymbol{\pi} = \beta \boldsymbol{\pi}. \quad (3.7)$$

The extrema conditions generated by the constraint requires that $\boldsymbol{\pi}$ is an eigenvector of the repelling Laplacian, $\boldsymbol{\nu}$, which gives:

$$H(\boldsymbol{\nu}) = \boldsymbol{\nu}^T L_r \boldsymbol{\nu} = \lambda R, \quad (3.8)$$

where λ is the associated eigenvalue, and the eigenvector has been normalized such that $\boldsymbol{\nu}^T \boldsymbol{\nu} = R$.

From Eq. 3.8, we see that the minimum is obtained when $\boldsymbol{\nu}$ is the eigenvector $\boldsymbol{\nu}_n$ associated to the minimum eigenvalue λ_n . Finally, we have that the minimum value of the Hamiltonian subject to the spherical constraint is the minimum eigenvalue of the repelling Laplacian.

This minimum is a global minimum for the system defined in Eq. 3.6 with the spherical constraint, and the associated one-dimensional embedding is given by the eigenvector $\boldsymbol{\nu}_n$. The resulting embedding is invariant under reflection about the origin. Furthermore, since we are most interested in the relative distances between nodes in this proximity-based embedding, the resulting normalized distance matrix generated from the embedding is also invariant under translations and re-scaling of the vector $\boldsymbol{\nu}_n$. In what follows we take $R = 1$. Note that unlike the other force based approaches which often employ non-convex optimization to locate local minima, finding the minimal value of the Hamiltonian with the spherical constraint reduces to a computationally efficient eigenvalue problem.

In the method proposed here, we only consider signed graphs which have at least one negative eigenvalue. For positive graphs, the repelling Laplacian is equivalent to the standard graph Laplacian. Since this matrix always has zero row sum, it has a 0 eigenvalue associated with the eigenvector $\mathbf{1}$, which is the smallest eigenvalue for the positive graph. From the physical interpretation of the system, this means the attractive forces have collapsed all the nodes onto a single point. It is also possible for signed networks to have no negative eigenvalues, for example if the positive sub-graph is fully connected [38]. We consider only graphs that have at least one negative eigenvalue, which occurs when the negative edges are sufficiently dense or well positioned in the graph such that the proximity embedding information is non-trivial.

3.3.3 Ground State Energy and Strong Balance

In this section, we propose a test statistic for bi-polarization, by exploiting the physicality of our system; the ground state energy, denoted as E^0 . Analogously to the Ising model, where the lowest ground state energy is achieved when there is no geometric frustration introduced by negative cycles, we show that the ground state energy of our Hamiltonian is minimized when the graph exhibits strong balance. If a test statistic is significant, it should be highly improbable on a null model. As the ground state energy depends on many aspects of the network structure, we focus on the following null model: fixing the graph topology, we randomly reshuffle the signs of the edges, while preserving the density of positive and negative edges, as in [132, 189]. If the real network has a lower ground state energy than the networks produced by the null model, we can conclude that the network has significant underlying bi-polarization structure. Since the ground state energy of our Hamiltonian is the minimum eigenvalue of the repelling Laplacian, we present here some results on the spectrum of the repelling Laplacian for complete-graphs, and then generalize the results numerically to non-complete graphs.

Repelling Laplacian: Spectral Results

An n -complete graph is a graph with n nodes, such that all possible edges exist between the nodes. We consider the set of complete signed graphs with n nodes, subject to the condition that $E^+, E^- \neq \emptyset$, denoted by $\{G^n\}$. We present three results: (1) The minimal eigenvalue λ_n is bounded from below by $-n$. (2) The lower bound $\lambda_n = -n$ is achieved when the graph has a perfect bi-partition (strong balance). And (3) Introducing frustration by switching edge signs will strictly increase the value of the ground state energy. These results motivate the use of the ground state energy as a bi-polarization measure, or measure of strong balance.

Theorem 1. *For an n -complete signed graph subject to the condition that $E^+, E^- \neq \emptyset$, the smallest eigenvalue of the repelling Laplacian $\lambda_n \geq -n$.*

The proof is deferred to Appendix, but the main steps involved are as follows. Beginning with the n -complete graph of all negative edges, one can easily show that the spectrum is $-n$ with multiplicity $n - 1$ and 0 with multiplicity 1 . Any other signed n -complete graph can be reached by successively changing an edge sign to positive, associated with the addition of a L^{shift} to the repelling Laplacian, which is positive semi-definite. Applying Weyl's inequality to the sum of the two Laplacians gives the required result.

Next, we show that the lower bound is reached when the graph is strongly balanced and admits a bi-partition.

Theorem 2. *Consider a strongly balanced graph $\tilde{G} \in \{G^n\}$, such that the nodes of graph $\tilde{G}$ can be partitioned into two sets V_1 and V_2 , where $V_1, V_2 \neq \emptyset$. Nodes inside each set are connected with positive edges, while the edges connecting the sets are negative. The smallest eigenvalue of the repelling Laplacian is $\lambda_n = -n$.*

Again, we defer the proof to Appendix. Broadly, we construct the eigenvector ν associated to the minimum eigenvalue λ_n by placing nodes in the same set at the same point such that for node $k \in V_i$ we set $\nu_k = x_i$ where $i \in (1, 2)$. Using the orthonormality of the eigenvectors, we can show that associated eigenvalue is $-n$,

$$E_{bal}^0 = \lambda_n = -n. \quad (3.9)$$

By Theorem 1, this is the minimum eigenvalue. By Theorems 1 and 2, the ground state energy reaches the lower bound when the graph is non-frustrated, and has a perfect bi-partition in accordance with strong balance. When the graph is frustrated we want to show that the ground state

energy, E_{frus}^0 is larger than the energy associated with the balanced graph E_{bal}^0 given in Eq. 3.9 as $-n$. We compare between graphs that have the same number of nodes, n , as well as the same density of positive and negative edges.

Theorem 3. *Consider a strongly balanced graph $\tilde{G} \in \{G^n\}$, such that the nodes of graph $\tilde{G}$ can be partitioned into two sets V_1 and V_2 , where $V_1, V_2 \neq \emptyset$. Introduce frustration into the graph by exchanging the signs of one positive and one negative edge, thereby preserving the overall density of positive and negative links. When the graph is sufficiently large such that $n > 4$, and $|V_1|, |V_2| > 2$, the ground state energy is strictly increased $E_{frus}^0 > E_{bal}^0$.*

The proof involves writing the repelling Laplacian of the frustrated graph as the sum of the repelling Laplacian of the balanced graph $L_r^{\tilde{G}}$, and a perturbation, L^{switch} , associated to the swapping of a positive and negative edge sign. The motivation for this swap comes from the definition of the null model defined above, where edge signs are reshuffled while preserving the number of positive and negative links. The new graph is frustrated because the two terms $L_r^{\tilde{G}}$ and L^{switch} cannot be simultaneously minimized. Finding a bound on the second smallest eigenvalue of the repelling Laplacian of the balanced graph $L_r^{\tilde{G}}$, gives an approximation for the energy gap between the ground and first energy levels. When the graph is sufficiently large such that $n > 4$, and $|V_1|, |V_2| > 2$, we can use the energy gap to bound the ground state energy from below, obtaining the inequality we require. In practice, $n > 4$ is a weak constraint, as most graphs of interest contain more than four nodes, so

$$E_{frus}^0 > -n = E_{bal}^0. \quad (3.10)$$

Bi-Polarization Measure

Following the intuition provided by the spectrum of the repelling Laplacian on complete signed graphs, we numerically generalize to the case of non-complete graphs, showing that the ground state energy is a statistically significant measure of bi-polarization, as compared to the null model. The test correctly identifies the polarization structures on various synthetic signed graphs with strong balance. We propose a bi-polarization score using the z-score of the ground state energy of the graph, compared to the null model energy distribution. We define:

$$Z(G) = \frac{E_0^G - \langle E_0 \rangle}{\sigma}, \quad (3.11)$$

where $\langle E_0 \rangle$ is the mean of the null model distribution, and σ is the standard deviation. A large negative value of $Z(G)$ indicates that the graph is significantly polarized.

There exist a number of proposed bi-polarization measures for signed networks, which typically focus on the local or global properties of the graph. For example, [116] proposes a balance measure

that counts the number of unbalanced cycles in the graph, inversely weighted by length. A more global measure is the frustration index, which counts the number of frustrated edges associated with the best bi-partition [19]. An intermediate scale measure, POLE, takes the correlation between a node’s signed and unsigned random walk, based on the assumption that link sign should be related to unsigned community structure [107]. Our measure incorporates both local and global structure, and allows for comparison between networks with fixed size and edge density. We refer the reader to the Appendix A.2 for a numerical test of the bi-polarization measure on synthetic signed networks with polarization structure.

3.4 Weak Balance and Extension to Higher Dimensions

A one-dimensional embedding provides sufficient insight to understand the graph’s alignment with strong balance, as the bi-partition of the nodes can be projected along the one-dimensional line. Weakly balanced graphs, however, are k -clusterable, and require partitioning the nodes into k factions. As a result, a one-dimensional projection may not adequately represent the relationships between the nodes. For noisy and un-balanced graphs, the extension to higher dimensions becomes essential; additional dimensions provide greater degrees of freedom to ensure that the distance between nodes in the embedding represents a proximity measure that takes into account the global graph structure.

When considering graph embeddings in higher dimensions, identifying the optimal dimension of the embedding is a challenging task. In the limiting case of a complete negative graph, intuition suggests that an optimal embedding should place each of the n equidistant from each of the others, a configuration which requires a $n - 1$ -dimensional embedding. We refer the reader to the Appendix A.3 for a detailed analysis of this case. To investigate the general case of non-complete graphs, we define, for each node i , a k -dimensional vector $\mathbf{x}_i$, which is the vector built using the i -th entry of each of the first k eigenvectors, i.e. $\mathbf{x}_i = (\nu_{ni}, \nu_{n-1i}, \dots, \nu_{(n-k+1)i})$. Let the distance matrix $D(k)$ correspond to the Euclidean distances associated to the positions of the nodes in the k dimensional latent space, such that:

$$D(k)_{ij} = |\mathbf{x}_i - \mathbf{x}_j|^2 = \sum_{\alpha=n}^{n-k} (\nu_{\alpha i} - \nu_{\alpha j})^2, \quad (3.12)$$

where $\nu_{\alpha i}$ is i -th entry of the α -th eigenvector of the repelling Laplacian. This implies that the distance matrix $D(k)$ is linear in dimension in the sense that if $D(\boldsymbol{\nu}_\alpha)_{ij} = (\nu_{\alpha i} - \nu_{\alpha j})^2$ are the entries of the distance matrix associated with the α -th eigenvector, then:

$$D(k) = \sum_{\alpha=n}^{n-k} D(\boldsymbol{\nu}_\alpha). \quad (3.13)$$

Since the Hamiltonian in Eq. 3.2 can be written in terms of the distance matrix $D(k)$:

$$H(D(k)) = \sum_{i,j} A_{+,ij} D(k)_{ij} - \sum_{i,j} A_{-,ij} D(k)_{ij}, \quad (3.14)$$

where in one-dimension,

$$H(D(1)) = \sum_{i,j} A_{+,ij} D(\boldsymbol{\nu}_n)_{ij} - \sum_{i,j} A_{-,ij} D(\boldsymbol{\nu}_n)_{ij} = \lambda_n, \quad (3.15)$$

it follows that the Hamiltonian is similarly linear in dimension:

$$H(D(k)) = \sum_{\alpha=n}^{n-k} H(D(\boldsymbol{\nu}_\alpha)). \quad (3.16)$$

With the spherical constraint given in Equation 3.8, we have that $H(D(\boldsymbol{\nu}_\alpha))$ is simply λ_α , the eigenvalue associated with the α -th eigenvector. Then:

$$H(D(k)) = \sum_{\alpha=n}^{n-k} \lambda_\alpha. \quad (3.17)$$

As a consequence of linearity, the Hamiltonian as a function of distance trivially decreases with an increase in dimension as long as $\lambda_\alpha \leq 0$. Our experiments on artificial and real-life networks show that the number of negative eigenvalues is usually relatively large, in any case too large for an efficient embedding of the signed network. At the same time, one can observe that each new dimension allows the nodes to be separated by greater distances, as the norm of the distance matrix will grow with each added dimension.

To compare the quality of the embedding at different dimensions, a natural choice is to combine these two quantities, and to normalize the Hamiltonian by the norm of the distance matrix, i.e. dividing by the term $\sqrt{(\sum_{i,j} (D(k)_{ij})^2)}$. This procedure defines the higher-dimensional generalized Hamiltonian:

$$\tilde{H}(D(k)) = \frac{\sum_{\alpha=n}^{n-k} \lambda_\alpha}{\sqrt{\sum_{i,j} (D(k)_{ij})^2}}. \quad (3.18)$$

In the case of the n -complete negative graph, we prove that the higher dimensional generalized Hamiltonian is minimized in $k = n - 1$ dimensions (see Appendix A.3). More generally, this choice is a heuristic, for selecting the optimal embedding dimension, and we demonstrate its effectiveness empirically in the following sections.

Based on this formulation and the intuition provided by the complete negative graph, we propose selecting k by minimizing $\tilde{H}(D(k))$ over k , thereby identifying the dimension that minimizes the ground state energy. Note that even if $\lambda_{k+1} < 0$, the addition of the $k + 1$ -th dimension may still increase the energy $\tilde{H}(D(k+1)) > \tilde{H}(D(k))$ due to the normalization factor.

This provides an alternative to the method proposed in [117], which uses the k smallest eigenvectors based on the largest eigen-gap, but formalizes the choice of k in terms of a physical energy

function. Let us emphasize again that, within our framework, the energy minimization is intended to find the dimension that produces the best proximity measure, and not the number of weak balance clusters. See the [22] for further visual illustrations highlighting the difference between these two objectives.

We evaluate our method using signed stochastic block models (SSBMs), as introduced in Chapter 2. Figure 3.2 shows the normalized energy versus dimension plots for various realizations of SSBMs with different community numbers, sizes and sign flip probabilities. Figure 3.2 (a) shows the normalized energy as a function of dimension for 3 community SSBM with edge probability 0.5, and randomly generated community sizes between 20 to 50. In all 10 realizations, the energy is minimized at $k = 2$, which is the expected dimension for representing three communities. Similarly, in Figure 3.2(b), the energy curves for noisy 3-community SSBMs (with a 0.1 sign-flip probability) also show a consistent minimum at $k = 2$.

Figure 3.3 shows a two-dimensional SHEEP embedding for two realizations of a 3 community SSBM, one with and one without noise on the edge-signs. When the number of communities is increased to six, as in Figure 3.2(c), the energy minimum shifts to $k = 5$ in all realizations, consistent with the expected dimension. For the noisy 6 community SSBMs shown in Figure 3.2 (d), the energy minimum is obtained at $k = 5$ for most realizations, but can vary depending on the faction sizes and noise parameter.

When the sign-flip probability increases, the optimal dimension may deviate from the number of clusters. This is because our method is designed for finding the best embedding to represent proximal relationships, not to recover the best node clustering, and the noise will naturally affect the proximal relationships between nodes. See [22] for an experiment illustrating the effect of noise on the optimal dimension.

3.5 Applications

3.5.1 Proximity Based Node Attributes

As the SHEEP Hamiltonian is a function of distance, this method is fundamentally designed to preserve node proximity, incorporating multi-scale information from local interactions (link sign) to global structure (factions from structural balance). In what follows, we build on this insight by demonstrating how SHEEP can be used to recover continuous node attributes. Unlike embedding methods designed for clustering, which are particularly adept at binary prediction tasks (ie, link sign prediction), SHEEP provides a continuous metric more suited to recovering continuous node attributes and determining relative extremism.

For all numerical experiments, we compare SHEEP to two other spectral signed network embedding methods, the opposing Laplacian [120] and SPONGE (Signed Positive Over Negative Generalized Eigenproblem) [53]. Like SHEEP, these methods induce a matrix operator on the graph, and

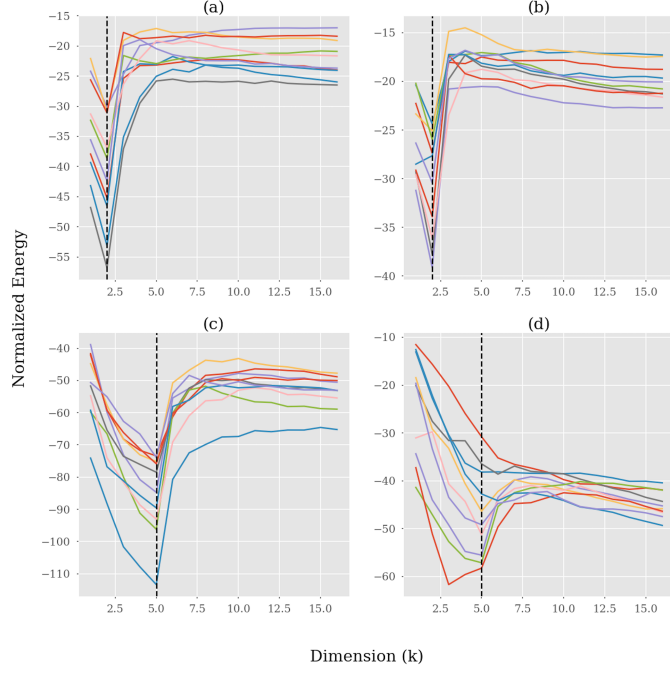


Figure 3.2: **Normalized Energy $\tilde{H}(D(k))$ vs dimension plots.** (a) Three-community SSBM (ten realizations). The energy minimum (dashed line) is obtained at $k = 2$ as expected. (b) Three-community SSBM with 0.1 sign flip parameter, the energy minimum is still obtained at $k = 2$. (c) Six-community SSBM (ten realizations), with energy minimum obtained at $k = 5$, as expected. (d) Six-community SSBM with 0.1 sign flip parameter. The energy minimum is still obtained at $k = 5$, for most realizations, but can vary depending on realization faction sizes and random variations.

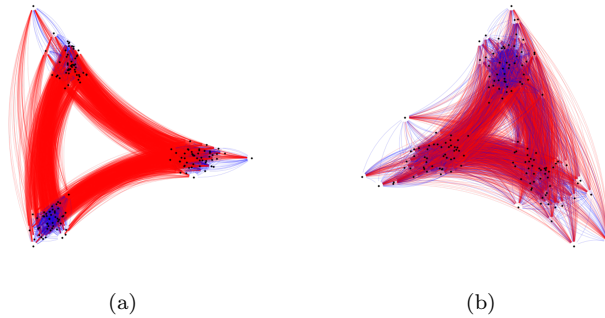


Figure 3.3: **Three-community SSBs With Varying Sign Flip Parameters.** Both figures depict the SHEEP embeddings for three-community SSBs with 50 nodes per community, and 0.5 edge probability. (a) has 0 sign flip probability and (b) has 0.2 sign flip probability.

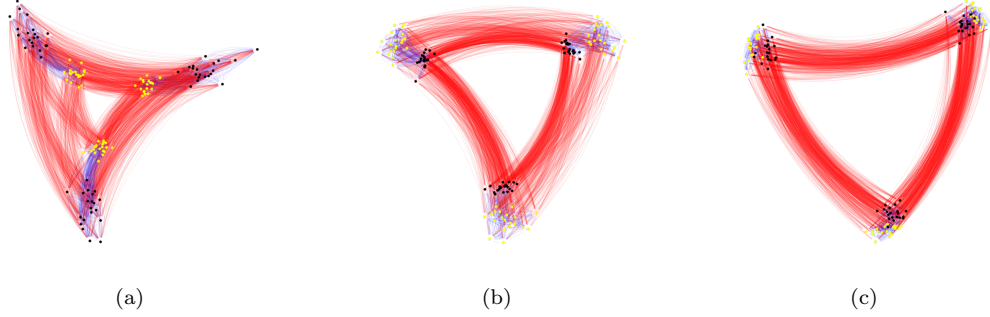


Figure 3.4: **Different embeddings for a modified three-community SSBM.** The graph is constructed such that the yellow nodes have fewer negative edges, compared to the black nodes. The embeddings are produced by a) SHEEP/Repelling Laplacian, b) opposing Laplacian, c) SPONGE.

construct an embedding from the k smallest eigenvectors. These methods were designed to identify the antagonistic factions characteristic of structural balance. In practice, the embedding dimension k is selected to match the expected number of factions, $k + 1$, and the factions are then recovered using k-means clustering in the latent space.

Figure 3.4 illustrates the distinction between the proximity-based SHEEP embedding and the two benchmark spectral clustering methods. It depicts the resulting embeddings produced for a modified signed stochastic block model (SSBM) which exhibits weak balance. Each signed community in the SSBM is separated into two subgroups: the yellow nodes have a lower density of negative edges to the other communities ($p = 0.3$), and black nodes which have a higher density of negative edges ($p = 0.7$).

The two clustering-focused methods produce embeddings that clearly delineate the three factions in the latent space, without capturing the variation between the yellow and black subgroups, or actually placing the less conflictual yellow nodes further away from the other factions, compared to the black nodes. SHEEP, however, places the yellow nodes closer to the origin, indicating the weaker negative relationships between these groups. This indicates SHEEP’s unique ability to represent the strength of node relationships, taking into account the presence (or absence) of edges, as compared to the other two methods. (See [22] for further visual examples.) In the following section, we refine this argument by quantitatively evaluating the three approaches on synthetic and empirical signed networks, with ground-truth proximity data.

Rankings on Generated Synthetic Graphs

As the SHEEP Hamiltonian is a function of distance, the sign of an edge between two nodes is related to the distance between them in the resulting embedding. In one-dimension, the positions generated by the SHEEP resembles a node ordering; nodes which are closer in ordering should have more positive connections. Here, we formulate a task to test the ability of SHEEP to recover ordinal

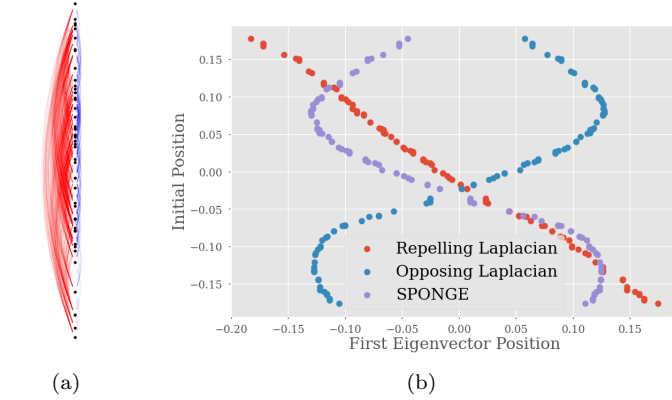


Figure 3.5: **Embedding results for well-ordered synthetic graphs.** (a) Shows a well-ordered synthetic graph, constructed from a 50-nodes complete graph, and edge signs assigned according to threshold of 0.1. (b) Results for a 100-node well-ordered graph with 0.1 threshold. The nodes initial position plotted against first eigenvector position for three spectral embedding methods, SHEEP (red), the opposing Laplacian (blue) and SPONGE (purple). There is a high correlation between the initial and SHEEP embedding position.

information, by generating a synthetic network in which edge signs correspond to a well-defined node order.

This well-ordered synthetic graph is constructed by first randomly distributing a set of points along a line, with uniform distribution between -1 and 1. The signed graph is constructed deterministically: nodes are connected by a positive (negative) edge if the distance between their position on the line is smaller (larger) than a chosen threshold. This process results in a complete signed graph, where the nodes have a continuous attribute; the value of its position along the line.

Figure 3.5 (a) shows the initial positions of 50 nodes, along with the edge signs generated for a threshold of 0.2. The one-dimensional SHEEP embedding is generated as the first eigenvector of the repelling Laplacian. To evaluate the success of the method at recovering the original node ordering, we compute the absolute value of the Kendall Tau (KT) rank correlation between the nodes' positions in the embedding and their true positions on the line. Note that we use the KT correlation over the Pearson correlation because we are interested in the ordinal relationship obtained by the embedding, which may be monotonic but not necessarily linear.

As in all subsequent experiments, SHEEP is compared to the opposing Laplacian and SPONGE. Figure 3.5(b) shows a comparison between the true node positions and the first eigenvector positions for a well-ordered synthetic graph with 100 nodes and a threshold of 0.1. Compared to the other methods, SHEEP effectively preserves the initial ordinal information. This task is repeated over several realizations of well-ordered synthetic graphs with varying parameters (see Appendix A.4). Across these trials, SHEEP consistently recovers the underlying node ordering, unlike the opposing Laplacian and SPONGE, demonstrating the usefulness of our proximity-based embedding measure.

Australian Rainfall Correlations

We evaluate our method on an empirical graph with a ground-truth proximity measure. The graph is derived from seasonal rainfall time series (in millimetres) collected at 305 stations across Australia. From this, we construct a signed network by taking the Pearson correlation between these measurements. This produces a complete signed graph with $n = 305$ nodes, each representing a station.

We impose that the graph is unweighted by masking the magnitude of the correlation, and setting each edge equal to the polarity, such that for all $(i, j) \in E$, $(i, j) \in \{-1, 1\}$. Note that the sign of Pearson correlation is not transitive, as it depends non-trivially on the magnitude of the correlation. Consequently, the resulting signed graph does not admit a perfect bi-partition. The same network was studied in [53], in which the authors used a $k = 6$ and $k = 10$ dimensional embedding to cluster the stations, finding the resulting clusters accurately reflected ground-truth geographical regions.

Minimizing the generalized Hamiltonian indicates that the optimal embedding dimension for the graph is $k = 1$. Although this may seem counter-intuitive, as spatially embedded systems typically require at least two dimensions, the result is consistent with meteorological findings. The Australian climate is largely governed by latitude, due to a high-pressure belt that migrates north and south over the course of the year, shaping seasonal rainfall patterns [56]. Because SHEEP preserves both proximal relationships and continuous node properties, we evaluate the embedding by comparing node positions to the latitude of the corresponding stations, which is closely tied to north-south spatial structure. As in the previous section, we compute the absolute Kendall Tau (KT) correlation between the first eigenvector of the repelling Laplacian (SHEEP), the opposing Laplacian, and SPONGE. The results are reported in Table 3.1, finding that SHEEP yields the highest KT correlation of the three methods. As depicted in Figure 3.6, the node position in the SHEEP embedding closely corresponds to the station latitude. While we note that the second eigenvector of SPONGE achieves a higher correlation with latitude than its first (Table 3.1), the advantage of SHEEP lies in its interpretability: proximity information is consistently captured in the first k eigenvectors. In contrast, the eigenvector that best reflects proximal structure in SPONGE varies across networks, limiting its utility in settings without access to ground-truth information.

	$\lambda_{n=305}^r$	$\lambda_{n=305}^o$	$\lambda_{n=305}^{sponge}$	$\lambda_{n=304}^{sponge}$
KT Corr. with Latitude	0.744	0.719	0.178	0.733

Table 3.1: **KT Correlation between Eigenvectors of L_r (SHEEP), L_o , SPONGE and Latitude for Australian Rainfall Correlation Network.**

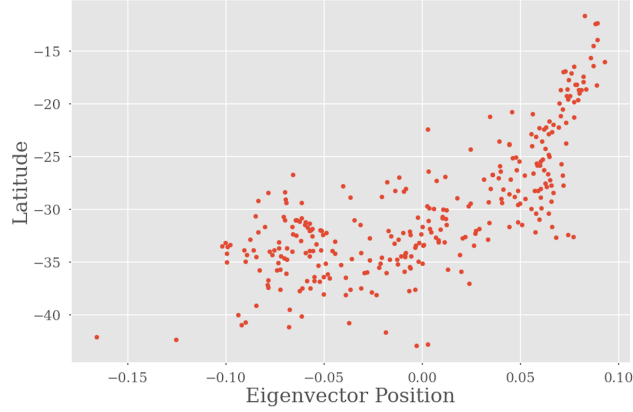


Figure 3.6: **Embedding results for Australia rainfall correlation graph.** Plot shows station latitude versus SHEEP embedding position, obtained from the first eigenvector of the repelling Laplacian. Edge weights are masked using the sign function, $(i, j) \in \{-1, 1\}$. KT correlation is 0.744.

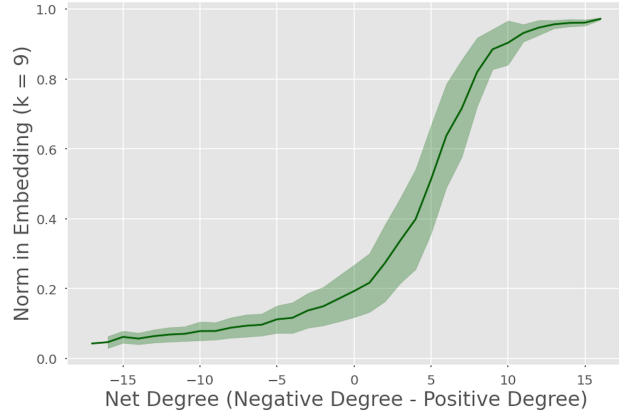


Figure 3.7: **Embedding norm vs net degree for several realizations of the signed configuration model.** Plot shows the mean distance to origin in SHEEP embedding (the norm in $k = 9$ dimensions) versus the net degree over 100 realizations of the model, for which $k = 9$ was the optimal embedding dimension. Shaded area gives the standard deviation. Nodes having relatively more negative edges are further from the origin.

3.5.2 Node Extremism Measure

Signed Configuration Model

In this section, we provide a second application of SHEEP, demonstrating that a node’s distance to the origin in the optimal embedding space gives a measure of its “extremism” or conflictuality. In general, this distance reflects both local measures, the positive and negative degrees, and global properties, such as the community structure of the graph. For instance, as Figure 3.4 demonstrates, SHEEP positions nodes closer to the origin if they have fewer negative connections to opposing factions.

Thus, the distance to the origin is determined by both the cluster structure of the graph, and by the intensity of the negative interactions between the nodes. In random graph models, such as a configuration model, local and global measures tend to coincide. The classical configuration

model generates a random graph from a given degree sequence, by randomly matching each edge “stub” with equal probability [154]. We modify this construction to the signed setting by generating two unsigned configuration models, and subtracting the adjacency matrix of one from the other to obtain a new signed random graph. Using both this signed configuration model, and SSBMs with varying levels of noise, we investigate the relationship between the node’s distance to the origin in the SHEEP embedding, the global graph structure, and the net degree (negative degree minus positive degree).

The signed configuration model used here is obtained by subtracting the adjacency matrix of one unsigned configuration model from another, each generated independently with 50 nodes and random degree sequences drawn from integers between 1 and 20. Over 5000 realizations of the signed configuration model, the most frequently observed optimal embedding dimension using SHEEP is $k = 9$. To ensure we compare across graphs with the same optimal dimension, in what follows, we restrict our analysis to 100 realizations of the signed configuration models for which $k = 9$ was optimal. For each node we compare the net degree to its distance from the origin in the latent space, i.e. the norm of its 9-dimensional position vector. Figure 3.7 shows the mean distance as a function of net degree, with the shaded region indicating the standard deviation across nodes and realizations. We observe an increasing relationship, where nodes with a higher proportion of negative edges are more “extreme” and thus further from the origin.

In the case when the graph is not random and has structure produced by structural balance, the norm gives a measure of extremism that cannot be explained by the net degree alone. In the Appendix A.5, we show that for networks which obey strong balance, the norm is determined primarily by the negative degree of the node. When the bi-polarization structure of the graph is perturbed by noise or other structural effects, the extremism measure can no longer be explained by the negative degree. Thus, when the structural balance assumption does not hold, SHEEP produces an extremism measure that reflects both local properties (such as degree) and global structural influences, adapting to the complexity of the graph.

Continuous Political Ideology

Finally, we evaluate the node extremism measure on an empirical network with ground-truth continuous node attributes, providing clearer intuition for how the measure reflects real-world node characteristics. The network, obtained from [18], represents (positive and negative) relationships between members of the USA House of Representatives for different congress sessions (here, we focus on sessions 110 and 144). A positive (negative) edge indicates that the two representatives have co-sponsored statistically significantly more (fewer) bills than expected by chance, giving some indication of their political alignment.

Given our choice to consider more recent congresses, which Aref et al. argue in [18] are highly polarized along Democrat-Republican lines, it is not surprising that the optimal SHEEP embedding dimension is $k = 1$. As SHEEP is designed to preserve proximal relationships in a continuous setting, we evaluate whether our method can recover members’ political ideologies on a continuous scale and examine whether distance from the origin in the embedding space correlates with political extremism.

Following the approach taken in [18], we compare our results against the Nokken-Poole (NP) ideology scores as a benchmark. The Nokken-Poole ideology scores place each member of congress on a continuous scale from +1 (conservative) to -1 (liberal). The score is frequently used in political science literature, and is the result of a three step multi-dimensional scaling method based on member’s voting habits in a given congress, derived by maximizing their utility [41, 157]. The Nokken-Poole scores are two dimensional, but the first dimension is often taken to represent political ideology along a left-right spectrum.

As in our previous experiments, we compare SHEEP to the opposing Laplacian and SPONGE. In this polarized setting, the opposing Laplacian and SPONGE would likely outperform SHEEP at detecting the two clusters associated to the Democrat and Republican parties, while SHEEP recovers the continuous ideology information reflected in the Nokken-Poole scores.

The 110th congress session is represented by a signed graph of $n = 452$ nodes; restricting to the largest connected component yields a graph of $n = 450$ nodes. The correlation between the eigenvectors and the NP scores are reported in Table 3.2. For this congress, the first eigenvector of the repelling Laplacian (SHEEP) has the highest correlation with the NP scores, compared to the other methods. We note that the fourth eigenvector of SPONGE obtains a better correlation with the NP scores compared to the first. Figure 3.8 (a) depicts the one-dimensional embedding of the network obtained by SHEEP, and the nodes are coloured by their NP scores.

The 114th congress session forms a signed graph of $n = 446$ nodes. Likewise, taking the largest connected component results in a signed graph with $n = 443$ nodes. The correlations obtained between the eigenvectors and the NP scores are reported in Table 3.3.

Again, SHEEP obtains the highest correlation with the NP scores, but we note that for this network, the first eigenvector of SPONGE performs better than the following eigenvectors. Observe that the position of the best-performing SPONGE eigenvector varies across different graphs. Identifying which eigenvector captures this proximal information is non-trivial, especially when ground-truth scores are unavailable. For this, SHEEP is superior, as the information is contained in first k eigenvectors deemed optimal by the embedding method. Figure 3.8 (b) shows the one-dimensional SHEEP embedding of the network, where the nodes are coloured according to the NP scores. Since the optimal embedding dimension for the 114th congress is $k = 1$, the KT correlation shows directly

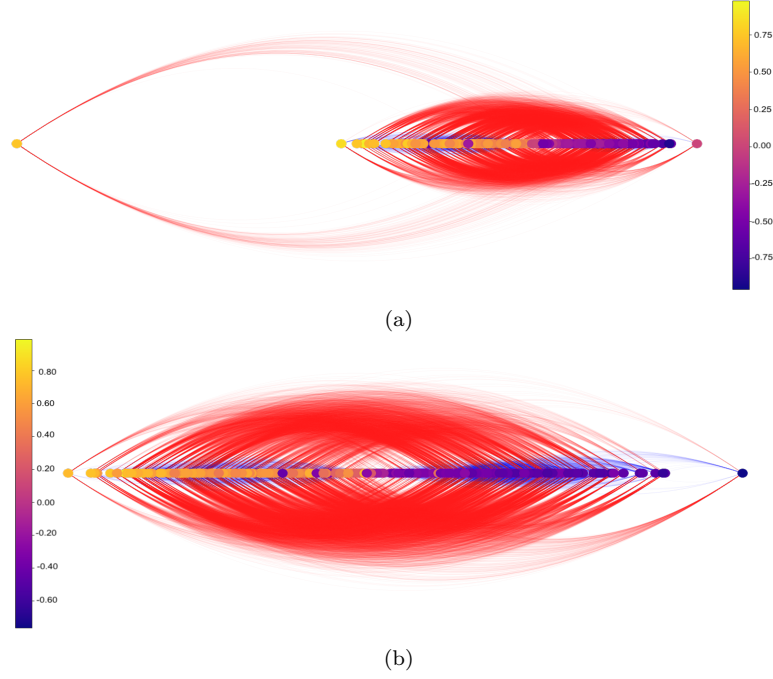


Figure 3.8: **One dimensional SHEEP embeddings of two congresses.** Congress 110 (a) and Congress 114 (b) signed networks. Nodes colours indicate Nokken-Poole political ideology scores obtained from [18].

the relationship between the node’s distance to the origin and its political extremism, using the NP scores.

	$\lambda_{n=450}^r$	$\lambda_{n=450}^o$	$\lambda_{n=450}^{sponge}$	$\lambda_{n=447}^{sponge}$
KT Corr. with NP Ideology Score	0.685	0.432	0.209	0.640

Table 3.2: **KT Correlation between Eigenvectors of L_r (SHEEP), L_o , SPONGE and Nokken-Poole Ideology Scores for USA House of Representatives Network, Congress Session 110**

	$\lambda_{n=443}^r$	$\lambda_{n=443}^o$	$\lambda_{n=443}^{sponge}$
KT Corr. with NP Ideology Score	0.698	0.382	0.626

Table 3.3: **KT Correlation between Eigenvectors of L_r (SHEEP), L_o , SPONGE and Nokken-Poole Ideology Scores for USA House of Representatives Network, Congress Session 114**

As the two congressional graphs are neither random, like the signed configuration model, nor perfectly bi-polarized, like the SSBMs, the node extremism measure represented by the NP scores cannot be entirely explained by either the net degree or the negative degree alone. Neither of these measures are sufficient for determining the extremism of the node; the positive edges and the graph structure both play a role, which is captured in the SHEEP embedding. Consequently, the SHEEP embedding is more highly correlated with the NP scores, as compared to the negative degree alone.

For a more detailed analysis, we refer the reader to the Appendix A.5.1.

SHEEP produces a node extremism measure which arises from the interplay between the negative degree, positive degree and graph structure. In random graphs, like the signed configuration model, local and global measures coincide, and the norm in the embedding is determined by the net degree. For graphs with perfect faction structure, the norm is determined by the negative degree. SHEEP is most useful for investigating the intermediate cases, graphs which are neither completely random, nor completely polarized. In such cases, SHEEP provides a meaningful way of understanding node extremism, as demonstrated by the application to political ideology.

3.6 Conclusions

In this chapter, we presented SHEEP, a spectral embedding algorithm for signed networks which characterizes node relationships by taking into account multi-scale information. The method is based on a physically inspired model: we construct a Hamiltonian that assigns attractive and repulsive forces to the positive and negative edges in the graph. We show that the Hamiltonian is intrinsically related to the graph’s repelling Laplacian, and that finding the minimum energy configuration reduces to an eigenvector problem. We further demonstrate that the resulting ground state energy, or minimum eigenvalue, is a statistical measure of the graph’s degree of strong balance. We extend our results to higher dimensions, presenting an energy-based approach to locating the optimal embedding dimension for the network. We propose an application of our measure to recovering continuous node attributes, demonstrating that SHEEP retains ordinal information on synthetic and empirical networks. We also show that the distance to the origin in the optimal embedding dimension gives a measure of node “extremism”, which is related to the interplay between local and global structure. Overall, this work contributes to the growing body of literature on spectral methods for signed networks, while simultaneously looking beyond the framework of structural balance.

We note that this method extends quite naturally to weighted signed networks as well. Given the form of the Hamiltonian, a weighted adjacency matrix is equivalent to loosening the requirement that each spring has a unit-weight spring constant, allowing the embedding to account for weighted proximity. In Chapter 6, we take precisely that approach, exploring the applications of SHEEP to a signed weighted social media dataset. We find that SHEEP provides fresh and additional insight into the political dynamics at play on the platform.

Chapter 4

Modelling random walks over signed networks

The chapter is based on the article *Babul, Shazia'Ayn, Yu Tian, and Renaud Lambiotte. "Strong and weak random walks on signed networks." npj Complexity 2.1 (2025): 4. [21]*

4.1 Introduction

Random walks play a central role in several aspects of network science [139]. Random walks are often used as a linear, tractable model for diffusive processes on networks, allowing us to explore questions such as the impact of certain network properties [125], including their community structure or their degree distribution, on the speed of the diffusion via the mixing time [134]. Random walks are also a key ingredient in several methods to extract information from large-scale networks. Important areas of application include node centrality which can be captured by the density of walkers on that node in the stationary state [87], community detection where groups of nodes can be defined as communities based on their tendency to retain flows of walkers for long times [177, 123], or the design of kernels capturing the similarity between nodes on a graph [78, 93].

Among the advantages of random-walk based approaches, one can cite their intuitive interpretation, in terms of flows along the edges, their connection to well-established theories, such as that of Markov chains, their linearity, opening up the toolbox of linear algebra, and their ease of use in a distributed setting, each of the walkers being independent from the others, and their capacity to probe the structure of the network at multiple scales.

In this chapter, our purpose is to explore how random walks can be defined on, and probe the structure of, signed networks. Random walks on signed networks have been studied in the past, either implicitly, by investigating the properties of the opposing Laplacian for node embedding [120] or explicitly, by defining rules for how traversing a negative edge affects the random walker [194], in a model that we will refer to as a strong random walk because it draws from the paradigm of strong balance.

As with many generalisations of “classical networks” techniques, as we observed in Chapter 3 with the Laplacian, different choices could be made to extend the notion of random walks to signed networks. The choice should be motivated either by the process that one aims to model or by the underlying structure that one seeks to extract. Random walks on unsigned networks are powerful due to their deep connection to specific structural patterns, communities and bottlenecks [46]. Communities can be seen as noisy connected components, i.e. connected components to which random edges have been added. The guiding question to define the *right* model of random walks on signed networks is thus: what is the prototypical structure that one aims to identify?

As in Chapter 3, the large-scale structure of signed networks is usually viewed through the prism of balance theory. It is well-known that the opposing Laplacian and the equivalent strong random walk allow for the efficient recovery of the bi-polarization in networks exhibiting strong balance [120]. It is also well-known that they may fail to properly characterise networks exhibiting a structure related to weak balance, and thus cannot serve as a model for the inference of a structure into more than two groups [198]. Alternative approaches have been proposed to solve this problem, for instance through other notions of Laplacian, e.g. SPONGE which can successfully cluster weakly balanced graphs even when sparse [53] or the SHEEP method presented in Chapter 3, which is most informative in the case of networks which are not completely balanced. These other Laplacians, however, do not have a random walk interpretation and thus do not inherit their advantages. The main contribution of this chapter is the definition of a random walk process, called the weak random walk, specifically designed for structures derived from weak balance, hence particularly fitted for clustering problems involving more than two groups in signed networks. In essence, the process consists of allowing the walker to explore the graph around a seed node only until it encounters a negative edge for a second time.

In this chapter, we introduce a first type of random walk on signed networks, called the strong walk, and show its connections to the opposing Laplacian. We discuss how to exploit such a random walk in order to define a similarity matrix between nodes of a signed network. We also identify the limitations of the strong walks for networks exhibiting more than two clusters, and introduce weak walks to remedy the problem. Finally, we use both random walks to construct similarity measures between nodes, directly leveraging these for downstream tasks such as clustering and semi-supervised classification. We compare the results of simple unsupervised and semi-supervised algorithms based on the two types of walks, showing the improvement in accuracy gained by the weak walks for certain families of networks. This contribution follows directly from Chapter 3, in that both chapters aim to probe the multi-scale structure of signed networks, but they do so from complementary methodological perspectives and with different objectives in mind. While SHEEP takes a spectral approach, the present approach is grounded in random walks, allowing us to probe the graph’s structure through a dynamical process. Moreover, SHEEP is explicitly designed for settings in

which the network is not fully balanced, as the method produces a proximity-based embedding rather than one designed to find clusters. The weak walk method, as we shall see throughout this chapter, is constructed with weak balance clustering in mind, and admits downstream applications for identifying clusters in weakly balanced networks.

4.2 Related Literature

In what follows, we consider an undirected signed graph, characterised by the adjacency matrix $A_{s;ij}$, following the notation presented in Chapter 2. As before, this matrix can be decomposed into its positive and negative parts as: $A_{s;ij} = A_{+;ij} - A_{-;ij}$. We refer to the degree of the node as the total number of connections of each node, positive or negative, is given by $k_i = \sum_j |A_{s;ij}|$, and define the diagonal matrix of degrees $D_{ij} = k_i \delta_{ij}$. We use $A_{u;ij} = A_{+;ij} + A_{-;ij}$ to denote the adjacency matrix of the graph where signs are not considered. Our purpose is to model diffusive processes on signed networks. Here, we will focus on discrete-time walks, but everything can be generalised to continuous times. We also only consider unweighted networks in the derivations and experiments to follow, but the generalisation to weighted ones is straightforward, where the edge weight can be interpreted as biasing the transitions of the walker.

As discussed briefly in Chapter 2, a discrete-time random walk on an unsigned network evolves such that a walker located on some node i jumps to one of its neighbours, with a probability equal to one over the number of neighbouring nodes. The density of walkers on the nodes is determined by the (unsigned) transition matrix $T_u = D^{-1}A$, of the associated Markov chain.

A vast body of network science literature is devoted to exploring the spectral properties of the transition matrix, and its relationship to the associated graph structure [139]. In particular, we note that T_u shares the same spectrum as the symmetric matrix $D^{-1/2}A_u D^{-1/2}$, which implies that its eigenvalues are all real and that its left and right eigenvectors, $\mathbf{u}_\alpha$ and $\mathbf{v}_\alpha$ respectively, are related by $\mathbf{u}_\alpha = D\mathbf{v}_\alpha$. The eigenvalues λ_α lie within the interval $[-1, 1]$, and the multiplicity of the dominant eigenvalue gives the number of connected components in the graph. The difference between λ_1 and the largest non-one eigenvalue, $\lambda_1 - \lambda_2$, called the spectral gap, determines the mixing time of the process and quantifies the presence of bottlenecks hindering the flow of walkers across the system [125]. When the graph is connected, its corresponding left eigenvector, denoted by π is given by $\pi_i = k_i/(2M)$, which characterises how walkers are distributed across the nodes in the stationary state.

Random walks are at the core of several algorithms designed for determining node centrality, community detection, or network embedding [123, 165]. They are particularly useful to capture the strength of indirect connections between nodes in situations when the graph is sparse, so that only a small fraction of the pairs of nodes are connected, which is the case for the vast majority of large-scale empirical networks. In these cases, random-walk-based kernels can be used to estimate the

proximity of pairs of nodes that are not directly connected, by incorporating, in an unbiased (or in a supervised) way, the information of walks of any length. In case of unsigned networks, the resulting similarity between two nodes is obtained from an appropriately weighted sum of the walks between them. Typically, the existence of many short walks between two nodes ensures their proximity. Important examples include the diffusion, or heat kernels [11, 127, 186], the mean commute time [77, 180], but also the random-walk with restart similarity [161, 195, 196]. For reasons that will appear clear later, the random walks in our model over signed networks tend to stop after a finite number of steps, and thus the random walk with restart generalises most naturally in the signed context.

Random walks with restart are a special type of random walk with teleportation. Take a node i , we consider a process where a walker may, at each step, either perform a jump to one of the neighbouring nodes with probability p , or teleport back to the original node i with probability $(1 - p)$. It is straightforward to show that its stationary state is given by

$$\mathbf{n} = (1 - p)\mathbf{e}_i(I - pT_u)^{-1}, \quad (4.1)$$

where $\mathbf{e}_i$ is a row vector made of zeros except for entry i where it takes the value 1. The stationary state is, intuitively, peaked around the pinned node i , and becomes more spread, giving more importance to distant neighbours, when p increases. When p approaches 1, greater importance is given to longer walks, and the vector converges to the stationary state π defined above. After taking a Taylor series of this expression, it is very similar to the kernel used in the Deepwalk embedding (see Chapter 2), and defines the weighted neighborhood of each node at different scales [171].

A similarity matrix can be obtained by repeating the process for each node and by stacking the N vectors, leading to

$$\kappa(p) = (1 - p)D(I - pT_u)^{-1}, \quad (4.2)$$

where we scale here each line by the node degree, to ensure that the matrix is symmetric. This quantity, or variations around it, have been used in a variety of contexts [78], including node classification and clustering.

Note that

$$\begin{aligned} \kappa_{ij}(p) &= (1 - p)k_i[(I - pT_u)^{-1}]_{ij}, \\ \kappa_{ij}(p) &= (1 - p)k_i\left[\sum_k p^k T_u^k\right]_{ij}. \end{aligned} \quad (4.3)$$

As each term $k_i T_{u,ij}^k$ takes the (i, j) element from $A_u D^{-1} A_u \dots D^{-1} A_u$, the similarity matrix $\kappa(p)$ is symmetric for any value of p . The symmetry comes from detailed balance, which holds when teleportation is done in proportion to the stationary distribution of the random walk without

restart. Moreover, the i^{th} line (row) sums to k_i , meaning that $\kappa(p)$ can be interpreted as a weighted graph, interpolating between the original graph (with the addition of self-loops) when p approaches 0 and low rank matrix obtained from the outer product of the stationary states, $k_i k_j / (2M)$, when p approaches 1. For any value of p , $\kappa_{ij}(p)$ can be viewed as a weighted network and be fed directly into standard methods for such graphs. For instance, its Newman-Girvan modularity [86] can be used to uncover its multi-scale modularity, an approach reminiscent of Markov stability [58, 121, 183].

4.3 Strong walks

4.3.1 Introduction and basic properties

To generalize the classical methods discussed in the previous section, and more generally, any algorithm based on random walks, it is necessary to first define a random walk on a signed network. A natural way is by considering the dynamics of two types of walkers, positive walkers and negative walkers, which allows us to maintain and quantify the notion of flow between nodes, while still distinguishing between walks that are positive or negative in nature [194].

In the following, we will therefore start by describing how the density of positive walkers, $n_{+,i}(t)$, and that of negative walkers, $n_{-,i}(t)$, on each node i evolve in time, so as to define different versions of random walks on signed networks. In this section we present one possible extension of the classical random walk to signed networks, based on the opposing Laplacian, which is often used in the literature [194].

The strong random walk, as we call it, is defined as follows. Positive and negative walkers perform a random walk by taking an edge, independently of its sign. If the walker takes a positive edge, it keeps its own sign; else it changes sign. The rate equation for the process is:

$$\begin{aligned} n_{+,j}(t+1) &= \sum_i (n_{+,i} T_{+,ij} + n_{-,i} T_{-,ij}), \\ n_{-,j}(t+1) &= \sum_i (n_{-,i} T_{+,ij} + n_{+,i} T_{-,ij}), \end{aligned} \quad (4.4)$$

where the transition matrices,

$$\begin{aligned} T_{+,ij} &= A_{+,ij} / k_i, \\ T_{-,ij} &= A_{-,ij} / k_i, \end{aligned} \quad (4.5)$$

indicate the transition probabilities for a walker from i to j and whether to keep or change the sign respectively. In matrix notation:

$$\begin{aligned} T_+ &= D^{-1} A_+, \\ T_- &= D^{-1} A_-. \end{aligned} \quad (4.6)$$

If we consider the dynamics of the walkers independently of their sign and consider the time evolution of $n_i(t) = n_{+,i}(t) + n_{-,i}(t)$, one recovers the classical equation for random walk on networks

$$n_j(t+1) = \sum_i n_i(t) T_{u;ij}, \quad (4.7)$$

where, as before, $T_u = D^{-1}A_u$. If, in contrast, we consider the time evolution of the difference $\Delta_i = n_{+,i} - n_{-,i}$, a quantity that can be understood as the total “charge” on node i , we find

$$\Delta_j(t+1) = \sum_i \Delta_i(t) T_{s;ij}, \quad (4.8)$$

where $T_s = D^{-1}A_s$, is the signed transition matrix from which one can re-construct the opposing Laplacian (see Chapter 3), whose spectral properties are closely associated to the notion of strong balance and have been used for node embeddings and clustering [120, 198].

Another interesting perspective is to consider the sets of equations (4.4) as one single system of equations driven by the matrix

$$T = \begin{pmatrix} T_+ & T_- \\ T_- & T_+ \end{pmatrix}. \quad (4.9)$$

The latter is simply the (unsigned) transition matrix of a larger graph, composed of $2N$ nodes (each node appears twice, once for positive walkers, and once for negative walkers), with adjacency matrix

$$A = \begin{pmatrix} A_+ & A_- \\ A_- & A_+ \end{pmatrix}. \quad (4.10)$$

The whole diffusive process can thus be understood in terms of the spectral properties of that transition matrix. The system of equations decouples when performing the following change of variable: for each i , we define $\Delta_i = n_{+,i} - n_{-,i}$ and $n_i = n_{+,i} + n_{-,i}$, where the determinant of the transformation matrix is 2, leading to the transition matrix

$$T' = \begin{pmatrix} T_s & 0 \\ 0 & T_u \end{pmatrix}, \quad (4.11)$$

for these variables. The spectral properties of T' are therefore determined by those of the two matrices T_s and T_u , with well-known special cases, e.g., when the graph is disconnected or (strongly) structurally balanced [194].

4.3.2 Walk lengths

The strong walk can be equipped with a teleportation mechanism in order to keep walkers concentrated in the neighbourhood of a target node. Given a teleportation probability $(1 - p_s)$, we consider the expected walk length before a teleportation. We know that if $p_s = 1$, the strong walk will have infinite expected length, thus we assume $p_s < 1$ here. For each node i , if the walk has length k , it means that teleportation does not happen at the first $(k - 1)$ time steps but the k th time step, thus the number of steps that teleportation does not occur follows a negative binomial distribution with

success probability being $1 - p_s$ and the number of success being 1. Hence, the expected walk length is the sum of the expected steps that teleportation does not happen plus 1, i.e.,

$$\mathbb{E}[l|X_0 = i] = \frac{p_s}{1 - p_s} + 1 = \frac{1}{1 - p_s}.$$

4.4 Strong walk based similarity matrix

Having defined the strong random walk over the signed network, we consider how these dynamics can be exploited to extract structural information from the graph. As in Section 4.2, a similarity matrix can be obtained from a random walk over an unsigned network, based on the principle that nodes connected by many short walks are structurally close, which is aptly captured by the flow of walkers between them. This concept can be naturally extended to signed networks, where similarity depends not only on the flow of walkers between nodes, but also on the valence of those walks. This motivates the construction of a signed similarity matrix that distinguishes between positive and negative walks. This can be accomplished precisely by using the strong random walk, resulting in a metric that follows the intuition from the notion of strong balance, and provides the main interface between the signed random walk dynamics and downstream tasks such as clustering.

A graph is said to be strongly balanced if it does not possess any cycle with an odd number of negative edges. Equivalently, this graph can be partitioned into two sets of nodes such that the positive edges are within each cluster, and the negative edges are between the two. This also implies that any walk between two nodes of the same set possesses an even number of negative edges, while any walk between two nodes of different sets possesses an odd number of negative edges. With this intuition in mind, we follow a similar procedure as in Section 4.2 required to produce the similarity matrix constructed from random walks with restart over an unsigned graph. For each node i , we consider a process where with probability p_s , walkers will jump to one of its neighbouring nodes, and possibly change sign, or with probability $1 - p_s$ teleport back to the pinned node. During the teleportation, the walker automatically becomes positive, independently of its sign. This choice is motivated by the fact that a node should be similar to itself. As before, the process defines for each node i a vector, but it is now a $2N$ vector, where the first N entries encode the density of positive walkers (arriving through walks with an even number of negative edges, or positive walks), while the last N entries encode the density of negative walkers (arriving through negative walks).

The walker begins as a positive walker on some node i , $\mathbf{e}_i$, a $2N$ vector. The state of the walker after k steps, $\mathbf{u}_k$ can be written in terms of the walker after $k - 1$ steps $\mathbf{u}_{k-1}$ as,

$$\mathbf{u}_k = (1 - p_s)\mathbf{e}_i + p_s\mathbf{u}_{k-1}T, \quad (4.12)$$

according to the dynamics defined by the transition matrix T . The stationary state of this process is defined by

$$\mathbf{n} = (1 - p_s)\mathbf{e}_i(I - p_sT)^{-1}. \quad (4.13)$$

The first N entries of $\mathbf{n}$ encode the flow of walkers on positive walks from i to the other nodes, hence capturing similarity between the nodes. The last N entries encode the flow of walkers on negative walks from i to the other nodes, hence encoding dissimilarity between the nodes. More precisely, stacking the first N elements of this vector defines the i -th row of a similarity matrix κ_+ . Stacking the last N elements of this vector defines the i -th row a dissimilarity matrix κ_- .

The two matrices can be seen as the positive and negative adjacency matrices of a graph where the entries incorporate information from longer walks that explore a larger area of the network, when p_s increases. Taking the difference between the matrices defines a signed similarity matrix, which encodes the relative importance of positive and negative walks between nodes. Note that in Eq. 4.13, we give the same importance to each node, in contrast to the Eq.4.2, as there is no straightforward way to make the similarity matrix symmetric by a weighting of the nodes. Other choices could be made, for instance, by weighting the nodes by their degree. In practice, in experiments where the full kernel is needed, we compute the signed similarity matrix first and then symmetrize it by adding it to its transpose.

4.5 Weak walks

4.5.1 Definition and basic properties

The walk (Eq. 4.4) appears to be appropriate to uncover the similarity of two nodes from a strong balance perspective, as walks with an even number of negative edges will increase the similarity between the nodes, while walks with an odd number of negative edges will decrease their similarity. However, under the weak balance paradigm, where graphs are k -clusterable, the strong walk may lead to conflictual information. Recall that a graph is weakly balanced if it does not possess cycles with one single negative edge. Figure 4.1 displays a small weakly-balanced network in which the two highlighted nodes are connected by both a positive (blue) and negative (red) path, according to the dynamics of the strong walk.

The weak balance condition, defined by the appearance of exactly one negative edge, suggests defining dissimilarity between nodes based on the existence of walks with one single negative edge. Any walks involving more than one negative edge are instead uninformative, and should not be considered when estimating the similarity between two nodes.

These arguments necessitate a modification of the strong walk, by imposing the condition that a walker can become negative only once, and disappearing upon encountering a second negative edge. This argument causes an asymmetry in the node states, as positive and negative walkers now play a different role. The dynamics is encoded by the enlarged transition matrix

$$T_w = \begin{pmatrix} T_+ & T_- \\ 0 & T_+ \end{pmatrix}, \quad (4.14)$$

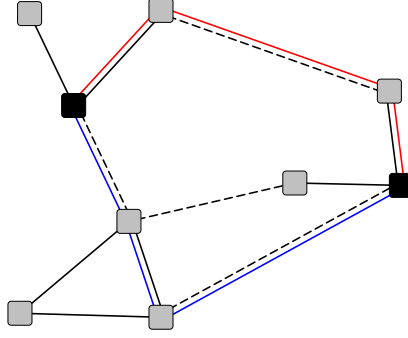


Figure 4.1: **Strong walk leads to conflictual signals.** When estimating the similarity between two nodes (marked in black) in a signed network with more than 2 clusters, the strong walk may display conflictual signals. The nodes are connected via both a positive walk (blue) and a negative walk (red). In this figure solid edges are positive; dashed edges are negative. Weak walks only consider walks with at most one negative edge and thus properly consider the two nodes as dissimilar.

After several iterations, the matrix retains its upper triangular structure, and such that:

$$T_w^n = \begin{pmatrix} T_+^n & U_n \\ 0 & T_+^n \end{pmatrix}, \quad (4.15)$$

where $U_n = \sum_{i=0}^{n-1} \binom{n-1}{i} (T_+)^i (T_-) (T_+)^{n-1-i}$. As defined here, the process leads to a leak of probability, as walkers disappear after taking a negative edge for the second time, leading to an asymptotically vanishing number of walkers (except in trivial cases).

To prevent the probability leakage and to retain the stochastic nature of the process, we adopt a random walk with restart; instead of disappearing, a negative walker will teleport back to its initial node after taking a negative edge.

Unlike in the case of the strong walk, where additional teleportation is necessary to keep the walks finite, the proposed weak walk is already finite, except in trivial cases (e.g., a connected component with only positive edges). Moreover, the length of the walks before the restart is not controlled by a parameter, but instead by the structure of the signed network.

Suppose we start with a positive walker on node x , given by the initial state vector, $\mathbf{e}_x$, where this is a vector of length $2N$. Then, the our proposed weak random walk with restart would be defined by the matrix:

$$T_x = \begin{pmatrix} T_+ & T_- \\ M_x & T_+ \end{pmatrix}, \quad (4.16)$$

where M_x is equivalently the zero matrix except in column x , which is set to the vector $\mathbf{v}$, defined as:

$$v_i = \sum_j T_{-;ij}. \quad (4.17)$$

This ensures that the matrix T is stochastic, as it has row sum equal to 1.

The block M_x enforces the behaviour that the negative walker will return to node x as a positive walker when it encounters another negative edge. The walk evolves by taking the product of the initial state vector $\mathbf{e}_x$ and the matrix T_x .

Although not necessary to keep the walks finite, the walk can also be modified to include an additional teleportation parameter, p_w . At each step, the walker will teleport back to the pinned node as a positive walker with additional probability $1 - p_w$, in addition to the restart in the case of a negative walker encountering a negative edge. For the same value of the teleportation parameter p_w , the weak walk exhibits a higher teleportation rate as compared to the strong walk, and thus a decrease in the expected walk length, as we will show.

Working from this framework, we can construct a signed similarity matrix in the same way as for strong walks. For a random walk starting as a positive one on node x , the state of the weak walker after k steps, $\mathbf{u}_k$ can be written in terms of the walker after $k - 1$ steps $\mathbf{u}_{k-1}$ as,

$$\mathbf{u}_k = (1 - p_w)\mathbf{e}_x + p_w\mathbf{u}_{k-1}T_x, \quad (4.18)$$

in much the same fashion as in Equation 4.12.

4.5.2 Walk length

We now characterise the weak walks from the perspective of expected walk length. We consider a weak walk where, as before, teleportation can happen at each step with probability $1 - p_w$, or upon a negative walker encountering another negative edge. Thus, unlike in the case of the strong walk, the walk length distribution must depend on the network structure.

Beginning with the enlarged transition matrix in Eq. 4.14. We encode the expected number of visit to each node in the matrix $Q \in \mathbb{R}^{2N \times 2N}$, where nodes are duplicated to separate the case when the node is visited via positive or negative walks. Then,

$$Q = \sum_{k=0}^{\infty} p_w^k T_w^k = (I - p_w T_w)^{-1},$$

where I is the identity matrix. Hence, the expected walk length starting from node i is

$$\mathbb{E}[l|X_0 = i] = \sum_{j=1}^{2N} Q_{ij}.$$

Note that we only count the first N rows in Q . It follows that the expected walk length for weak walk is

$$\mathbb{E}[l] = \sum_{i=1}^N \mathbb{E}[l|X_0 = i]P(X_0 = i).$$

We validate the theoretical results with Monte Carlo simulations on three-community Signed Stochastic Block Models (SSBMs; see Chapter 2 for details). These SSBMs are generated with edge

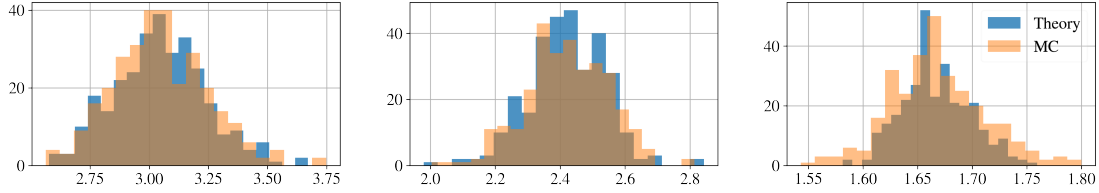


Figure 4.2: **Distribution of the expected walk lengths.** Walks start from different nodes, and the figure shows both the theoretical results in blue and Monte-Carlo simulations in orange on three-community symmetric SSBM with 100 nodes in each community, $p = 0.05$ and different levels of teleportation (left: $p_w = 1$; middle: $p_w = 0.8$; right: $p_w = 0.5$, where the strong walk with $p_s = 0.8, 0.5$ has expected walk length 5, 2, respectively).

probability 0.05 and zero sign flip probability, such that the graphs obey weak balance. To simulate the walks, we consider a uniform initial distribution to select the initial node, and perform $n_s = 1000$ random walks from each starting node to. The walk length is obtained by the number of steps until either teleportation happens or the walker encounters the second negative edge. Figure 4.2 displays the distribution of the expected and simulated walk lengths from various nodes in the graph.

4.6 Applications

In this section we present an application of the weak random walk to signed network clustering. As is common practice in the signed network literature, we evaluate the method’s performance at detecting antagonistic groups of nodes, following the framework of structural balance. This task is most challenging when the graph deviates from perfect structural balance, as is often the case in real-world datasets. While both the strong and weak random walks are motivated by the intuitions provided by strong and weak balance respectively, their use is not limited to perfectly balanced graphs, and as we will see, we find them to be applicable and informative even on incompletely balanced graphs.

We first visually inspect the resulting signed similarity matrices generated by the weak and strong walks for two small empirical signed networks, for which ground truth clusters exist. Subsequently, we use the random walk matrix to perform semi-supervised clustering on a series of synthetic SSBM networks, with varying levels of noise, community numbers, and link densities. For this task, we compare our proposed weak walk against two other state-of-the-art methods, (1) the strong walk and (2) the Signed Random Walk with Restart (SRWR), a three-parameter signed network walk designed for semi-supervised clustering on signed networks [109]. The SRWR model considers a random walker that changes its sign as it takes positive or negative edges according to two parameters, as well as a teleportation parameter, and generates a ranking of nodes in signed networks which can be used downstream to perform signed link prediction with high accuracy. We nevertheless find that the weak walk outperforms the other two methods in certain key scenarios, despite being parameter free.

4.6.1 Computational Design

In the experiments to follow, we adopt a computationally efficient construction of the signed similarity matrices; we approximate the stationary state by taking the state of the walk vector at $k = n$, where n is large, and the walk is converging to stationarity. We can compute the state vector at each step with an iterative process, thus solving for the similarity matrix as a whole in one computation.

As matrix multiplication scales well, this proves more efficient than employing matrix inversion as required by strong walk Equation 4.13. Moreover, the weak walk transition matrix defined in Equation 4.18 is specific to each starting node x , meaning that computing the full similarity matrix via inversion would require N matrix inversion operations. Thus our choice here reduces the computational complexity of the algorithm, particularly important for large-scale graphs.

Under this construction, we define the strong and weak similarity matrix as follows. Beginning with a positive walker on node x , we set:

$$\mathbf{u}_0 = \mathbf{e}_x. \quad (4.19)$$

For the strong walk, at each step the new vector $\mathbf{u}_k$ is computed from $\mathbf{u}_{k-1}$ as:

$$\mathbf{u}_k = (1 - p_s)\mathbf{e}_x + p_s\mathbf{u}_{k-1}T, \quad (4.20)$$

where T is the $2Nx2N$ strong walk transition matrix, and p_s is the strong walk teleportation parameter.

For the weak walk, the process is defined as in Equation 4.18, although it is computationally easier to use the non-Markovian matrix T_w to compute $\mathbf{u}_k$ from $\mathbf{u}_{k-1}$ as:

$$\mathbf{u}_k = (1 - (p_w\mathbf{u}_{k-1}T_w) \cdot \mathbf{1})\mathbf{e}_x + p_w\mathbf{u}_{k-1}T_w, \quad (4.21)$$

where p_w is the weak walk teleportation parameter, and $\mathbf{1}$ is the all-one vector, thus $(p_w\mathbf{u}_{k-1}T_w) \cdot \mathbf{1}$ sums over the probability that has already been allocated to the corresponding neighbours and the first term re-sets the lost probability to the original node x as a positive walker.

The advantage of using an iterative process over the matrix given in Equation 4.16 is that in practice the computation can be made for several initial states simultaneously, while the matrix in Equation 4.16 is specific to the state where the walker begins, i.e. node x . However, the processes are equivalent since the term M_x in Equation 4.16 returns all probability leakage back to node x as a positive walker, exactly as the first term in Equation 4.21.

Finally, if the state of the walker (strong or weak) is given by $\mathbf{u}_k$, the similarity between node x and any other node is given by the vector $\mathbf{s}_x$:

$$\mathbf{s}_x = \mathbf{u}_{k=n}[1 : N] - \mathbf{u}_{k=n}[N + 1 : 2N], \quad (4.22)$$

since we subtract the negative walks from the positive. For the experiments on the empirical networks for which we calculate the entire similarity matrices, the kernel is made symmetric by adding it to its transpose.

In our experiments, we tune the strong walk teleportation parameter p_s , by sampling across a range of values and selecting the one with the best performance for each graph. However, we fix $p_w = 1$ for the weak walk, ensuring that the weak walk teleportation rate is determined entirely by the graph structure. This design choice keeps the weak walk parameter-free and emphasizes its structural sensitivity.

4.6.2 Clustering on Empirical Networks

Highland Tribes

Despite increasing prevalence of signed network research, there is a scarcity of empirical datasets annotated with ground truth cluster labels. Here, we demonstrate the distinction between the strong and weak walk kernels on two small networks for which such labels exist.

The first network encodes inter-tribal relationships, representing the friendly (positive) or hostile (negative) relations between 16 tribes based in the Eastern Central Highlands of New Guinea [172]. These tribes can be clustered into three broader communities, which characterise the higher order groupings of the society [96].

Figure 4.3 (left) depicts the adjacency matrix of the Highland Tribes network, where the nodes, representing the tribes, have been ordered according to their higher-order community assignment. While the network is not perfectly balanced, it tends towards weak balance, (see [15] for a detailed characterization of the network’s local and global structure). As the network has more than two communities, and thus is not strongly-balanced, it provides precisely the setting in which the weak walk is expected to outperform the strong walk at detecting the three communities.

Figure 4.3 provides the strong and weak walk kernels (middle, and right, respectively) generated by performing the random walks for 500 steps. Observe that the strong walk kernel exhibits far more incorrect signs as compared to the weak walk, particularly showing several positive connections between the second and third communities.

As the walks produce signed similarity matrices, they can be used downstream, analogously to the adjacency matrix, as input into community detection methods. Consider, for example, the spin-glass community detection method for signed networks based on simulated annealing, which takes a signed matrix representation of the graph, and returns community assignments in an unsupervised fashion [200]. This method typically takes the adjacency matrix as input; it equally could be used in conjunction with the strong or weak walk kernel. When ground-truth clusters are available, the accuracy of the returned labels can be evaluated using the Adjusted Rand Index (ARI), a similarity measure between two clusterings compared against random chance, and which ranges

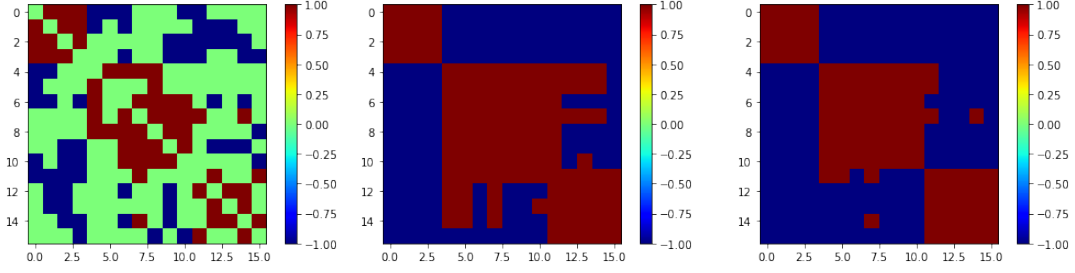


Figure 4.3: **Matrices representing the Highland Tribes network.** (left) Adjacency matrix (centre) Strong walk kernel resulting from a 500 step walk with teleportation parameter $p_s = 0.7$. (right) Weak walk kernel resulting from a 500 step walk with no teleportation $p_w = 1.0$. Positive and negative values have been set to $+1$ or -1 respectively for clearer visualization.

from -1 (dissimilar), 0 (random) and 1 (perfect alignment) [84]. Using the ARI enables a direct comparison across the three methods.

As our goal here is to compare the performance of the weak walk against the strong, we take the parameter free weak walk with no additional teleportation. This is assessed against the strongest performance of the strong walk across all parameter choices. Specifically, we select the optimal teleportation parameter for the strong walk by inputting its kernel into the spin-glass algorithm, comparing the community assignments with the ground truth, and selecting the parameter that yields the highest ARI. For the Highland Tribes graph, the optimal parameter is $p_s = 0.7$, corresponding to the matrix displayed in Figure 4.3 (middle). In this case, as the graph is quite small, dense and almost balanced, the clustering task is straightforward; all three matrices combined with the spin-glass algorithm yield identical community labels, achieving an ARI of 1.0 .

Sampson’s Monastery

In this section, we present the three kernels obtained for another small empirical signed network with ground truth labels. The Sampson’s Monastery network is derived from an ethnographic study conducted within a New England monastery, focusing on the interpersonal dynamics between the 18 novice monks. Based on his observations, Sampson, the ethnographer, designated each relationship between the monks as positive (friendly) or negative (antagonistic), and categorized the men into four groups according to the signed community structure he observed [182].

Although this network is quite small, comprising only 18 nodes, the graph is more sparse than the Highland Tribes network, and has more noise in terms of edge signs that violate structural balance (see [15] for a comprehensive analysis of the network’s structural balance). Figure 4.4 (left) depicts the network adjacency matrix, where the nodes, representing the novice monks, have been arranged according to the four communities designated by Sampson.

As this network has more than two communities, the weak walk should outperform the strong walk at amplifying the community structure in the network. Figure 4.4 shows the strong and weak

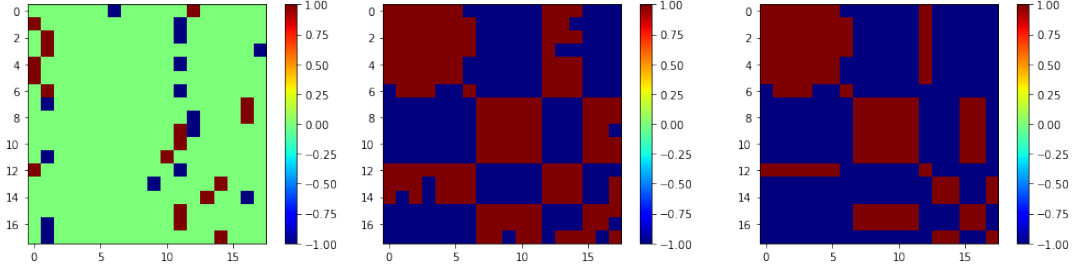


Figure 4.4: **Matrices representing the Sampson’s Monastery network** (left) Adjacency matrix (centre) Strong walk kernel resulting from a 500 step walk with teleportation parameter $p_s = 0.8$. (right) Weak walk kernel resulting from a 500 step walk with no teleportation. Positive and negative values have been set to +1 or -1 respectively for clearer visualization.

walk kernels (middle, and right) generated by performing the random walks for 500 steps. As in the previous example, we employ the parameter-free weak walk kernel without teleportation. For the strong walk kernel, we determine the optimal teleportation parameter by maximizing the ARI, following the same procedure used for the Highland Tribes network. In this case, the optimal teleportation probability is found to be $p_s = 0.8$. The strong walk kernel exhibits far more incorrect signs as compared to the weak walk, generating several positive connections between the second and third communities, and the third and fourth communities. When combined with the community detection spin-glass algorithm, the parameter-free weak walk kernel obtains the best ARI (0.639) compared to the adjacency matrix (0.442) and the strong walk kernel (0.518).

4.6.3 Semi-supervised Clustering on SSBMs

In this section, we use the kernels generated by the weak walk, the strong walk and the Signed Random Walk with Restart (SRWR) to perform semi-supervised clustering on several families of Signed Stochastic Block models (SSBMs) which exhibit signed community structure, modulated by a noise parameter ν . (See Chapter 2 for a brief introduction to SSBMs). As these graphs are generated as SSBMs, all nodes are associated with a ground-truth community label by design.

This task is semi-supervised in nature; we use prior knowledge of the labels of a select subset of nodes, and use this information to label the rest of the graph, for which the labels are initially hidden. The semi-supervised clustering for each random walk is methodologically similar to that used in the PageRank algorithm [126]; we begin with (positive) walkers on a set of labelled seed nodes. The similar vector generated by the (signed) walk is used to obtain a ranking of each of the other nodes in the graph. Each unclassified node inherits the label of the seed node for which it achieved the highest ranking.

In these experiments, we use the parameter-free version of the weak walk without teleportation ($p_w = 1$). The strong walk has one adjustable parameter controlling the teleportation (p_s), while SRWR has three parameters: two which control if the walker changes sign when it takes a negative

edge or not, as well as a teleportation parameter. For the latter two walks, we sweep through the parameter space for the strong walk and the SRWR and report results corresponding to the parameters that achieve the highest ARI.

Consequently, these methods perform better than the weak walk in certain instances, but they may be less useful in practice, since the ideal parameters are not known a priori and indeed tend to change with each iteration of the SSBM, even when the SSBM parameters are held constant. Moreover, we show that the weak walk, despite being parameter free, exhibits competitive and, in some cases, superior performance in certain cases.

To evaluate the accuracy of the clustering assignment obtained in each case we compare against the ground truth clusters by taking the ARI. These methods can be applied both to completely balanced and incompletely balanced graphs, and the experiments in this section provide a sense of the lower bound on the degree of balance required for these methods to work.

Strong Balance: Two-Community SSBMs

As a benchmark, we first apply the semi-supervised clustering to a graph characterized by strong balance; a two-community SSBMs with 100 nodes in each community, and 0.05 probability of link existence between each pair of nodes. Clustering is performed on multiple realizations of this graph, with varying levels of frustration, controlled by the sign-flip parameter ν . The strong and weak walk similarity vectors are derived from the state vectors of the walks at $k = 100$ steps. Analogously, for the SRWR, we set the max iterations to 100 to obtain the similarity vector.

The other parameters for the SRWR and strong walk are chosen by parameters sweep, and we find that the optimal parameter differs across each iteration of the graph. For each noise parameter ν , we generate ten independent graph realizations, perform the semi-supervised clustering, and report the mean and standard deviation of the ARI as a function of ν . The results, given in Figure 4.5, reveal that the SRWR performs the best, a feature to be balanced with the fact that it has the most parameters. The strong walk outperforms the weak walk; this aligns with the expectations of strong balance, and this walk benefits from the additional advantage of teleportation.

Weak Balance: Three-Community SSBMs

We now shift our focus to the setting of weak balance. We focus here on three-community SSBMs with 100 nodes in each community, and 0.05 probability of link existence between each pair of nodes. As in the previous example, we perform the semi-supervised clustering on multiple realizations of the graph with varying levels of frustration. As before, the strong and weak walks and SRWR similarity vectors are all computed with 100 iterations. The parameters for the strong and SRWR walks are chosen by parameters sweep, changing for each iteration of the graph.

The mean ARI and corresponding standard deviation across 10 realizations for each value of ν are given in Figure 4.6 (left). We observe that the SRWR again outperforms the other two methods.

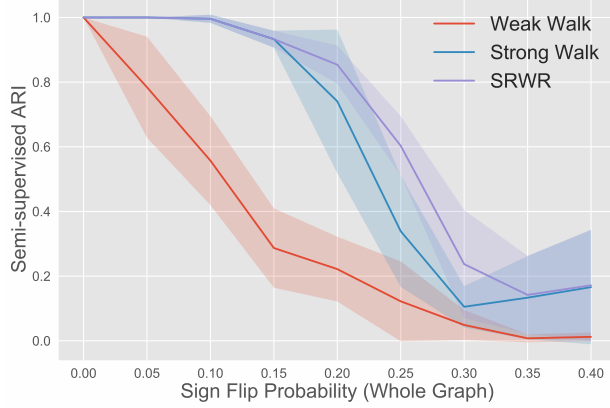


Figure 4.5: **Results on two-community SSBM with strong balance.** ARI from the semi-supervised clustering for the weak walk, strong walk and SRWR for various graph realizations and sign flip probabilities.

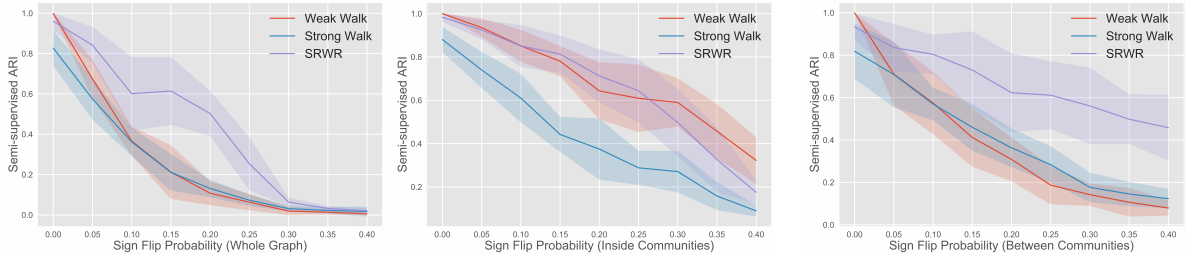


Figure 4.6: **Results on three-community SSBM with weak balance.** ARI from the semi-supervised clustering for the weak walk, strong walk and SRWR for various graph realizations and sign flip parameters. (left) Frustration probability applied to whole graph. (middle) Frustration probability restricted to inside communities. (right). Frustration probability restricted to between communities.

The strong walk and weak walk exhibit similar performance in this weak balanced setting, despite the strong walk having one extra parameter.

In empirical networks, noise may be present along positive and negative edges, and may occur between or within communities with different probabilities. To model this behaviour, we introduce parameters $\nu_{inside} \in (0, 1)$ and $\nu_{between} \in (0, 1)$ which separately control the level of frustration on intra-community and inter-community edges, respectively. We investigate the results in two additional settings when the frustration is restricted to 1) inside the communities by varying ν_{inside} and setting $\nu_{between} = 0$, or 2) between communities by varying $\nu_{between}$ and setting $\nu_{inside} = 0$.

The weak walk is more robust to noise inside communities (Figure 4.6 (centre)), and indeed achieves similar performance to the SRWR. When the noise is between the communities (Figure 4.6 (right)), the weak walk is less robust and the SRWR outperforms both the strong and weak walk methods.

Weak Balance: Six-Community SSBMs

The task of clustering on signed networks increases in difficulty with the number of clusters. We perform the same task as in the previous experiment, extended to a six-community SSBMs with 50

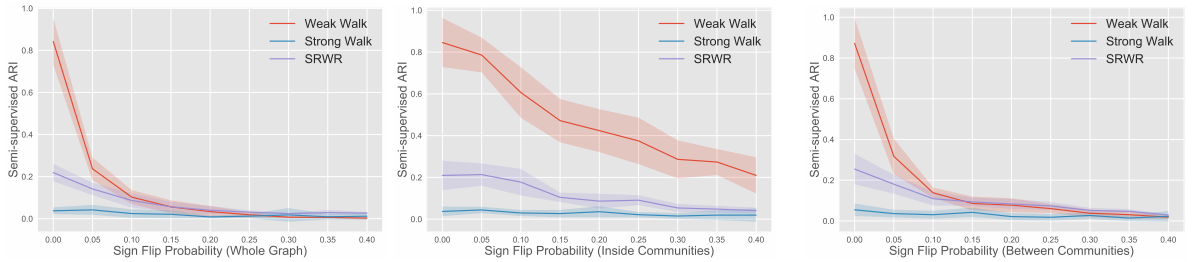


Figure 4.7: **Results on six-community SSBM with weak balance.** ARI from the semi-supervised clustering for the weak walk, strong walk and SRWR for various graph realizations and sign flip parameters. (left) Frustration probability applied to whole graph. (middle) Frustration probability restricted to inside communities. (right). Frustration probability restricted to between communities.

nodes in each community, and 0.05 probability of link existence between each pair of nodes. Due to the larger size of these graphs, we increase the maximum number of iterations (or walk length) to $k = 300$, but as before the other parameters are selected by parameters sweep. The results are provided in Figure 4.7 (left). In this setting, the weak walk outperforms the strong walk and the SRWR, although the performance decreases as the noise level increases.

Following the same procedure as in the previous section, we investigate the results when the frustration is restricted to either inside the communities, as in Figure 4.7 (centre) or between the communities, as in Figure 4.7 (right). The weak walk indeed proves once again to be more robust to noise inside communities, compared to when the noise is between the communities, outperforming both the strong walk and the SRWR by a larger margin.

The distinction between the methods is more pronounced in the six-community setting. The weak walk, despite being parameter-free, outperforms the other methods when the weak balance structure is clearly present, highlighting its effectiveness under such conditions.

Asymmetry

In the literature on signed network clustering, algorithm performance is commonly benchmarked on SSBMs generated with uniform inter-community link density. In practice, in empirical settings, this assumption may not hold; the likelihood of connections between communities can vary significantly. In the case of a social network, links may be more likely to exist between two communities due to a higher interaction probability, which can be influenced by external factors such as physical proximity. To model such settings we introduce here the term *asymmetric SSBM*, which represents an SSBM with non-uniform inter-community link density.

We define an asymmetry parameter a which modulates the inter-community link density. For a three-community asymmetric SSBM the adjacency matrix would take the form,

$$A = \begin{bmatrix} 0.05 & 0.05a & 0.05 \\ 0.05a & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 \end{bmatrix}.$$

One realization of the adjacency matrix for a three-community asymmetric SSBM with 100 nodes in each community and asymmetry parameter $a = 0.2$ is depicted in Figure 4.8.

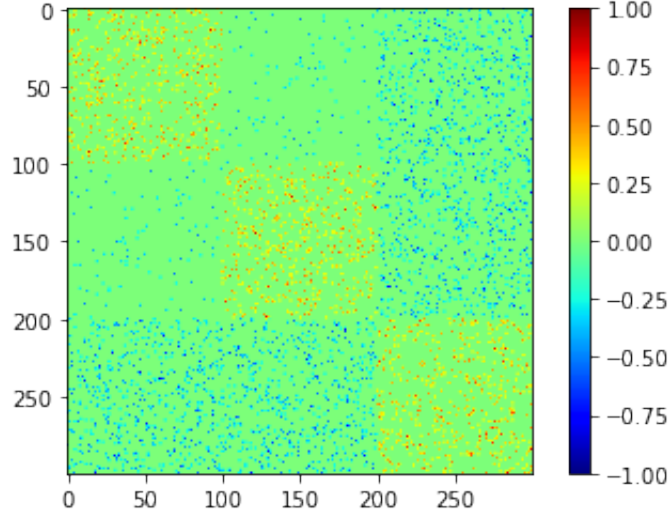


Figure 4.8: **Adjacency matrix of example asymmetric SSBM.** This is a 300 node three-community asymmetric SSBM, with asymmetry parameter $a = 0.2$.

Here, we perform the same task outlined in the previous sections on 300 node three-community SSBMs with varying levels of asymmetry. As before, the strong and weak walks and SRWR similarity vectors are all computed with 100 iterations. As depicted in Figure 4.9, when the graph exhibits asymmetry, the weak walk outperforms both other methods by a large margin.

This can be explained by understanding that a walk between nodes in the first two communities can take two different paths; 1) by taking a negative edge between the communities, or 2) by passing through the third community, a walk which requires two negative edges. This second walk cannot happen in the case of the weak walk, where the walker teleports back to the seed node upon encountering a second negative edge. Thus, when the asymmetry is high, when a is close to 0, the latter case is more likely. This consequently introduces noise into the strong walk kernel, which does not occur in the case of the weak walk.

4.7 Conclusions

Random walks are foundational in network science, and are at the core of several algorithms, with applications from community detection to network embedding. Random walks allow us to explore the indirect patterns of connectivity between nodes, and their spectral properties are directly related to important network patterns, such as the cut size. Defining the specific rules governing a random walk process is non-trivial, and tuning the parameters of the model, e.g. to favour network exploration in depth versus in width, may lead to significant improvements, as was demonstrated by node2vec

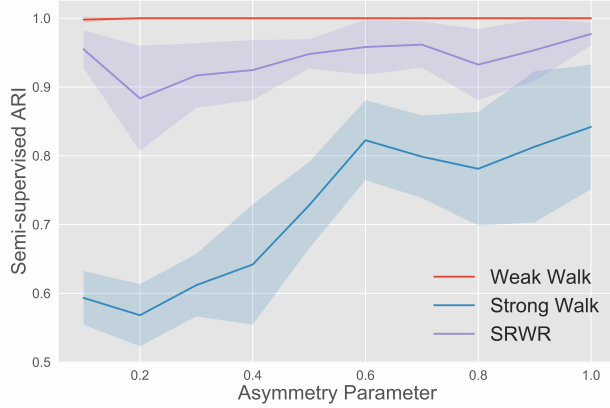


Figure 4.9: **Results on three-community SSBM with varying asymmetry.** ARI from the semi-supervised clustering for the weak walk, strong walk and SRWR.

[93], for instance. More generally, the choice of a random walk model should be driven by the type of structure that one aims to isolate or by the nature of the diffusive processes taking place over an empirical graph [124].

This work contributes to the growing family of random walk models over signed networks, seeking one that is “optimal” for the weak balanced setting. We argue that the random walk process associated with the standard matrices for the mining of signed networks, e.g. the opposing Laplacian, is inherently (and implicitly) biased to the search of a strong balance structure. When a system exhibits multiple clusters, as is characteristic of weak balance, we argue that another random walk model, called the weak walk, is more compatible with the structure.

After introducing the model and exploring its properties, we have constructed three classical random-walk based measures of node similarity, based on the strong walk, the weak walk and the so-called SRWR walk respectively. Through a series of experiments on synthetic and empirical data, formulated as unsupervised and semi-supervised clustering tasks, we showed that the weak walk outperforms the strong walk when the graph has more than two communities, or exhibits asymmetry in the density of links. We have also shown that it is competitive against the three-parameter SRWR walk in many cases as well, despite being parameter-free.

A wide range of algorithms for network mining rely on trajectories generated by a random walk model. In designing the similarity matrix, we made a specific, yet classical, choice for aggregating the multiple walks connecting a pair of nodes. However, alternative formulations are possible; for example, the similarity matrix could be defined by summing over the states of the walks at each step, by taking the state vector after k steps, taking the stationary state of the walk as in PageRank [126], or by combining the results from walks of different lengths in a non-linear fashion to give greater weight to longer or shorter walks as in the similarity kernel introduced in [78].

Future research perspectives for this method include feeding weak walk trajectories instead of strong walks ones into other algorithms and exploring applications to other typical signed network

tasks such as unsupervised clustering [198] and link prediction [135]. As a specific example, the strong walk has been used to learn stances for users in signed social networks [168]. Combining weak walks with these state-of-the-art methods could yield, at a minimal developing cost, an improvement in performance and flexibility for empirical networks, especially when the assumption of strong balance does not hold, or when the graph exhibits asymmetry in link density. Another challenge for future research is the identification of the "right" number of factions when confronted by real-world data, in order to guide the user towards methods building on strong or weak walks in practice.

Chapter 5

Social contagion of music taste over heterogeneous social ties

This chapter is based on the article *Babul, Shazia'Ayn*, Desislava Hristova*, Antonio Lima, Renaud Lambiotte, and Mariano Beguerisse-Díaz. "Link Me Baby One More Time: Social Music Discovery on Spotify." Proceedings of the International AAAI Conference on Web and Social Media. Vol. 19. 2025. (* indicates joint first authors)*[106]

5.1 Introduction

A significant branch of network science is devoted to the study of dynamical processes that take place *on top* of a network, where the network topology shapes the evolution of these processes. Traditionally, these models are fundamental for both extracting information about the network structure, using random walks, or for simulating spreading processes, finding applications in models of disease transmission or the dissemination of information or behaviours over social networks [139, 155]. These processes are typically formulated on networks with uniform edges; however, they can be adapted for heterogeneous edges, as in the case of the signed random walk we presented in the previous chapter. In this chapter, we shift our focus towards a different type of heterogeneity, examining how ties that differ in intensity, modality or reciprocity, affect the dynamics of diffusion. Rather than relying on a simulated diffusion process, we present an empirical study that characterizes the spread of music taste over a social network.

In general, models of diffusion of ideas, behaviours or other innovations over social networks often are inspired by epidemic spread, by relating the activation probability to the social network structure and the speed of the spread [155]. Such models are frequently employed to understand information diffusion in an online context [94]. Other social processes require social reinforcement to spread effectively, as in Granovetter's threshold model, where activation probability increases as a function of number of previously activated neighbours [89]. The successful adoption of behaviours requiring social reinforcement, called complex contagion, is deeply impacted by network topology,

and spread more easily over clustered networks [44]. Empirical evidence for complex contagion has been found to describe behaviours from information sharing on social media [149], political campaign donations [199], person-to-person product recommendations [130], and even the decision to exercise [14].

In these works, there is often no evidence of “contact” events between individuals, which must instead be inferred from the network topology and sequence of activations. In the absence of contact event information, it can be difficult to understand which spreading process is occurring. Several works have proposed methods for inferring the process type and parameters using network topology and the observable activations [43, 66]. Others have taken a global approach, to understand how diffusion processes can lead to cascades on a social network [188, 129].

The notion of contagion itself differs from context to context. In epidemiology, the notion is rather straightforward; a contagion event refers to a healthy individual contracting a disease following a contact event with an infected individual [152]. In the theoretical network science literature, concerned with the dynamics of adoption of behaviours, contagion refers to peer-to-peer influence, where the diffusion patterns are created by casual social influence between neighbours [13]. In empirical settings, however, causality is often difficult to establish, and observed “contagion processes” frequently reflect temporal correlations shaped by individual attributes and shared environments [184]. In this chapter, we adopt an observational, network-based notion of contagion, where contagion refers to an increased probability of adoption following a social exposure, conditional on the structure and characteristics of the social network. In particular, we investigate how music spreads over an online social network, which is characterized by diverse ties. The work here elucidates the social aspect of music sharing, describing how person-to-person recommendations facilitated by established social relationships lead people to discover and appreciate new music. Our objective is to identify the network and contextual factors that drive the diffusion of music taste in a heterogenous networked setting.

To that end, we use a set of data obtained from Spotify, an audio streaming platform where users can listen to music and share it with others. Online streaming platforms have transformed the ways that people discover, listen, and share music with each other [204, 148]. Social music sharing is facilitated by a function that allows users to send Spotify content links over various social media and messaging platforms, leading directly back to the Spotify application. When these person-to-person recommendations are successful, the receiver engages with the shared artist’s music, potentially becoming a new fan over time, here measured as an active engagement with the artist following the share. We elucidate the social and contextual factors that lead to a successful recommendation (or “activation”), creating a predictive model for short-term engagement, and examining the spread of new music discoveries over the Spotify social network.

There have been few works applying spreading models specifically to the musical context. A recent study showed how live music events generated increased listenership amongst concert attendees and their friends [192]. Another observational study compares song download time-series to epidemic spread curves to identify mechanisms that contribute to the spread of a song [176]. This macro-level approach is common in other works on empirical datasets across different contexts, seeking to observe evidence of complex contagion on social networks using just the adoption curves [75].

In contrast to other related works, we have evidence of the sharing events between users, in the form of link shares, as well as information about both the sender and receiver, including prior listening history, relationship history, and their role in the social network. This information enables us to investigate how these social and contextual factors affect the probability of a successful social recommendation.

Our work here, like that of [130], investigates person-to-person recommendations. These processes are distinct from those studied in the majority of works on social cascades which focus on information propagation, such as cascades over social media. In these cases, the action of “adopting” the post is intertwined with the action of socially spreading it - by clicking to re-share. In our case, we are instead interested in how the behaviour of an individual changes after receiving a recommendation - the two actions, adoption and propagation are therefore distinct, and sharing requires a higher level of individual effort.

In [48], the authors characterise different diffusion protocols based on level of individual effort and social cost of not participating. Sharing music to a friend has a high individual effort, while there is a low social cost to engaging more with the shared artist. The authors assert that such protocols spread more effectively with social reinforcement, and over stronger social ties. Peer influence has been posited to be more effective over stronger and reciprocal ties [7], and this has empirically been found to be the case in an experiment studying political mobilisation on Facebook [34]. As part of this work, we explore the effect of relationship strength between the individuals sharing music, finding results consistent with the role of strong social ties.

We examine three key propositions regarding the social dynamics of music discovery through link sharing on Spotify. First, we expect that diffusion is more likely to happen between individuals who display overlapping music preference, based on their historical listening history prior to the contact event. Second, that spread is more likely to happen along stronger, established ties and through more intimate modes of contact, such as direct messaging, rather than broadcast over a public social media. This hypothesis is based on the established literature on social influence. As the underlying social network is heterogeneous, we are able to explore various definitions of social tie strength, such as the number of social interactions, and the mode of contact. And lastly, that the diffusion is more likely to take place in the presence of “social cohesion” amongst the receiver’s social contacts.

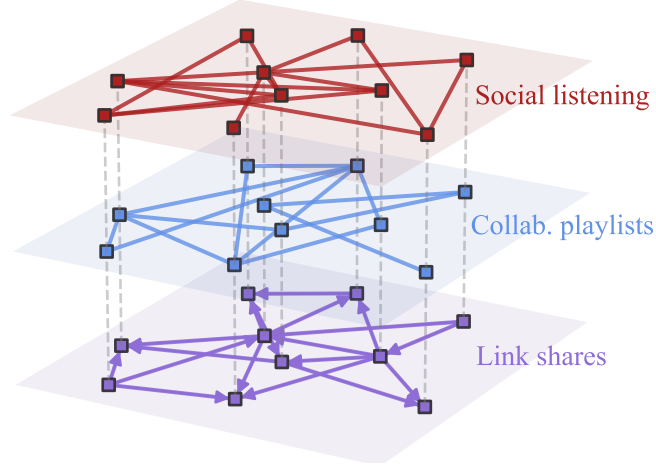


Figure 5.1: **A synthetic network representing the structure of Spotify social network.** The Spotify social network is a multiplex network, each layer corresponding to a different social interaction. Social listening sessions and collaborative playlists represent undirected layers, while link sharing is a directed layer.

Inspired by complex contagion literature, we examine the proportion of the receiver’s contacts who are already engaged with the shared artist, the degree of network clustering, and the overlap between the sender’s and receiver’s social circles. Our findings suggest that musical preferences diffuse most readily between users with similar tastes, but are also facilitated by strong social ties, direct modes of contact, and the reinforcing effects of social cohesion.

In this chapter, we first describe the dataset used in this empirical study, as well as the method used to sub-sample the data. We describe how we classify a successful person-to-person recommendation using the post-share engagement. Next, we investigate the three hypotheses by analysing the relationship between future engagement probability and the relevant features. Finally, we build a Random Forest model to predict short-term engagement following a link share event, and explore the relative importance of the contributing features to identify the ones most predictive of a successful diffusion event.

5.2 Data and Methodology

Spotify has over 574 million users, and over 100 million available tracks [187]. The platform also offers a social aspect; users have various ways of interacting with each other directly on the platform, from following the activity of their friends, sharing music, to collaboratively building playlists together. For this study, we use datasets describing the listening and sharing history of users on the platform. We construct the underlying social network from the social interactions that users engage with on the platform.

5.2.1 Spotify Social Network

The Spotify social network comprises its users (nodes) and their interactions on the platform (edges). Each edge has an associated timestamp, indicating when the interaction occurred. This thesis is primarily concerned with networks characterized by heterogeneous ties. The first two chapters dealt with signed networks; while the Spotify social network is unsigned, it exhibits heterogeneity in a number of other respects.

The network is itself multiplex, each layer representing a different type of interaction that can occur over the platform; collaborative playlists, social listening sessions and link shares. A collaborative playlist is a special type of playlist that can be modified by a group of users. Although these are group interactions, we represent them as multiple pairwise interactions (edges) between each pair of collaborating users. The edges representing these interactions are undirected, as they are a reciprocal interaction. A social listening session allows multiple users to add songs to a real-time queue, listening to the same music either remotely or together. Like collaborative playlists, these are reciprocal interactions, and the edges are undirected. A link share is a link to a track, album or artist sent from one user to another via a messaging or social media platform. This interaction is directed, and is represented in the network by an edge from the sender to the receiver.

Figure 5.1 gives a depiction of this layered structure on a synthetic network. Each layer can be described by an adjacency matrix L , an $N \times N$ matrix where N is the number of users. The values of the matrix L_{ij} are the number of interactions occurring from user i to user j . In the case of an undirected layer, where the interactions are reciprocal, the matrix is symmetric and $L_{ij} = L_{ji}$.

The resulting social network is constructed from three layers, the first represents the social listening sessions, an undirected layer with adjacency matrix L_s . The second layer represents the collaborative playlists, an undirected layer with adjacency matrix L_c . And, the final layer represents the link sharing, a directed layer with adjacency matrix L_l .

When aggregating the adjacency matrices of all the layers as $A = L_s + L_c + L_l$, the network has $\sim 2.2\text{B}$ edges between $\sim 518\text{M}$ users (nodes). The aggregation is performed in different ways, depending on the metric being computed. More details are provided in the Results and Discussion section.

5.2.2 Link Sharing Events

We focus here on “contact” events, defined as instances where one user attempts to spread their music taste to another user, accomplished via sharing a link to a Spotify track over a social media or messaging platform. These shares give a direct signal of a spreading event taking place on top of the social network, which we use to investigate the social mechanisms which influence the music discovery process on Spotify.

Even in the presence of explicit social sharing, the user experience on Spotify is shaped by the platform’s recommender algorithm, which influences not only music discovery, but also the sequences of songs consumed over time [10]. To mitigate the influence of Spotify’s recommender system on observed user behaviour, we make several methodological choices as follows, designed to capture intentional rather than algorithmically-driven behaviours.

We restrict our analysis to “discovery” share events, where the receiver has not listened (on Spotify) to the artist of the shared track prior to receiving the link. Our definition of “discovery” follows from previous literature exploring user behaviour on Spotify, the first time a user comes into contact with a piece of content. In [148], the authors make a distinction between discovery that is programmed, where a song is chosen by the recommender system, or *organic*, when the song is searched and selected by the user. We introduce here a third form of discovery, which occurs via peer-to-peer recommendation.

The data is further filtered to include only link shares of tracks for which 1) the link was opened (comprising 50% of all shares), and 2) the receiver’s Spotify account registered a playback event of the shared content for at least 30 seconds (64% of all opened shares). This ensures that the contact event was significant enough for the spread to be possible; the receiver did indeed come into contact with the music for a significant enough time to establish their response as meaningful.

To account for popularity bias, we randomly sample up to 1M share events stratified by artists across various popularity bins. Our final sample consists of ~ 2.72 M events across 3,102 distinct artists, where the share event occurred between April-June 2023.

Each link sharing event contains anonymised identifiers for the sender and receiver, as well as identifiers of the track, album and primary artist. The share event dataset also includes track metadata, including the artist’s popularity on Spotify at the time of the share, and the album release age at the time of the share. Moreover, the data captures the context of the contact event; the social media application type used to send the link (e.g. WhatsApp), the timestamp of the share and the timestamp at which the link was opened. The identifiers and timestamps are used to construct the dataset of link sharing events, which are the basis of our analysis. The track metadata and share context are features used in our analysis and predictive model, as detailed in Table 5.1.

We delineate two types of social media application, direct or broadcast, corresponding to one-to-one or one-to-many modes of communication. The applications we identify as direct are: WhatsApp, Facebook Messenger, SMS, Line, Instagram Direct, and Samsung Messenger. The applications we identify as broadcast are: Instagram Stories, Facebook Feed, Facebook Stories, and X, formerly Twitter. Some sharing events have an unknown destination platform or fit into neither category. These events are excluded from the section of analysis involving share application mode. We recognize that the applications identified as “direct” also have functionalities supporting one-to-many communication (group chats), but the key distinction here is in the magnitude of the number of

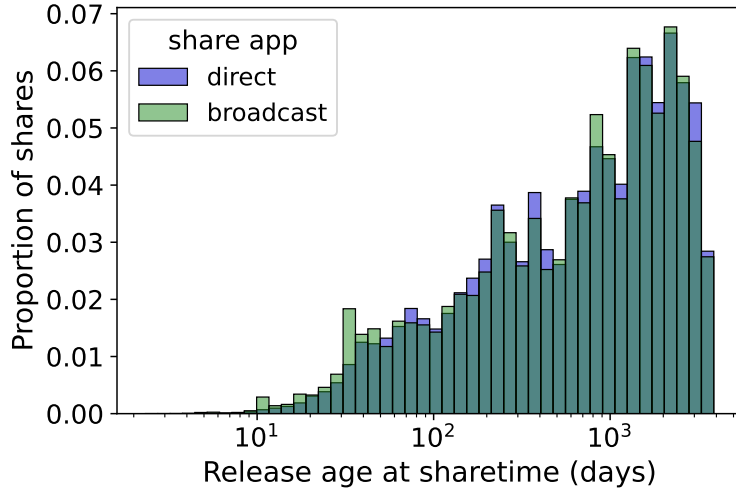


Figure 5.2: **Distribution of the album release age at share-time for the link sharing events.** Distribution separated by broadcast and direct application types.

simultaneous recipients. Figure 5.2 gives the distribution of album release age at share-time for our link sharing event sample, separated by direct and broadcast integration. The figure indicates that older tracks are more likely to be shared; this is a consequence of the increased prevalence of old content on the platform compared to new.

5.2.3 Music Vector Embeddings

In the standard literature on diffusion processes over social networks, each individual in the network is typically considered identical, exhibiting equal susceptibility to “contagion” following a contact event [23]. In practice, this assumption rarely holds; individuals have varying levels of susceptibility to diverse influences, ranging from infectious diseases, to opinion shifts, to adopting novel behaviours. In this work, we leverage a rich dataset which allows us to move past this assumption. Using each user’s prior listening history on the Spotify platform, we examine how music spreading relates to taste similarity and homophily, the tendency for individuals to connect with others who are similar to themselves. Specifically, we identify factors that render an individual more or less susceptible to adopting a new artist following a contact event.

To this end, we employ the vector embeddings of Spotify users and content in latent space, which are pre-trained and available internally at Spotify. These vector embeddings are re-computed daily using the word2vec algorithm [147] trained on user-generated playlists, with the playlists as the “documents” and the tracks as the “words”. This process produces for each track an 80-dimensional embedding vector, where tracks that are often co-listed in playlists appear in high proximity in the embedding space. Users frequently build genre-themed playlists, such that similar songs will exhibit high co-occurrence across playlists [10]. These user generated playlists capture an aspect of

intentionality that is lost if actual listening sequences are used, for example, where subsequent song choices are often selected by the recommender system.

Users are embedded in the same latent space, their vectors are generated by averaging the vectors of all the tracks they have listened to during a certain time window. For details on the vector embedding generation process, see [10]. The vectors can be used to compute the music taste similarity between two users, or between a user and track, by taking the cosine similarity between the vectors. We note that there are other possible methods to capture music taste similarity, but we use this method because it the standard used in Spotify’s internal systems, has been discussed widely across the literature and the information captured by the embedding space methodology has proven to be a good representation of the way that users engage with music [148, 10]. For each user and track in our sample, we take the corresponding vector computed at the beginning of our analysis period, to ensure consistency and ensure the vectors have only been influenced by listening history prior to the share events.

5.2.4 User-Artist Engagement

To identify a successful share, we quantify the receiver’s engagement with an artist following a share event, by defining a metric called “user-artist engagement”.

We denote the number of unique tracks by the artist α that the user i streams for longer than 30 seconds on a particular day t as $n_{i,\alpha}(t)$. The user-artist engagement is defined as $e_{i,\alpha}(t) = \log_{10}(n_{i,\alpha}(t)) + 1$, if $n_{i,\alpha}(t) \geq 1$, and 0 otherwise.

User-artist engagement captures the intensity of a user’s engagement with the artist each day, which can be aggregated over time as:

$$E_{i,\alpha}(t_0, K) = \sum_{t=t_0}^{t_0+K} e_{i,\alpha}(t), \quad (5.1)$$

where t_0 is the day when the link share is opened, and K is the number of days over which we aggregate, which can be either positive or negative indicating aggregation over time before or after t_0 respectively.

Several alternative metrics could be chosen to measure a successful share. For example, the total number of streams over a window of time gives a broad indication of interest; the number of unique tracks listened to over a window of time gives an indication of deep interest; or the number of days listened to an artist gives a longitudinal metric, indicative of loyalty. Our choice of E in Eq. 5.1 balances all three by looking at the unique number of daily tracks listened over a time window. The deeper (more daily tracks) and extended (more days) the listening behaviour, the higher the user-artist engagement factor. The advantage of using this metric, instead of simply counting the number of unique tracks that the user listens to over the same period of time, is that $E_{i,\alpha}(t_0, K)$ is higher if the number of listening events are dispatched over a larger number of days. To illustrate,

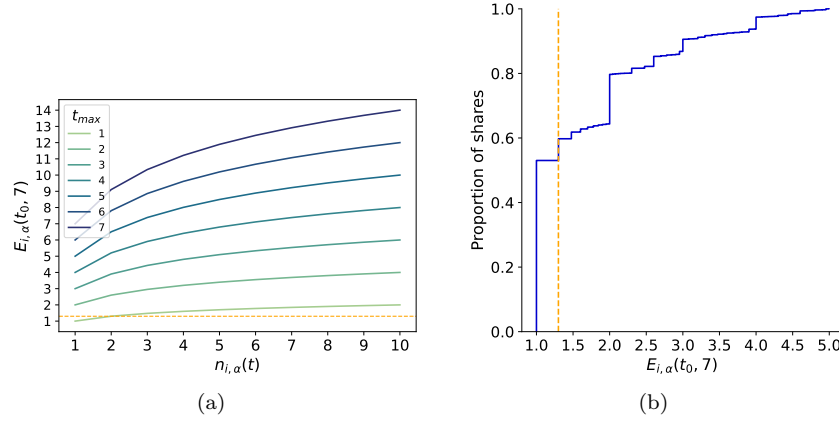


Figure 5.3: **User-artist engagement threshold choice.** (a) User-artist engagement curves as a function of the number of tracks $n_{i,\alpha}(t)$ for different number of days listened $t_{\max}$ within a 7 day period. (b) Cumulative distribution function (ecdf) of the link shares and the associated receiver-artist engagement 7 days post share ($E_{i,\alpha}(t_0, 7)$). The orange line indicates our chosen threshold for a successful share event.

Fig. 5.3 (a) presents the user-artist engagement E as a function of the number of daily unique tracks listened $n_{i,\alpha}(t)$, for various total number of days listened $t_{\max}$ within a seven day period. The different user-artist engagement curves define the metric for different number of tracks, and engagement days, demonstrating that the metric E rewards high and repeat engagement.

This behaviour is desirable when describing the propagation of a social process which tends to result in the development of a habitual relationship, such as becoming a fan of an artist. Thus, we capture the delineation between a state of longer-term engagement, and an intense, but time-limited, fixation with the artist. We use this metric to characterize both the receiver’s engagement with the artist following the share, and the engagement of the sender with the shared artist, prior to the share.

A successful share is defined as the binary outcome, by thresholding the receiver’s user-artist engagement with the shared artist in the seven-days following the event. Figure 5.3 (b) gives the cumulative distribution function of $E_{i,\alpha}(t_0, 7)$, the receiver-artist post-share engagement. We observe that 52.8% of shares yield a seven-day engagement factor of 1.0, indicating that the user only listened to a track by the shared artist once in the seven days following the share.

Consequently, we classify a successfully “activated” receiver as one who has listened to two or more tracks by the artist in the seven-days following the share, a threshold which is equivalent to $E_{i,\alpha}(t_0, 7) > 1.3$, since $\log_{10} 2 + 1 = 1.3$. This threshold is shown in Fig. 5.3 (b) by the dotted orange line.

This choice is grounded in the behavioural interpretation of engagement; intuitively, exploring an artist’s discography or returning to the artist’s music after the initial exposure signals that the user is engaging further with the artist and that the recommendation was indeed successful. This choice of threshold therefore identifies receivers who are beginning to form an attachment with

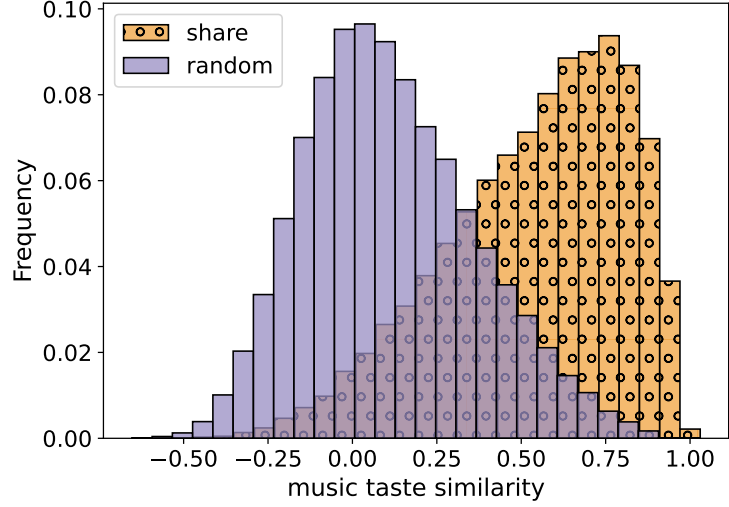


Figure 5.4: **Measuring homophily in the link sharing events.** Distributions of music taste similarity for pairs of users who shared music in our sample (share) and a random permutation of those pairs (random).

the artist in the short-term and who may become potential fans. With this threshold, 47.2% of the link shares in our sample result in successful engagement, representing a fairly balanced split between activated and non-activated users. As previously discussed, other choices could be adopted to define a successful share, however, here we are interested in identifying the mechanisms that cause a behaviour change on the short-term timescale, directly following a contact event. Moreover, as discussed, users listening sequences represent a mix of both organic and programmed content. The user profile vectors, defined by aggregating over their most recent song embeddings, are re-computed frequently and are fed into the platform’s recommender system to suggest *programmed* content [10]. In order to ensure we are capturing an intentional re-engagement with the discovered artist, we chose a short timescale, making the assumption that user engagement within a short seven-day window is more likely to reflect organic and intentional behaviour, rather than passive exposure to algorithmically suggested content.

5.3 Results and Discussion

We evaluate the three key propositions presented in the Introduction by examining the probability of engagement $p(\text{engaged})$, i.e. the proportion of link share events where the receiver’s engagement is $E_{i,\alpha}(t_0, 7) > 1.3$, with respect to: firstly, the music taste similarity between the two interacting users, and the shared track, secondly, the nature of the social relationship between the users, and lastly social cohesion, the penetration of that artist within the receiver’s social network neighbourhood. We then construct a model to predict successful receiver-artist engagement following a share event, using a broader set of social and contextual features.

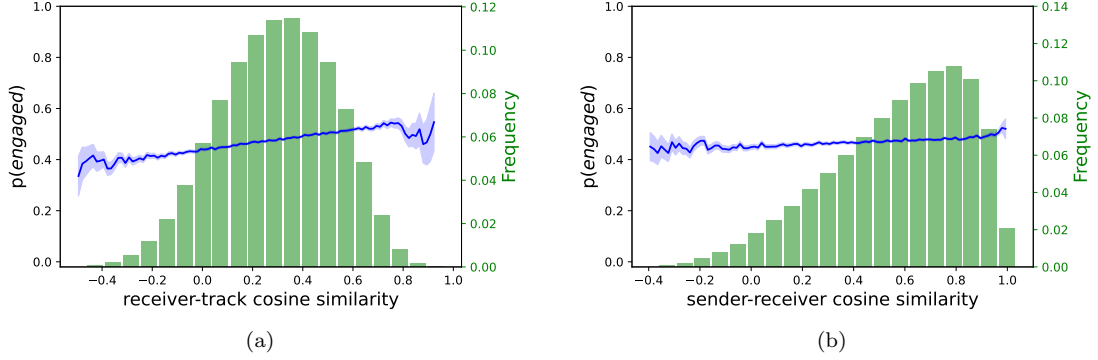


Figure 5.5: **Analysis of music taste similarity.** The blue line shows the probability of becoming engaged with an artist as a function of the feature, where the shaded area is the 95% confidence interval (main y-axis). The histograms give the distribution of the feature in our sample in the second y-axis (Frequency).

5.3.1 Music Taste Similarity

Social networks often exhibit homophily, the tendency for social ties to form between individuals who share similar characteristics [145]. Consequently, it is often difficult to determine whether correlated patterns of behaviours between connected individuals can be attributed to a causal influence (social contagion) or arise as a result of the pre-existing homophily in the network. Several works have proposed methods to delineate the effects of the two phenomena, though the problem remains challenging, particularly in the absence of longitudinal temporal data [202, 13].

We quantify the similarity in music preference by taking the cosine similarity of the embedding vectors described in the preceding section. To obtain an estimate of the level of homophily in our sample of link-sharing users, regardless of whether the shares were successful, we design the following process. We generate a null model by reshuffling the users in our sample, creating random permutations of sender and receiver pairs, and measuring the cosine similarity between their taste vectors. Figure 5.4 compares the distribution of the user-user cosine similarities between the randomly generated pairs, and between the pairs of users who shared and opened links in reality.

The figure illustrates that the distribution of user pair cosine similarities for users involved in link sharing is skewed towards a higher similarity with a mean of 0.54, compared to the random user-user pairs, for which the distribution has a mean of 0.11. The difference in distributions between the two groups is statistically significant as confirmed by the Kolmogorov-Smirnov test ($D = 0.44, p < 0.01$). This gives some indication of the level of homophily in the Spotify social network. This pattern may arise because users preferentially form friendships with individuals who share similar music tastes, because they chose to share content with their friends who have similar tastes, or because the music taste of connected users begins to converge over time.

As a consequence, it is difficult to determine whether the engagement we observe following a share is due to homophily or to social contagion. However, while we may not be able to disentangle the two mechanisms at play, we can still investigate the social and contextual factors that lead to

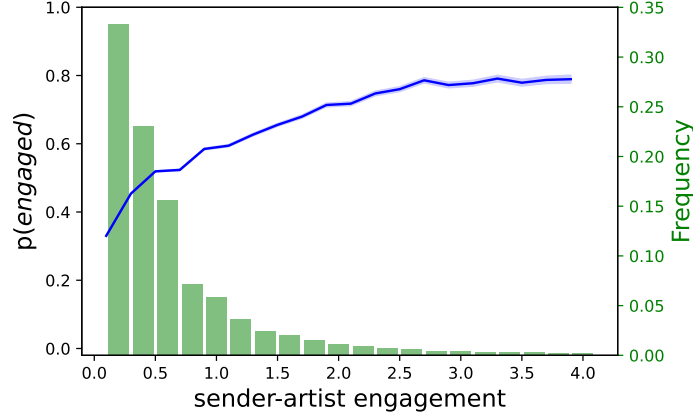


Figure 5.6: **Analysis of Sender-artist engagement aggregated 7 days prior to share.** The blue line gives the probability of receiver engagement (main y-axis) as a function of the feature, where the shaded area is the 95% confidence interval. The histogram (second y-axis in green) gives the distribution of the feature in our sample.

successful engagement following a share event. We first test the hypothesis that the probability of engagement will increase when users have similar taste in music (sender-receiver cosine similarity), or when the shared track is similar to the listening history of the receiver (receiver-track cosine similarity). The results are given in Fig. 5.5 (a), showing the distribution of receiver-track cosine similarity in our sample, and $p(\text{engaged})$ as a function of the similarity. Similarly, Figure 5.5 (b) shows the results for sender-receiver cosine similarities. We observe that while this hypothesis does hold, the receiver-track overlap presents a stronger effect than the sender-receiver overlap; $p(\text{engaged})$ increases linearly with cosine similarity.

Sender-Artist Engagement

Next, we investigate the effect of the sender-artist engagement with the artist prior to the share on $p(\text{engaged})$. We measure this by taking the aggregate user-artist engagement (described in the previous section) for the seven-days prior to the share. This signal indicates how engaged the sender was with the shared artist, prior to making contact with the receiver. This metric corresponds to Equation 5.1, where i is the sender, t_0 is the day of the share, and $K = -7$. The feature itself could be a proxy for the strength or intensity of the sender’s recommendation. Users have the ability to add messages when sharing a Spotify link, and while we do not have access to this textual data, it is possible that the associated message might be more emphatic when the sender is more engaged with the artist. Note that the sender-artist engagement can be 0 if the sender sends a track to someone without listening to it first, as these are separate functions on the Spotify interface. The results are shown in Fig. 5.6, which illustrates that the probability of a successful activation does increase with increased sender-artist engagement.

5.3.2 Tie Strength Effects

Following [7] and [34], behaviours and trends propagate more effectively over stronger social ties. In studies of spreading processes on networks with uniform edges, it is inherently difficult to quantify this effect. In this study, because the underlying social network is heterogeneous and characterized by diverse ties, we have the unique ability to investigate our second proposition; that the probability of engagement $p(\text{engaged})$ will be higher when the contact event takes place between users with stronger and more intimate relationships.

We measure the tie strength between users in three different ways; reciprocity, the mode of social media application used to share the track, and the sum of social interactions. Reciprocity is a binary variable indicating whether the receiver (i) has ever shared a link with the sender (j), prior to the current share time. Formally, a link is reciprocal if $L_{l,ij} > 0$. Figure 5.7 (a) presents the distribution of sender-receiver cosine similarities over all share events, separated by reciprocal and non-reciprocal ties. The mean sender-receiver cosine similarity between users connected via reciprocal links is 0.58, compared to 0.5 for non-reciprocal links. The difference in distributions is statistically significant as confirmed by the Kolmogorov-Smirnov test ($D = 0.12, p < 0.01$), indicating that users connected by reciprocal ties tend to have a greater taste similarity than those who do not. Figure 5.7 (b) shows the probability of a successful share event, $p(\text{engaged})$, along reciprocal and non-reciprocal links, holding constant the receiver-track cosine similarity. Even controlling for the taste of the receiver, the probability of future engagement is more likely to occur along a reciprocal tie. Variability is observed at very large and very small receiver-track cosine similarities, where there is limited data available.

The second measure we consider is the mode of social media application used to share the link. As previously outlined, we have delineated the applications into two categories, *direct* (representing one-to-one communication) and *broadcast* (representing one-to-many communication). Although not strictly a measure of tie strength in the traditional sense, this choice is motivated by the intuition that one-to-one channels (like Whatsapp or Facebook Messenger) are more intentional, intimate or occur between closer individuals, as compared to one-to-many channels (e.g. posting an Instagram story) which will reach a broad untargeted audience. For this analysis, we restrict the data to include only events for which the application mode is unambiguously identified, retaining only known direct (66%) or broadcast (5%) share events.

Figure 5.7 (c) presents the distribution of sender-receiver cosine similarities over the subsampled share events, separated by direct or broadcast events. Among direct-mode shares, the mean user-user cosine similarity is 0.54, compared to the 0.57 for broadcast shares. The difference in distributions between the two groups is statistically significant as confirmed by the Kolmogorov-Smirnov test ($D = 0.03, p < 0.01$), indicating that users selectively engage with music shared by friends with

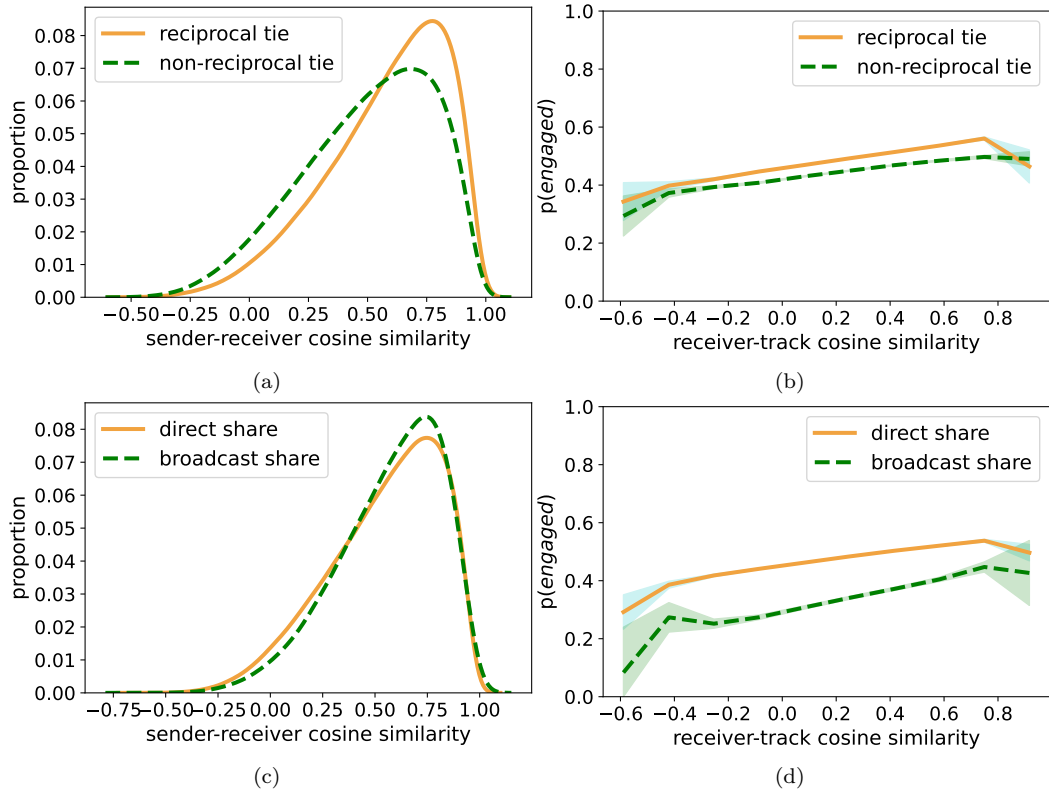


Figure 5.7: **Effects of reciprocity (a-b) and share media type (c-d) on the probability of becoming engaged with the artist.** (a) Distribution of sender-receiver cosine similarities in our sample, separated by reciprocal and non-reciprocal tie. (b) Probability of engagement following link share as a function of receiver-track cosine similarity, separated by reciprocal and non-reciprocal tie. (c) Distribution of sender-receiver cosine similarities in our sample, separated by direct or broadcast application. (d) Probability of engagement following link share as a function of receiver-track cosine similarity, separated by direct or broadcast application.

similar taste on social media, while broadcast platforms may afford greater opportunities to explore less familiar or diverse musical content.

Figure 5.7 (d) shows the probability of engagement for share events via direct versus broadcast mediums, conditional on the receiver-track cosine similarity. Even while controlling for the taste of the receiver, the probability of a successful share is more likely to occur when the share is delivered via a direct medium, even when the track does not reflect the receiver’s historical listening patterns. Note that the probability of engagement approaches zero for broadcast shares at large receiver-track cosine similarities.

The final measure of tie strength, the sum of social interactions, represents the number of times that the sender (j) and receiver (i) have interacted (on Spotify) prior to the focal share event. This includes social listening sessions, collaborative playlists, and link shares (from the sender to receiver or from the receiver to the sender). The sum of social interaction metric is equivalent to $L_{s;ij} + L_{c;ij} + L_{l;ij} + L_{l;ji}$, where we have included the link shares in both directions. All interactions are required to have occurred prior to the share event under analysis.

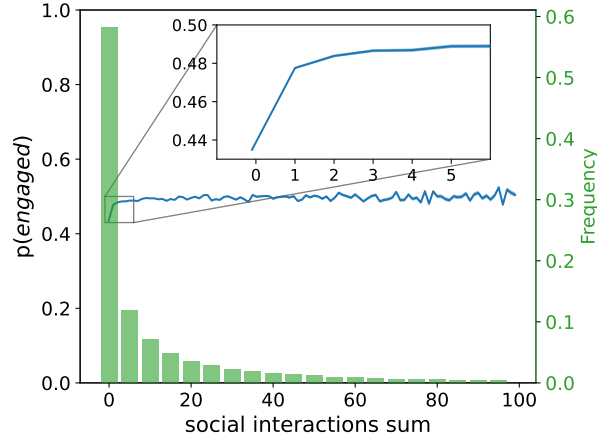


Figure 5.8: **Analysis of sum of social interactions.** The histogram gives the distribution of the feature in our sample. The blue line gives the probability of engagement as a function of feature, where the shaded area is the 95% confidence interval. The inset gives a closer look at the probability of engagement at the lower end of total interactions.

This measure quantifies whether the tie is newly formed or well-established. Figure 5.8 presents both the distribution of the sum of social interactions over the pairs of users in our sample, as well as the probability of engagement along a link as a function of the sum of social interaction metric. The probability of engagement increases from 0.44 to 0.48 after just one previous interaction and rises gradually thereafter, reaching 0.5 at five interactions. Beyond this point, the probability of engagement appears to plateau; this indicates that while engagement is more likely to happen along an established tie, there is a “saturation point” after which more prior interactions do not increase the probability.

Crucially, while we develop these three measures of tie strength to leverage the heterogeneity of the underlying social network, we do not assume that they are mutually independent. In particular, both reciprocal ties and direct shares are more likely to occur between users with established history of interacting on the Spotify platform. Shares made via a direct medium occur between users with an average of 24 social interaction sum, compared to an average of 6 for shares made via broadcast channels. Similarly, users connected by reciprocal ties have an average of 42 social interaction sum, compared to 3 for those with non-reciprocal ties. Nevertheless, we retain all three features in the analysis and predictive model as they represent different aspects and granularity of tie strength, and indeed show different relationships with the probability of post-share engagement, underscoring the advantages of preserving information on edge heterogeneity when examining social diffusion processes.

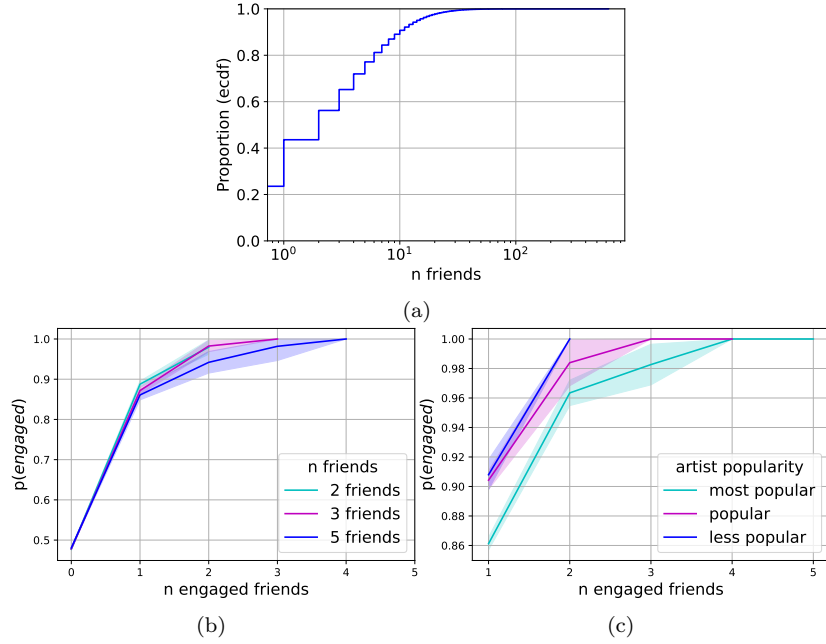


Figure 5.9: **Analysis of social cohesion features.** (a) Distribution of number of friends for all receiving users in our sample. (b) Probability of becoming engaged versus number of friends who are already engaged with the artists, for receivers with different numbers of friends. (c) Probability of engagement versus number of friends who are already engaged, stratified by artist popularity.

5.3.3 Social Cohesion

Finally, we investigate the effects of social cohesion, addressing our third hypothesis. These results elucidate how the level of engagement with the shared artist within the receiver’s neighbourhood of contacts influences the probability of a successful share. We define the receiver’s contacts or *friends* as anyone with whom they have a reciprocal relationship on the platform. These relationships can arise from social listening sessions, collaborative playlists, or bi-directional link sharing. More specifically, the receiver (i) is friends with a user (u) if $L_{s;iu} > 0$, or $L_{c;iu} > 0$, or $L_{l;iu} > 0$ and $L_{l;ui} > 0$. Figure 5.9 (a) shows the distribution of number of friends over receivers in the sample. The mean number of friends is 3.95, but the distribution is skewed to the left; most users have zero or one friend.

A receiver’s friend is considered to be engaged with the shared artist if their user-artist engagement factor indicates listening to an average of at least one track per day by the shared artist for the six months prior to the analysis period. This is equivalent to the condition that $E_{u,\alpha}(t_0, -180) > 180$, where u is the friend user, α is the shared artist and t_0 is the start of the analysis period. We expect to observe the phenomenon we call *social cohesion*, that the probability of engagement following a share will increase with the number of the receiver’s previously engaged friends, arising either from social influence, or homophily in the social network.

We group all receivers who have the same total number of friends, and measure the probability

of engagement as a function of the number of engaged friends, thereby holding constant both the absolute number of friends and the fraction who are engaged. The results are displayed in Fig. 5.9 (b) for receivers with 2,3 and 5 friends. The shaded area gives the 95% confidence interval. In all cases, the probability of engagement increases as the number of engaged friends increases, consistent with the hypothesis that social cohesion increases the probability of engagement.

One possible confounding factor for the observed effect is the artist popularity. In particular, the fraction of listener friends who are engaged with an artist may simply be a proxy for the artist’s general popularity, as these artists have more engagement overall. To disentangle the two effects, we proceed by classifying the artists according to their popularity on Spotify, and compute the probability of engagement as a function of number of engaged friends. The results are shown in Fig. 5.9 (c). The probability of engagement increases with the number of engaged friends, even when accounting for artist popularity. Moreover, we notice that the effect of *social cohesion* is indeed more pronounced in the case of less popular artists, perhaps indicating that social effects play a more significant role in music discovery for lesser known artists.

As discussed in the Introduction, these measures are reminiscent of those used to investigate complex contagion in the spread of social behaviours. Complex contagion as defined in [45], models how behaviours spread over social networks while incorporating social reinforcement; the probability of adoption increases with the number of previously-adopted contacts. Evidence supporting this model has been found in empirical data describing the spread of campaign donations [199], exercising [14], and marketing recommendations [130]. Our dataset does not contain evidence of multiple contacts; users are rarely sent a link to the same artist by multiple people. As in [202], it can be difficult to disentangle the evidence of complex contagion and homophily, as both mechanisms produce similar patterns. However, our results on the relationship between social cohesion and engagement probability suggest that these social mechanisms contribute to explaining our observations, even if the actual contact must be inferred rather than directly observed.

These mechanisms are typically related to clustering in the network. Consequently, we also examined the relationship between engagement probability and two additional measures; the receiver’s neighbourhood clustering coefficient, and the edge overlap between the pair of users. The receiver’s clustering coefficient is the proportion of friends who are also friends with each other. Higher clustering has been shown to facilitate the spread of complex contagion [44]. The edge overlap as defined in [142] is the proportion of friends shared by the sender and receiver. These measures were calculated by taking the unweighted version of the social network. Neither of these measures proved to have a significant relationship with engagement probability. This may be because the Spotify social network is merely a partial representation of the true underlying social network, and the metrics calculated may not accurately reflect the complete set of social relationships, which are unavailable to us.

5.3.4 Predicting Post-share Engagement

In this section, we integrate the insights from the three hypotheses, together with other contextual factors, to predict a successful engagement following a link share. To that end, we formulate a supervised binary classification task, using the data available at share-time to predict engagement in the week following the share. The purpose of this task is two-fold; we are interested in both training a model capable of predicting a successful person-to-person music recommendation, and in further elucidating *which* social and contextual factors have the greatest influence on the outcome.

Features

In Table 5.1 we describe the set of features used to inform the binary prediction task. They are separated into groups according to the hypothesis they most align with.

The Taste Similarity (TS) features describe the music taste similarity between the sender and receiver, and between the receiver and the shared track. These capture distinct effects; the sender-receiver cosine similarity measures the overlap between the listening histories of the interacting users. The receiver-track cosine similarity captures the level of targeting; how accurately the sender anticipates the receiver’s musical taste. This feature set is informed by our results in the Music Taste Similarity section. The sender-artist engagement (SA) describes the sender’s level of engagement with the artist prior to the share, similarly explored in the Music Taste Similarity section.

Social Tie Strength (ST) features capture the strength of the relationship between the sender and receiver, quantified by the number of times they have previously interacted on the platform, the relationship reciprocity, and the mode of social media application by which the link was sent. This set of features is informed by our results in the Tie Strength Effects section.

The Social Cohesion (SC) features describe the popularity of the artist amongst the receiver’s social network connections, and allows us to capture the social effects of engagement probability. This feature is informed by the results provided in the Social Cohesion section.

Finally, we include additional features that describe both the network context, and the historic context of the artist, track, and user’s platform usage. The Social Network Degrees (SN) features characterize the role of the sender and receiver in the Spotify social network in terms of their degrees. The Track Context (TC) features provide the track album age and popularity of the artist on Spotify. This allows us to control for the general popularity of the shared track, which may reflect its inherent quality, causing a higher probability of engagement. The last feature group describes the receiving user’s activity levels on Spotify. The aptly named Receiver Platform Usage (PU) measures are included because the engagement threshold is fundamentally dependent on the receiver’s Spotify usage. For infrequent users, it may be impossible to meet the engagement threshold, even if they appreciated the shared track. These features provide information about the receiver, including whether they use the paid subscriber version of Spotify (as opposed to the free version), how much

music they streamed in the 7 days prior to the share, and the number of days that they spent on the platform prior to the share-time.

Model Training

We train a Random Forest binary classification model using the scikit-learn python framework [162] on the full set of features and data (~ 2.72 M link sharing events) using a randomised hyper-parameter search with 5-fold cross-validation and 500 fits varying the parameters of the model. The best model performance is obtained with 300 estimators, and a maximum tree depth of 60 resulting in a ROC-AUC of 0.73 (see Table 5.2 Full Model).

Feature Importance

From the model we directly obtain a measure of feature importance using Mean Decrease in Impurity (MDI), as provided in Fig. 5.10. The Taste Similarity features have a significant impact on the model performance, as do the receiver-platform engagement, sender-artist engagement, and various social features from the social network.

As in Fig. 5.10, several of the feature sets ranked surprisingly low in importance for this prediction task, despite demonstrating a relationship with the outcome variable in the preceding independent analysis. In particular, the results in the Tie Strength and Social Cohesion sections indicated that the number of engaged friends, social media application type and reciprocal link sharing were important predictors of engagement, but these features do not emerge as important in the context of the Random Forest classifier.

This hints at the fact that the information encoded in these features can be learned by the model during training using a combination of other features. As a trivial example, the model can infer that the number of engaged friends is greater than zero using the sharer-artist engagement for the week prior to the share, and the number of previous shares between the receiver and the sender to determine if the relationship is reciprocal.

Moreover, feature importance measures are significantly impacted by feature correlation. Unsurprisingly, the taste similarity features are highly correlated with a 0.42 Pearson correlation ($p < 0.01$). Several of the social network features are also correlated; for example, the sum of social interactions and the receiver in-degree have a Pearson correlation of 0.48 ($p < 0.01$). Intuitively, the sum of social interactions is inherently related to the amount that the receiver shares in general. Likewise, the sender out-degree and the sum of social interactions have a correlation of 0.32 ($p < 0.01$). This can explain why the MDI scores for some features are lower than expected, as the model can take one of the correlated features as redundant. To get a complete picture of the importance of various feature sets, we perform an isolation test, as follows.

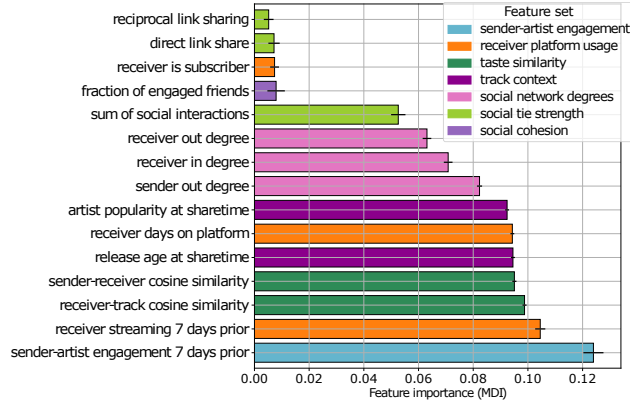


Figure 5.10: **Feature importances for the full model.** Importance measured using Mean Decrease in Impurity (MDI). Bars are coloured according to their feature set membership.

Feature Set Isolation Test

We formulate a procedure to determine the relative importance of the various feature sets defined in Table 5.1. Instead of conducting a more traditional ablation test (removing one feature from the model at a time), which reveals little new information due to the variables being correlated, we conduct a feature set isolation test. We train the several iterations of the Random Forest model (with the same hyper-parameters described above) with different isolated groups of features, and compare the model performances to explore how predictive strength varies. We report the ROC-AUC, Precision, Recall and Average Precision. Since our dataset is not perfectly balanced, we use ROC-AUC score as an indicator of overall model performance. The results are summarised in Table 5.2, where the best performance metrics are highlighted in bold. All standard errors of the metrics are within 0.0007, indicating stability across the 5-folds.

The sender-artist engagement (SA) and the receiver platform usage (PU) feature sets both perform well alone, obtaining the highest ROC-AUC. As previously discussed, the sender-artist engagement could be a proxy for the strength of the sender’s recommendation via a text-based addition to the link-share, information which was unavailable for this study.

The Taste Similarity features (TS) also perform reasonably well alone, as might be expected. However, they do not constitute the highest-performing feature set; this is likely due to the specific context of the share. The embedding vectors capture a notion of music preference, based on aggregated information over many contexts, which may have a smoothing effect on users with diverse taste profiles.

Note that the Social Cohesion feature (fraction of engaged friends) produces the lowest ROC-AUC, but the highest precision. This feature is incredibly sparse, only available in approximately 1% of the cases. For this reason, it cannot have high recall thus leading to lower overall performance. However, when the feature is available, it has the highest precision out of the feature sets, an effect we also observed in the single-variable analysis provided in the section on Social Cohesion. Despite

some sets of features performing better in isolation, all sets of features perform better than a random baseline (ROC-AUC = 0.5) and indeed, best model performance is achieved by combining all feature sets.

5.4 Conclusions

In this chapter, we present an empirical analysis of how music taste diffuses over Spotify’s social network through person-to-person recommendations. We leverage the advantages of the diverse ties that comprise the network to better understand the mechanisms that facilitate the spreading process, incorporating both the edge weights and the multiplex nature of the network.

We examine various social and contextual effects that contribute to a successful share event, using information about the users’ context within the Spotify social network, their listening behaviours and their relationship history. We have shown that the probability of a successful share is higher when the share occurs between (1) between users with similar taste, when (2) the share is highly targeted to the receiver’s tastes, and (3) when the sender is more engaged with the artist they shared. We use the heterogeneity of the underlying social network to show that the process spreads more effectively along stronger and reciprocal ties, and is facilitated by intimate one-to-one modes of contact. Finally, following much of the prior literature on social influence, we explored how social cohesion affects the probability of engagement, demonstrating that users with a higher fraction of engaged friends will be more likely to become engaged with the shared artist. Using the identified social and contextual features, we construct a model to predict engagement following a person-to-person music recommendation.

We recognize that while the model archives strong performance, especially within the context of predicting social behaviour, there may indeed be other important factors we are unable to capture with the data available to us, such as the type of relationship between the sender and receiver, or the context in which the link was shared.

As we focus here on short-term engagement resulting from a link share, one future research direction would be to investigate the lifetime of such engagement, and whether sharing can lead to more permanent behaviour changes, resulting in the receiver becoming a long-term listener, or even a fan, of the shared artist. Since we have focused on the dynamics of pairwise interactions in the Spotify social network, other possible future directions include the study of higher-order group interactions, especially facilitated through group social listening sessions or playlists constructed by larger groups.

This work contributes to a wider body of literature on how behaviours, tastes and opinions diffuse over a social network, especially those with diverse interaction types. Using explicit evidence of contact events together with rich contextual and social information, we have shed light on the dynamics of social music discovery in the age of online music streaming.

Feature	Description
Social Tie Strength (ST)	
Sum of social interactions	Sum of the number of social interactions between the sender (j) and receiver (i) prior to share-time. Includes collaborative playlist additions, social listening sessions, and link shares between the two users in either direction. Sum of social interactions = $L_{s;ij} + L_{c;ij} + L_{l;ij} + L_{l;ji}$.
Direct link share	Boolean, True if the social media application type used to share the link was direct (e.g. Whatsapp), as opposed to broadcast or unknown.
Reciprocal link sharing	Boolean, True if receiver has ever shared a link with the sender prior to share. This is equivalent to $L_{l;ij} > 0$.
Social Network Degrees (SN)	
Receiver link share in-degree	Number of link shares received by the receiver (i) (from anyone) prior to share-time. This is equivalent to $\sum_k L_{l;ki}$
Receiver link share out-degree	Number of link shares sent by receiver (i) prior to share-time (to anyone). This is equivalent to $\sum_k L_{l;ik}$.
Sender link share out-degree	Number of link shares sent by the sender (j) before the share-time (to anyone). This is equivalent to $\sum_k L_{l;jk}$.
Social Cohesion (SC)	
Fraction of engaged friends	“Friend” is defined as someone with whom the receiver has a reciprocal relationship in the social network prior to share-time. A friend is considered engaged if $E_{u,\alpha}(t_0, -180) > 180$, where u is the friend user, α is the shared artist and t_0 is the start of the analysis period.
Taste Similarity (TS)	
sender-receiver cosine similarity	Cosine similarity between receiver and sender.
receiver-track cosine similarity	Cosine similarity between receiver and track.
Track Context (TC)	
Artist popularity at share-time	Rank of the primary artist in Spotify by number of streams at the time of share.
Release age at share-time	Time elapsed between album release time and share-time (seconds).
Sender-Artist Engagement (SA)	
Sender-artist engagement 7 days prior	Sender-artist engagement aggregated over 7 days prior to share-time.
Receiver Platform Usage (PU)	
is subscriber	Boolean, True if the user was a paid subscriber to Spotify when the share was opened.
receiver streaming 7 days prior	Number of hours of streamed music in the previous 7 days to opening the share.
receiver day on platform	Number of days passed since the user registered to the platform to the day the share was opened.

Table 5.1: **Model features and description.**

Feature Set	ROC-AUC	Precision	Recall	Avg Precision
Social Cohesion (SC)	0.5144 $\pm$ 0.0001	0.8779 $\pm$ 0.0004	0.0329 $\pm$ 0.0003	0.4868 $\pm$ 0.0003
Social Tie Strength (ST)	0.5327 $\pm$ 0.0003	0.4997 $\pm$ 0.0002	0.1561 $\pm$ 0.0001	0.4933 $\pm$ 0.0003
Social Network Degrees (SN)	0.5469 $\pm$ 0.0002	0.5181 $\pm$ 0.0004	0.4091 $\pm$ 0.0005	0.5235 $\pm$ 0.0003
Track Context (TC)	0.5731 $\pm$ 0.0001	0.5356 $\pm$ 0.0003	0.5497 $\pm$ 0.0002	0.5365 $\pm$ 0.0003
Taste Similarity (TS)	0.5847 $\pm$ 0.0004	0.5346 $\pm$ 0.0005	0.5088 $\pm$ 0.0004	0.5935 $\pm$ 0.0005
Receiver Platform Usage (PU)	0.6102 $\pm$ 0.0005	0.5536 $\pm$ 0.0006	0.5573 $\pm$ 0.0004	0.6018 $\pm$ 0.0007
Sender-Artist Engagement (SA)	0.6954 $\pm$ 0.0001	0.6487 $\pm$ 0.0005	0.5737 $\pm$ 0.0003	0.6634 $\pm$ 0.0003
Full Model	0.7336 $\pm$ 0.0003	0.6734 $\pm$ 0.0004	0.5874 $\pm$ 0.0003	0.7498 $\pm$ 0.0004

Table 5.2: **Feature Set Isolation Test Results.** All metrics are reported with their standard error.

Chapter 6

Detecting ideological extremes in social media networks

This chapter is based on the article *Candellone, Elena*, Shazia'Ayn Babul*, Özgür Togay, Alexandre Bovet, and Javier Garcia-Bernardo. "Negative ties highlight hidden extremes in social media polarization." Network Science 13 (2025): e10. (* indicates joint first authors) [39]*

6.1 Introduction

This thesis is primarily concerned with adapting traditional network science methods to admit heterogeneous edges, arguing that retaining edge type diversity provides greater insight into the structure of the network, or the processes that unfold on top of it. We have explored two major types of heterogeneity; tie sentiment, presenting methods for extracting latent structure in signed networks in Chapter 3 and Chapter 4, and tie strength or modality, which prove crucial for describing social diffusion dynamics, as we demonstrate in Chapter 5. In this final chapter, we weave together these learnings, presenting an empirical study on an online social network for political discussion, characterized by its signed, weighted and multiplex nature. Using two network embedding methods, including the SHEEP algorithm from Chapter 3, we investigate how the interplay of these heterogeneities can uncover hidden individual-level ideological positioning and extremity, patterns that would remain obscured if the network were reduced to uniform or unsigned edges.

Online social networks are a natural context for this work, as interactions are inherently heterogeneous, and extracting the latent structure of the network is especially crucial, as these platforms have changed how individuals interact with news content, inform themselves on current events, and form opinions on related topics [137]. When platform discussions take on a political nature, they provide a new mechanism for information diffusion, which can facilitate an increased polarization around controversial issues, and amplify political divisions and conflicts [83, 6, 52, 104, 25, 72, 76, 197]. Consequently, political polarization is increasing not only within small, active partisan groups [151] but also more broadly across the general population [151, 4, 3].

In Chapter 3 and Chapter 4, we used the term *polarization* following traditional signed network literature, to refer to networks that can be partitioned into communities with positive edges within, and negative edges between communities. In the context of this politically-motivated empirical study, however, we draw from political science literature, in which *polarization* is more nuanced, delineating between ideological, affective, and structural (or interactional) polarization [104, 63, 35, 128, 26, 151, 70, 6, 181, 207].

Ideological polarization refers to the widening gap between the political beliefs of different groups; affective polarization captures the emotional hostility and negative attitudes between political factions; and structural polarization examines the division of social interactions into homogeneous groups with minimal cross-group engagement.

With the increasing availability of data on online interactions, structural polarization has emerged as a rich area of study [207, 181, 73]. These prior works have shown that social media platforms can amplify polarization, reinforcing group differences by facilitating the formation of echo chambers (clusters of like-minded users) and the rapid diffusion of biased or false information [57, 50, 25, 74].

Ideological polarization can also be recovered from the network of interactions between social media users, under the assumption that users preferentially interact with other users (or content) with whom they share a similar ideological stance [26, 174]. The *ideal points model*, used to measure both voting behaviours [69, 167, 166, 51, 208] and ideology [150, 205, 24], is predicated on the same assumption: the probability of voting (positively or negatively) on a piece of content depends on the latent ideological difference between the individual and the content. The model has been applied across various contexts, from a legislator voting on a piece of legislation, to a user voting on a social media post.

The methodologies typically employed to study structural or ideological polarization are consistent with traditional unsigned networks comprising uniform edge weights, which consider all interactions equivalent. This is primarily because such studies remain constrained by limited data, as social media platforms readily provide information on positive interactions (e.g., likes or retweets on Twitter/X) or neutral interactions (e.g., mentions on Twitter/X), while negative interactions (e.g., downvotes on Reddit) are either unrecorded or only accessible to researchers at an aggregated level. Thus, prior studies construct network representations that make no distinction between interactions that have positive or negative valence.

In this work, however, we consider such distinctions crucial; drawing from *Emotional Information* theory, which argues that negative sentiment is a strong driver of negative links between individuals, we interpret observed negative links as a sign of discord or tension [27]. As we note throughout this thesis, social interactions can be incredibly diverse, representing different modalities, or sentiments ranging from support to hostility. Negative interactions significantly impact offline social networks

and individual outcomes [159]. Following directly from Chapter 3 and Chapter 4, we employ a signed network representation, which offers a powerful tool for capturing this complexity.

Existing works on structural polarization in online media using signed networks have primarily been limited to three platforms: Epinions [175] and Slashdot [133], where users explicitly label each other as friends or foes, and Wikipedia, where users can cast positive or negative votes in administrator elections [131]. More recently, signed network representations have been constructed from Reddit and Twitter data, by inferring the valence of an interaction using sentiment analysis applied to comments and posts [169, 168, 114].

As in discussed Chapter 2, signed networks are often studied through the framework of *structural balance*, which in the context of online media, allow for a refined analysis of polarization by identifying communities with internal coherence and cross-group antagonism [99, 42, 102, 54]. Several methods exist to partition the signed graph into these factions, quantify graph-level polarization, or extract ideological information from the network structure, including the two methods presented in Chapter 3 and Chapter 4 [19, 65, 200, 116, 107, 17, 22]. These studies suggest that negative ties provide a deeper understanding of structural polarization by revealing hidden antagonisms or patterns that may not be apparent in networks composed of positive interactions alone [200, 65, 16, 22, 114].

In this chapter, we evaluate the contribution of including heterogeneous (negative and weighted) ties, through an empirical case study on Menéame, a social media platform with naturally occurring signed signals. Menéame ¹ is a popular Spanish social media platform analogous to Reddit, functioning primarily as a news aggregator. On Menéame, users can post links to news stories; the posts appear in the Menéame newsfeed with a hyperlink, information about the posting user, and a brief description of the story. Other users in the ecosystem can subsequently vote (up or down) or comment on the post [67].

The posted articles cover a broad selection of topics, ranging from sports to local and international politics. Given how users interact on the Menéame platform, a signed network representation naturally emerges; the signed signals can be obtained directly from votes that users leave on the articles and comments posted by other users. While Menéame has been studied before from various perspectives [110, 88, 12], to our knowledge, no prior work has exploited the naturally signed structure of the social network ecosystem to study polarization. The design of the platform is such that the main newsfeed appears the same for all users, because there is no personalized recommendation algorithm; thus, any observed structural polarization is due only to user preferences.

In general, there exist several measures to identify and quantify polarization at the network level, on both unsigned and signed networks (see Chapter 3 for a brief review) [108]. In this chapter, however, we adopt an individual-level perspective, seeking to quantify users' ideological positioning on a continuous scale, to obtain a granular user-level characterization of the network.

¹<https://www.meneame.net/>

The primary contribution of this chapter lies in our assessment of the value of incorporating negative interactions when measuring online polarization at the individual level. In what follows, we construct both a signed and unsigned network representation of the Menéame data, comparing the resulting patterns of structural polarization. Structural polarization is identified on unsigned social networks which admit a clustering of individuals into opposing groups with limited (but existing) communication between them [26, 25]. This can be measured using an embedding generated via Correspondence Analysis (CA), first presented in Chapter 2, applied directly to the network [90].

On signed networks, polarization is traditionally identified by antagonistic factions [200, 65, 114, 71, 190]. Because empirical networks rarely exhibit perfect structural balance, we use the SHEEP embedding algorithm (Chapter 3) to assign each user a score capturing their extremity within the network [22].

To quantify the value of negative ties to the study of polarization, we take a dual approach, leveraging both SHEEP and CA. These methods produce lower-dimensional embeddings of network data using spectral techniques and matrix decomposition, respectively, and have been used for latent ideology analysis [22, 26, 76, 72, 73, 164]. A side-by-side comparison of the two methods reveals the added value of incorporating negative ties into network models of political polarization, and allows us to compare how signed and unsigned representations capture individual extremity and alignment within the same underlying social network.

Thus, our study addresses the key research question; *What do negative interactions reveal about individual-level polarization that positive interactions cannot?*

In what follows, we first describe the data collection process, network construction and the embedding techniques. We then present our findings, comparing the insights gained from the signed and unsigned networks, focusing on topic-specific polarization and the role of the negative interactions. We find that Menéame users can be grouped into two main ideological factions that exhibit structural polarization, more pronounced in discussions around controversial topics. While the signed and unsigned methods largely agree in identifying ideological groups, negative ties uncover critical patterns, particularly at the political extremes, that remain hidden in unsigned networks. Finally, we discuss the implications of our findings, highlighting the complementary nature of combining methodologies to enhance our understanding of polarization.

6.2 Data and Methodology

6.2.1 Menéame Dataset

Menéame is a Spanish platform created in 2005, which functions as both a news aggregator and a social media, designed to promote community participation in sharing and discussing news. Users can submit posts containing a hyperlink to an external source, such as news outlet articles or other social media posts, which appear on the newsfeed with the link, information about the posting user,

and a short description, as depicted in Fig. 6.1. The platform is organized into several sections, but this study focuses on the main page; a newsfeed displaying the most popular stories, and the queue, where new stories are listed. Users can interact with posts via up or down votes and comments, they can also comment or vote on other comments. Stories that generate high positive engagement, but little negative engagement, appear at the top of the feed.

While the platform guidelines state that users should use negative votes to report spam ² and to help remove content that violates platform guidelines, in practice, negative votes are used primarily to express disapproval.

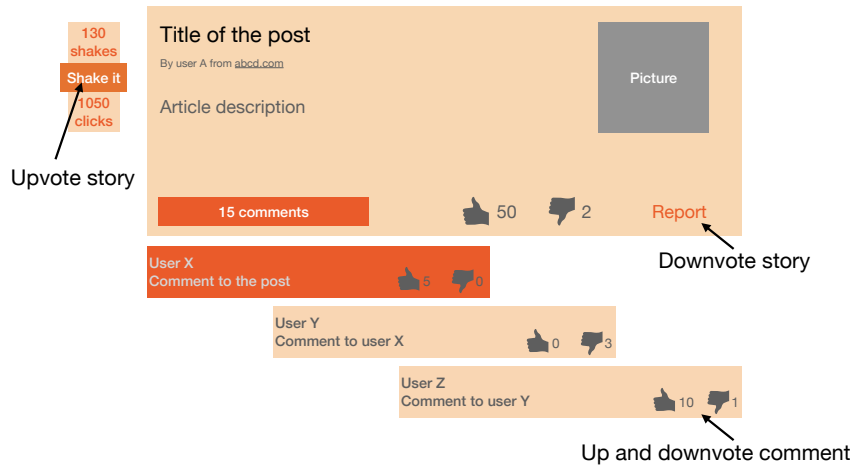


Figure 6.1: **Schematic representation of a Menéame platform story.** Users can upvote and downvote stories, and comments on the story. Downvoting stories is possible only for registered users through the "Report" button, while anyone, including unregistered users, are able to upvote. Only registered users may vote on comments. Comments with many positive votes are highlighted in orange on the platform.

The dataset employed in this analysis contains all stories, comments, and votes appearing on the platform during the period spanning December 1st, 2022, to August 8th, 2023. In aggregate, the dataset comprises 47,887 stories, 1,869,190 story votes, 704,636 comments, and 3,113,863 comment votes, from 11,156 unique users voting on stories, and 8,604 users voting on other users' comments. This data was obtained with ethical approval, and the final dataset is available in aggregated and anonymized format (see [39]).

From this data, we construct two types of undirected signed networks: a *user-to-user* network, using the votes on comments; and a *user-to-news outlet* network, from the votes on stories. These two networks allow us to validate our methods internally, comparing results across the two interaction types, and externally, against external classifications of news outlet ideology. To enable topic-specific analyses, we further divide each network into layers based on the topic of the associated story, identified from its short description. Details on the network construction can be found in Subsection 6.2.3.

²<https://www.meneame.net/faq-es>

6.2.2 Topic Modelling

Topic modelling refers to classifying texts into a finite number of broad thematic categories, or topics [31]. This can be framed as a supervised task when existing labels exist, or unsupervised, as in our case, when no prior topics are indicated in the data. Our objective here is to identify the topic of each posted story to separate the data into subsets, allowing us to construct a multiplex network where each layer corresponds to a different topic. Given the absence of ground truth labels against which to benchmark our topic classifications, we adopt a comparative strategy. We employ two different topic modelling algorithms, BERTopic [92] and hierarchical Stochastic Block Model (hSBM) [85], on the same text corpus, and compare their degree of agreement. Using this comparison method, we are able to obtain a robust classification for a restricted subset of stories. In what follows, we describe the two algorithms in detail and the methodology used to evaluate their correspondence.

BERTopic

The *BERTopic algorithm*, developed by Grootendorst [92], is a modular text-embedding-based approach for identifying topics from a text corpus. Using the language model BERT (Bidirectional Encoder Representations from Transformers) [60], the algorithm first generates text embeddings, subsequently performing dimensionality reduction and clustering on the embeddings to identify coherent topics. Each topic is associated with a set of keywords, derived using a variation of the Term Frequency-Inverse Document Frequency (TF-IDF) procedure, which facilitates the interpretation of the identified topics. A detailed account of the algorithmic steps is provided in Appendix B.2.1, and for more details on the parameter choices, see [39].

hSBM Topic Model

The TM-hSBM, proposed by Gerlach et al. in [85], is an application of community detection methods for topic inference on a text corpus. Unlike BERTopic, which is based on sentence embeddings, this approach constructs a *bipartite* network from the text and document information. A bipartite network is one with two distinct types of nodes, and edges exclusively exist between nodes of unlike type. In this case, the two groups of nodes are the words and documents (equivalently, comments or stories). An edge exists between a word and a document if it appears therein. The hSBM procedure then applies a Bayesian hierarchical Stochastic Block Model (hSBM) inference [163] to the bipartite network, a method typically employed to detect communities in networks. In this context, words are grouped into topics if they share similar connectivity patterns; that is, they tend to appear in the same sets of documents and be absent from the same others [85]. More details on the procedure is located in Appendix B.2.2.

Combining the results of BERTopic and hSBM

We apply both BERTopic and TM-hSBM to the text corpus of story descriptions. BERTopic identifies 180 topics containing 38,883 stories in aggregate, labelling 16,633 stories as outliers. See [39] for a full list of the recovered topic names and keywords. We implement an outlier reduction procedure based on cosine similarity between text embeddings, after which we further consolidate the results by reducing the number of topics to 30.

TM-hSBM produces topics at different hierarchical levels. At the highest level of granularity, it identifies 114 topics, which are subsequently aggregated into 20, 4, and 1 topics, respectively (See [39] for a full list of the topic keywords at each level). Comparing the outputs of the two methods, applied independently to the corpus, allows us to obtain more robust topics in absence of ground truth. We compute the classification accuracy score and find that BERTopic with 30 reduced topics most closely aligns with the second hierarchical level of TM-hSBM.

To interpret the BERTopic results, we examine the most representative keywords for each topic produced based on their TF-IDF weights. As an informative example, Topic 1 is associated with keywords such as “*ukraine, russia, war, russian (male), russian (female), putin, russian (plural), nato, military*”, which we manually label with the broader term, the “*Russia-Ukraine War*”. To merge results across methods, we assess the overlap between BERTopic and hSBM topics and use this overlap to construct robust macro-topics.

Continuing with the example above, *hSBM Topic 1* (keywords: “*ukraine, russian (male), war, russia, ukrainian (male), trump, putin, militar, nato, invasion*”) and *hSBM Topic 16* (keywords: “*nuclear, chinese (male), ukraine, russian (male), militar, fire, russia, wagner, china, american*”) exhibit overlaps of 52% and 46% with *BERTopic Topic 1*, respectively, and are grouped together under the same macro-topic, the *Russia-Ukraine War*.

In cases where multiple BERTopic topics are aligned with the same hSBM topic, we merge the BERTopic topics. For instance, the *hSBM Topic 6* (“*rent, healthcare, hospital, union, labor, doctor, strike*”) is aligned with both the *BERTopic Topics 11* (“*health, strike, doctors, hospitals*”), *14* (“*workers, labor, work, employment*”), and *Topic 18* (“*education, students, teachers, schools*”), and thus is manually labeled as the broader term “*Public Services*” (Figure B.1 in the Appendix displays a confusion matrix describing the topic overlaps).

Following this procedure to combine the results from BERTopic and hSBM results in seven robust macro-topics: *Broad Politics, Russia-Ukraine War, Public Services, Crime, Climate Change, Cryptocurrencies/tech, Inflation*. We retain only the subsample of stories classified as the same topic by *both* algorithms.

To validate this approach, we conducted a manual inspection of a random sample of comments. Among 100 comments where both algorithms agreed, 93 classifications were approved and 7 disapproved. Among 100 comments where the algorithms diverged, we approved BERTopic’s classification

in 34 cases, hSBM's in 26 cases, neither in 39 cases, and both in only one case. Table 6.1 reports the final number of stories associated with each macro-topic. While only a small subset of the initial corpus is retained for our downstream analysis, this refined categorization provides a more reliable basis for analysing polarization and the role of negative ties on a topic-specific granularity. In what follows, we focus our analysis on the two largest topics: *Broad Politics* and *Russia-Ukraine War*.

Topic	Number of Stories	Number of Votes	Upvotes	Downvotes
Broad Politics	7,427	411,591	378,528	33,063
Russia-Ukraine War	2,394	77,828	64,240	13,588
Public Services	1,621	85,154	81,637	3,517
Crime	1,537	73,465	69,074	4,391
Climate Change	1,458	41,847	39,661	2,186
Cryptocurrencies/Tech	814	26,533	24,414	2,119
Inflation	763	26,724	25,316	1,408

Table 6.1: **Dataset summary.** Number of stories, votes, upvotes, and downvotes per macro-topic in the dataset.

6.2.3 Network Construction

This section details the procedure of constructing the two types of interaction networks used in our analysis; a network of *user-to-user* interactions and a *user-to-new outlet* network.

The *user-to-user* network is our main dataset under study in this work. We consider a network $G = (U, E)$, where U is the set of users, and E , the set of edges. We consider there to be $|U| = n$ number of users, such that the resulting adjacency matrix has form $A \in \mathbb{R}^{n \times n}$. Each entry A_{ij} represents the sum of the signed votes between user i to the comments of user j , and the votes of user j to the comments of user i , such that the resulting network is undirected and A is symmetric. When users i and j have never interacted, or have interacted equal number of times positively as negatively, $A_{ij} = 0$. This construction is motivated by the assumption that the absence of interaction between two users is equivalent to a net sum of zero interactions, that is, when the number of positive votes equals the number of negative votes. Moreover, such cases are incredibly sparse; positive and negative votes cancel out exactly in only 0.74 % of pairs. The value of A_{ij} is bounded between $([-n_{ij}, +n_{ij}])$, where n_{ij} is the number of interactions between users i and j , such that A_{ij} attains the extremum when there are *only* negative or *only* positive interactions between the users. We subsample the data to retain only active users who cast more than 10 votes on the platform during the period of observation. Although the format of the data and the platform could naturally produce a directed network representation, we make the modelling choice to remove the directions from the graph to focus on the impact of the signed and weighted edges. Empirically, we find that voting behaviour between users tends to be symmetrical, as illustrated in Tables B.1

and B.2 in the Appendix. In what follows, we refer interchangeably to the *user-to-user* network, or the unipartite network, since all nodes are of the same type (users).

The *user-to-news outlet* network is used primarily to validate our approach, as we can benchmark the ideological positions of the news outlets generated via the embeddings, against external classifications of news outlet ideology. In this case, we consider a bipartite network $B = (V, E')$, where, as before, V is the set of nodes, and E' , the set of edges. Here, the nodes fall into two distinct classes; users U or news outlets O , forming two disjoint sets, such that $V = U \cup O$. As edges connect nodes from one subset to the other only, we have that $E' \subseteq U \times O$.

The network can be represented by an adjacency matrix with shape:

$$A = \begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix},$$

where M is the incidence matrix that has shape $|U| \times |O|$, such that each entry M_{kl} is the sum of the signed votes of user k on news outlet l . As a user can vote on many different stories from a given news outlet, we aggregate votes at the outlet level.

Following common practice in latent space models, where ideological positioning is inferred from the balance of positive and negative votes, we remove the news stories that received only positive votes [167]. In the following analysis, we find this step to be fundamental in quantifying the impact of controversial stories, reflecting the fact that stories receiving exclusively positive votes do not polarize the discussion.

6.2.4 Network Embeddings

We employ two different network embedding techniques, using SHEEP for the signed network, and Correspondence Analysis (CA) for the unsigned network. SHEEP, presented first in Chapter 3, is a spectral embedding method capable of representing proximal information between nodes in signed graphs. CA, as presented in Chapter 2, is a matrix decomposition technique that produces embeddings of categorical data.

In the case of SHEEP, we apply the method to each topic network for both the unipartite signed *user-to-user* vote networks and the bipartite signed *user-to-news outlet* vote networks. We employ a normalized version of the adjacency matrix, where the entries are divided by the total degree of the network. Following this procedure yields an embedding for each user and news outlet, which we use in our downstream analysis. As nodes with a preponderance of negative votes are repulsed from the other nodes in one dimension, we project the results using principal component analysis (PCA), retaining only the first dimension.

Correspondence Analysis is typically applied directly to interaction matrices. In this case, the interaction matrix is the adjacency matrix of the *user-to-news outlet* bipartite graph, where the rows are the users, the columns are the news outlet domains, and the entries are the number of times

each user interacted with the news outlet. To mirror the one-dimensional embedding obtained for the signed graph using SHEEP, we take only the first dimension of the recovered CA embeddings for the following analysis. In practice, the embeddings are obtained via the *prince* Python package [97], which implements CA amongst other statistical techniques.

6.2.5 News Outlet Ideology

To assess the validity of the results obtained by applying CA and SHEEP to the *user-to-news outlet* network, we benchmark against two independent “ground truth” measures of the left-right ideological position associated with each news outlet.

The first ideological measure is obtained via another network-based approach applied to data from Twitter (now X). Using the Tweepy Python package [100] and the Twitter API v2 with Academic Research access, we collected all the tweets authored by the official accounts of major Spanish political parties, where influence is defined by share of the popular vote. We exclude regionalist parties that primarily tweet in languages other than Spanish. As such, the set of political parties included were *PACMA*, *Más País*, *VOX*, *Izquierda Unida (IU)*, *Podemos*, *Ciudadanos (CS)*, *Partido Socialista Obrero Español (PSOE)*, and *Partido Popular (PP)*. We additionally identified the 20 most frequently mentioned accounts by each party’s account, applying a manual filtering procedure to include only accounts belonging to politicians or party affiliated institutions. The complete list of the Twitter handles included in the analysis can be found in the Appendix B.3.

We subsequently aggregate the number of tweets per political party that mentioned one of 40 Spanish news outlets, identified by their website domains, which frequently appear in Menéame posts. These domains include: *rtve*, *abc*, *elmundo*, *atresplayer*, *cope*, *okdiario*, *larazon*, *ondacero*, *telecinco*, *vozpopuli*, *youtube*, *elespanol*, *europapress*, *elconfidencial*, *telemadrid*, *cuatro*, *canalsur*, *eltorotv*, *elindependiente*, *cadener*, *eleconomista*, *elpais*, *lavanguardia*, *esdiario*, *libertaddigital*, *20minutos*, *elperiodico*, *lasexta*, *lavozdegalicia*, *eldiario*, *huffingtonpost*, *facebook*, *twitch*, *infolibre*, *publico*, *laultima hora*, *elsaltodiario*, and *gaceta*.

Using Correspondence Analysis directly applied to the interaction matrix between political parties and news outlets, we obtain an embedding for the political parties, which closely mirrors the established left-right division in Spain (see Fig. B.2 in the Appendix). Thus validated, we take the first dimension of the CA embedding as a “ground truth” measure of the left-right ideological position of news outlets. This procedure closely follows other latent space models for obtaining a measure of ideological positioning [26].

For additional robustness, we also use an independent ideological classification provided by PoliticalWatch³. The qualitative analysis of PoliticalWatch evaluates 30 media outlets’ characteristics,

³<https://politicalwatch.es/blog/political-watch-publica-primer-media-bias-chart-espana-2021/>

including their word-choice, and fact-checking standards on a sample of articles to assess their ideological leaning. The correlation between the ideological position of the news outlets as determined by PoliticalWatch and the Twitter-based method is high (96%), as illustrated in Fig. B.3 in the Appendix.

Together, these two independent ideological measures serve a two-fold purpose; we are able to both validate the results of SHEEP and CA on the *user-to-news outlet* network, and to further understand the results of the embedding methods on the *user-to-user* network. Broadly, if a user is classified as left-wing due to their interactions with other users, this position can be validated by measuring their propensity to vote on stories from left-leaning outlets on Menéame.

6.2.6 Characterizing User Voting Behaviours

Associating User Embeddings with Political Ideology

To validate whether the structural embeddings derived from SHEEP and CA capture meaningful political ideology, user positions in the embedding space are mapped onto a left-right ideological spectrum, using the following procedure. For each embedding, users were grouped into bins of size fifty, based on their embedding position. Within each bin, voting behaviour was characterized by comparing how often users voted positively for each media outlet relative to the platform-wide average. Let V_{um} represent the average vote sign of user u towards media outlet m , and $\bar{V}_m$ the average vote to media m across all users, then the deviation from the average D_{um} is computed as:

$$D_{um} = V_{um} - \bar{V}_m$$

A positive (negative) value of D_{um} indicates that the user casts a higher (lower) frequency of positive votes towards that outlet, as compared to the platform average. The ideology of the media (I_m), extracted from the *user-to-news outlet* network, is combined with the voting deviation to obtain the weighted deviation, W_{um} between user u and outlet m . This is computed as:

$$W_{um} = D_{um} \cdot I_m$$

This quantity provides an estimate of the ideological preference of a user u ; a higher W_{um} indicates a higher share of positive votes towards right-wing media, or a lower share of positive votes towards left-wing media. To account for variations in user activity level, the measure is scaled by the ratio of the user's voting frequency F_{um} to outlet m to the average frequency $\bar{F}_m$ across all users, as:

$$R_{um} = W_{um} \cdot \frac{F_{um}}{\bar{F}_m}$$

This adjustment ensures that outlets garnering disproportionately high or low activity are appropriately weighted in the analysis. Finally, we calculate $R_{bin,m}$ as the average R_{um} for all users in the bin and $F_{bin,m}$ as the voting frequency of users in the bin toward outlet m . The aggregate ideological score for each bin is then defined as:

$$I_{bin} = \frac{\sum_m R_{bin,m} \cdot F_{bin,m}}{\sum_m F_{bin,m}}$$

This value provides an aggregated bin-level estimate of the ideological positioning of users, derived from their voting behaviour toward media outlets, which we use to interpret the results of the unipartite embeddings.

Clustering Interaction Patterns

Next, we analyse interaction patterns between users located across the SHEEP and CA embedding spaces. For each embedding, users are binned according to their position, and two types of interaction matrices are constructed. The first matrix, V_e , is constructed such that $V_{e;ij}$ gives the total number of votes from users in bin i to the comments of those in bin j , where the bins are defined by their position in embedding e . The second matrix, A_e is similarly constructed such that $A_{e;ij}$ gives the average sign of the vote from users in bin i to the comments of those in bin j , for embedding e . For details on the binning process, see [39].

To reduce bias introduced by variations in user activity levels, we apply an iterative normalization process to the total number of vote matrices, A_e . Each row and column of the matrix is iteratively normalized by dividing by its respective sum, ensuring each sum to one. This normalization reflects the relative strength of interactions, independent of individual user activity levels, preventing highly active users from skewing the results, and enabling meaningful clustering of interaction patterns across ideological groups.

The normalized matrices serve as inputs to a k-means clustering algorithm, to obtain macro-level groupings of users according to interaction patterns. These clusters are subsequently used in our downstream analyses and visualizations. Details on the k-means process are provided in [39].

6.3 Results and Discussion

6.3.1 Mapping the Ideological Dimensions of News Outlets

In this section, we use the *user-to-news outlet* network to infer the ideological positions of the outlets. Beyond using these positions to understand the ideologies of the users themselves, the results are themselves valuable; obtaining ideological positions of news outlets that exist in the present polarized media environment gives insight into how contentious topics shape public discourse.

Our analysis rests on the assumption that users are more likely to vote positively for outlets aligned with their own ideological position, and conversely, to either vote negatively or abstain from

voting when the outlet’s stance diverges from their views. The ideological position of both a user and an outlet may very well be issue-dependent, for instance, individuals across the political spectrum may converge in their support for Ukraine in the Russia–Ukraine war. As such, we conduct a topic-specific analysis and focus on the two largest macro-topics we identified on the platform during the period of observation: *Russia-Ukraine War* and *Broad Politics*.

We use SHEEP and CA to quantify the structural position of each news outlet, comparing these results with the external benchmarks of left-right ideology described in the Methods section. By incorporating information from negative ties, SHEEP uncovers patterns of structural polarization that remain obscured in the unsigned network, an effect that is especially pronounced on divisive issues, like the Russia-Ukraine War.

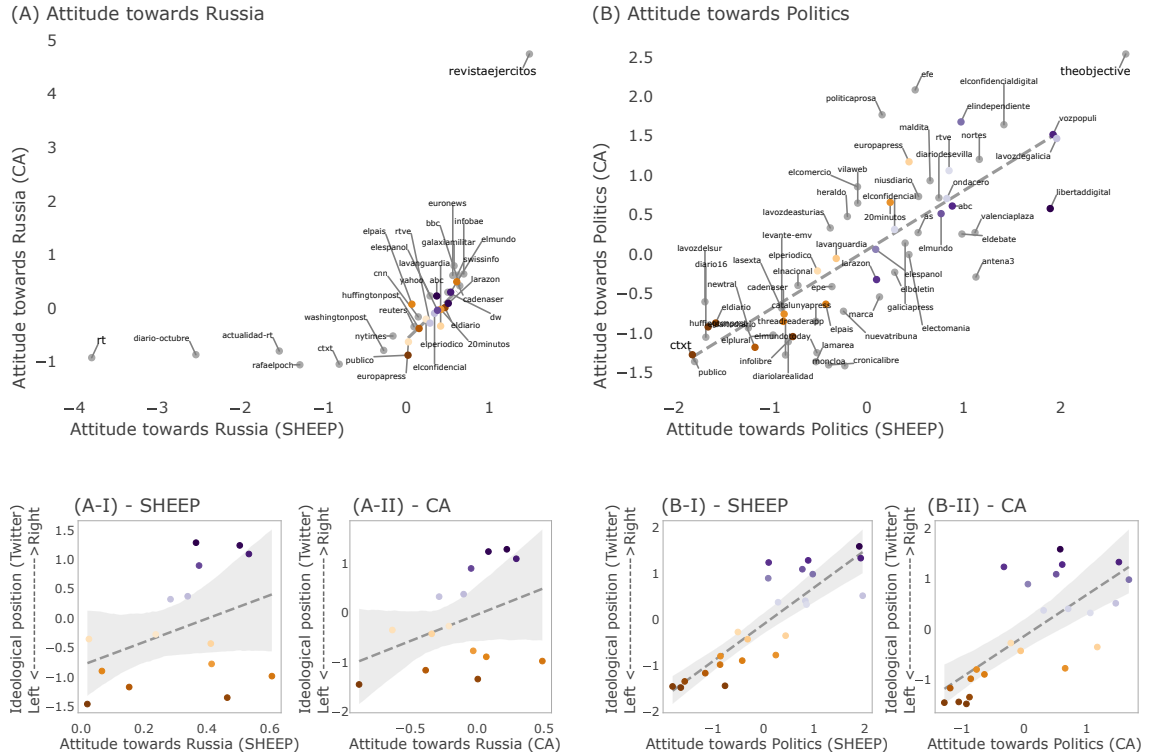


Figure 6.2: Comparing News Outlets’ Attitudes towards the Russia-Ukraine War and Broad Politics. The main panels (A–B) compare the SHEEP and CA embeddings obtained for the news outlets using the topic-specific networks. Colours indicate outlets’ Twitter-derived ideology (brown = left-wing, purple = right-wing, grey = unclassified). Sub-panels (A-I, A-II, B-I, B-II) compare each embedding method with Twitter ideology for a subset of outlets. On the Russia-topic network, both methods identify the defence-focused outlet *Revista Ejército* as an outlier, whereas only SHEEP distinguishes the pro-Russia outlets *Russia Today*, *Diario Octubre*, and *Actualidad RT* from other left-leaning media (*ctxt* or *público*). Neither method is significantly correlated with the broad ideology scores obtained from Twitter. On the politics-topic, the two methods produce highly correlated embeddings, which are also highly aligned with the ideology obtained from Twitter (see panels B-I and B-II).

Russia-Ukraine War

We compare the news outlet embeddings obtained using the SHEEP and CA algorithms on the bipartite *user-to-news outlet* network constructed from stories classified under the *Russia-Ukraine War* topic. The two embeddings are moderately correlated, with a 58% Pearson correlation as depicted in Figure 6.2A. They diverge, however, due to the unique ability of SHEEP to (correctly) identify the outlets *Russia Today (rt)*, *Diario Octubre*, and *Actualidad RT* as representing extreme pro-Russia positions. The CA embedding, in contrast, does not distinguish these Russia-funded outlets from other independent left-wing sources, including *CTXT* or *Publico*, placing them in close proximity. This divergence emphasizes the utility of negative ties at distinguishing left-wing positions from pro-Russia positions. At the opposite end of the spectrum, both methods identify *Revista Ejército*, a publication which, according to their own website “aims to influence political agendas, highlight defence gaps, and promote public investment in Spain’s defence industry”, as extreme in the opposite direction.

We further evaluate the similarity between the SHEEP and CA embeddings with the independent measure of left-right ideology derived from Twitter data in Fig. 6.2A-I and Fig. 6.2A-II respectively. Both methods display only a moderate Pearson correlation with this measure (38% and 39%). This suggests that while Twitter-based ideology, inferred from political party interactions, captures broad partisan alignment, it does not fully reflect issue-specific stances on highly polarized topics. In particular, we note that the left-leaning outlets *eldiario*, *20minutos*, and *cadenaser* are found to be strongly opposed to Russia in this context.

Broad Politics

We subsequently apply the same procedure to the *user-to-news outlet* network generated from stories classified under the Broad Politics macro-topic, which encompass various aspects of Spanish politics. In this case, SHEEP and CA display a high Pearson correlation (76%), as in Fig. 6.2B. The embedding results for both methods are also highly correlated with the ideology derived from Twitter, achieving Pearson correlations of 86.6% for SHEEP and 73.1% for CA, as in Figures 6.2B-I and B-II.

The results indicate that within the Broad Politics topic, both methods largely produce embeddings that align with the left-right political spectrum, while on the Russia-Ukraine War topic, they diverge, revealing issue-specific dimensions that emerge only when negative ties are incorporated.

6.3.2 User-Level Structural Polarization

In the following section, we examine the structural polarization in the *user-to-user* network, to obtain a clearer understanding of what negative interactions reveal about polarization, beyond that which can be obtained from positive ties alone. To that end, we obtain embeddings for the users

by applying SHEEP and CA to the unipartite *user-to-user* networks, constructed from users voting on the comments of other users. Consistent with the previous analysis, we focus on two topical domains: *Russia-Ukraine War* and *Broad Politics*.

Russia-Ukraine War

We begin by visualizing the network using the Fruchterman-Reingold Spring Layout algorithm [81], considering only positive interactions (Fig. 6.3). This algorithm, a refinement of the Spring Embedder method discussed in Chapter 2, similarly assigns attractive forces to connected nodes and weaker repulsive forces to un-connected nodes, to produce a visually informative two-dimensional layout. Both the Spring Layout visualization and the embeddings generated by SHEEP and CA (indicated by node colours in Figures 6.3A–B) reveal the presence of two distinct factions in the network.

By manually inspecting the users’ comments, we identify the two factions obtained via Spring Layout as predominantly representing views that can be defined as 1) anti-NATO and 2) pro-arming Ukraine. An example of a comment from the anti-NATO faction states: “You speak as if Russia is preventing Ukraine from joining UNICEF. NATO is not just any “supranational body”, it is a military alliance with its missiles pointed at Russia.” A similarly representative comment from the pro-arming Ukraine faction is: “Letting Russia do whatever the hell it wants is much more dangerous than arming Ukraine. Among other things because otherwise in a year you would have 3 or 4 other major countries taking example and going over the top of international laws knowing that they have more legitimacy than before.”

While SHEEP and CA demonstrate significant agreement in their overall classification of users within these factions, they diverge in identifying the most extreme individuals. SHEEP pulls disengaged users with few votes toward the centre of the embedding, while this pattern is less pronounced for CA; the Spearman correlations between degree centrality and SHEEP and CA embeddings are 36% and 22% respectively. Importantly, SHEEP identifies extreme users as those with a more active, hostile behaviour, characterized by numerous negative links. These users are identified as dark red in Fig. 6.3A.

The embeddings generated by SHEEP and CA are themselves highly correlated (Fig. 6.3C) with a Spearman correlation of 88%. As illustrated in the network visualization, the two methods diverge primarily in their identification of the most extreme users. For this topic, both SHEEP and CA reveal bimodal distributions of users (histograms in Fig. 6.3C), which correspond to two primary ideological factions characterized by pro and anti-NATO sentiment.

Both methods identify a larger prevalence of users with a negative stance towards NATO. The bimodal distribution identified by both methods is indicative of structural polarization in the network. We note that the SHEEP embedding distribution displays a long tail at both extremes, while

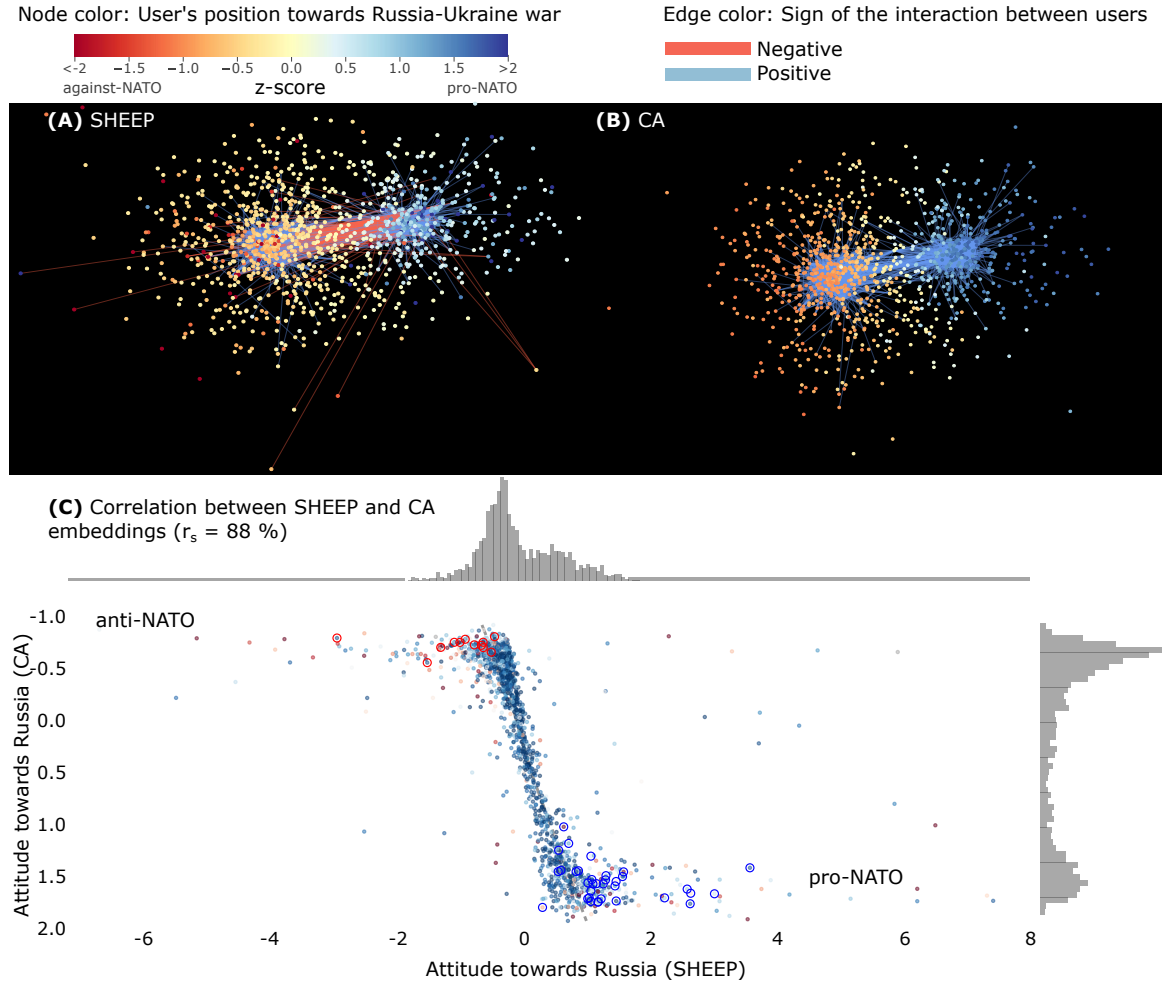


Figure 6.3: **Visualization of user-to-user network for the Russia-Ukraine topic.** Panels A and B show a Spring Layout visualization of a random sample of 3,000 users. Positive edges are drawn in blue, negative edges in red (only in SHEEP), and edges with absolute weight below 3 are omitted. Node colour represents standardized embedding values from SHEEP (A) and CA (B), capped at 2, where red indicates anti-NATO and blue pro-NATO attitudes. The layout reveals two ideological factions with significant cross-faction interaction. In SHEEP (A), extreme users (dark red or blue) are well integrated within the network, while in CA (B) such users tend to have few votes and therefore no visible edges. Panel C compares user attitudes obtained by SHEEP and CA. Each point represents a user, coloured by their voting tendency (dark blue = positive, red = negative). Blue and red circles highlight users who positive towards *Revista Ejércitos* and *Russia Today*, respectively. Histograms show the distribution of users across the embedding space for each method. SHEEP captures extreme negative voters in both factions, and overall the methods are strongly correlated (Spearman correlation of 88%).

CA does not. These long tails are driven by negative votes; this is evident in the colour coding in Fig. 6.3C, where red indicates users who predominantly cast negative votes, and blue indicates those who predominantly vote positively.

Users located at the tail-ends of the SHEEP embedding frequently cast negative votes, and engage with the two most extreme outlets in opposite directions: *RT* and *RevistaEjércitos*. Users who vote positively for these outlets are respectively marked with red and blue circles in Fig. 6.3C. In contrast, CA does not clearly distinguish extremist pro-Russia users from left-wing users who criticise NATO in the years preceding the Russian invasion.

We subsequently follow the procedure outlined in Section 6.2.6 to obtain cluster assignments, using the embeddings and the voting behaviour of the users within. For each embedding method, we construct the normalized voting probability (Fig. 6.4C–D) and average vote sign (Fig. 6.4E–F) matrices. Clustering these matrices with k-means yields ideological factions within the user base and facilitates the interpretation of the embedding results. For both SHEEP and CA, the procedure identifies three factions of users, which are differentiated by colour in the histograms in Fig. 6.4, and by the boxes in the matrices. As the normalized voting probability matrices show, users predominantly vote on other users within their own faction.

As in Section 6.2.6, we use user–outlet voting propensities to characterize the ideology of the three factions, which can be classified as anti-NATO, moderate, and pro-NATO. The embeddings differ in the way they treat users with sparse voting history: CA considers users who have few votes, but vote predominantly on extreme users, to be themselves extreme, SHEEP considers these users as moderate, as there is not enough information about their latent ideology to pull them far away from the centre (see also Fig. 6.3). As such, SHEEP identifies a larger “moderate” faction.

The faction ideologies obtained via the user-media outlet interactions, align well with the voting propensities in the user-user network, as reported in the average vote sign matrices. As Figures 6.4E–F illustrates, users in the pro and anti-NATO factions vote positively for users with similar attitudes and negatively for users with different attitudes (Figures 6.4E–F). We also find that while moderates do not tend to vote negatively in general, they predominantly cast positive votes for the anti-NATO faction, indicating a closer ideological affinity with this group. This pattern also emerges from the clusters identified by the CA method, which excludes negative ties, suggesting that the information encoded by negative ties is partially available in positive ties—i.e., the absence of interaction is related to the propensity to vote negatively.

Overall, both CA and SHEEP identify the polarization apparent on the platform in discussions regarding the Russia-Ukraine war, identifying a bimodal user distribution. However, only SHEEP captures the inter-faction hostility that we observe as negative voting patterns between extreme users in opposing factions.

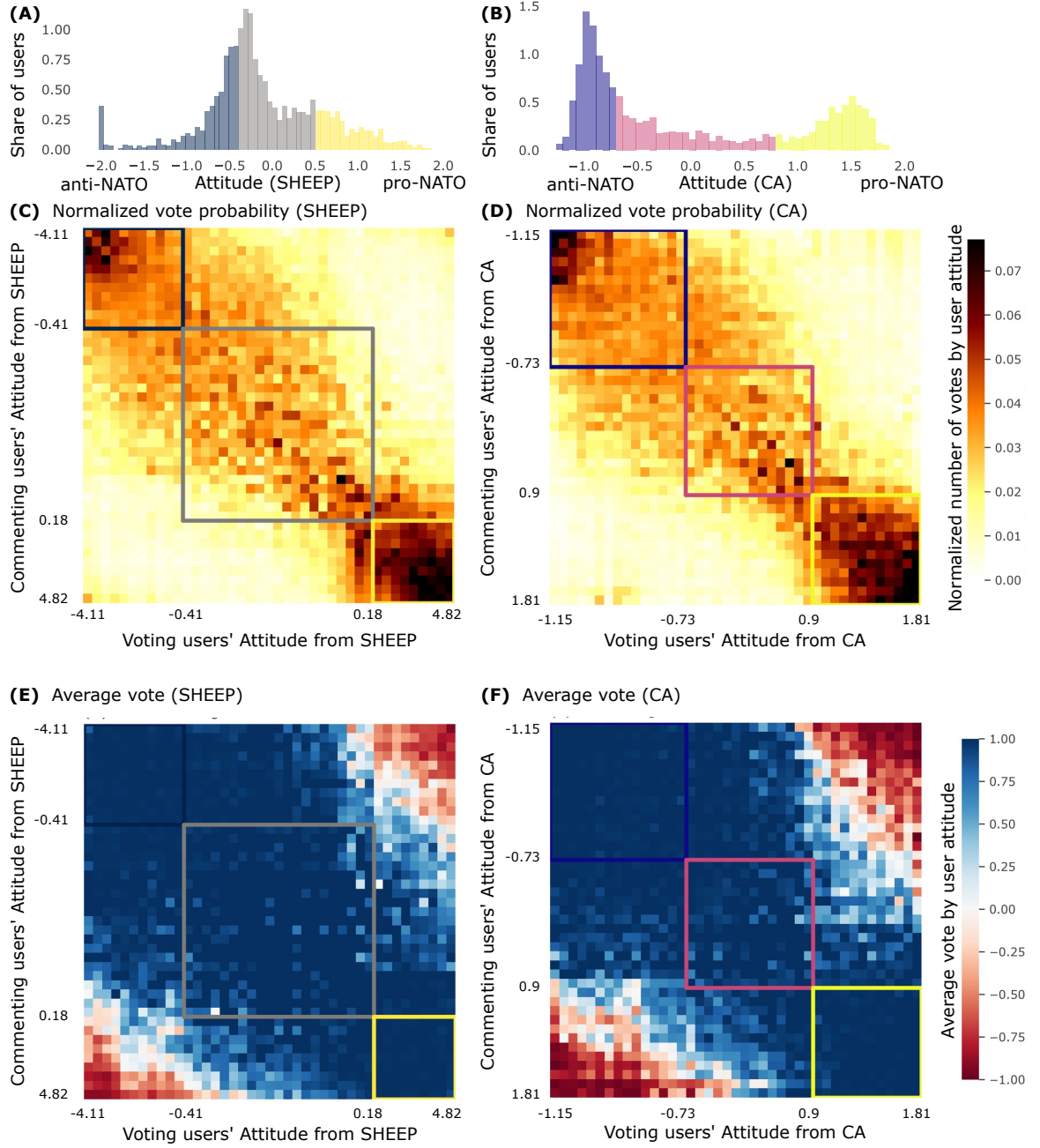


Figure 6.4: **Analysing users' views on Russia-Ukraine topic.** Panels A and B show the distribution of users across the embedding space for SHEEP and CA, with bar colours indicating the obtained k-means clusters (see 6.2.6 and 6.2.6 for details on the procedure). Panels C and D present heatmaps of the normalized vote probability matrices for SHEEP and CA, where the rows correspond to voter and the column to the comment author. Most voting occurs between users with similar attitudes, producing concentrations along the diagonal. Panels E and F show the average vote probability matrices, ranging from -1 (all negative) to +1 (all positive). Users typically vote positively, but extreme users are more likely to downvote the opposite extreme. Clusters identified by k-means (matching panel A/B colours) are highlighted as boxes on the heatmaps: anti-NATO (blue, dark purple), moderate (grey, pink), and pro-NATO (yellow)

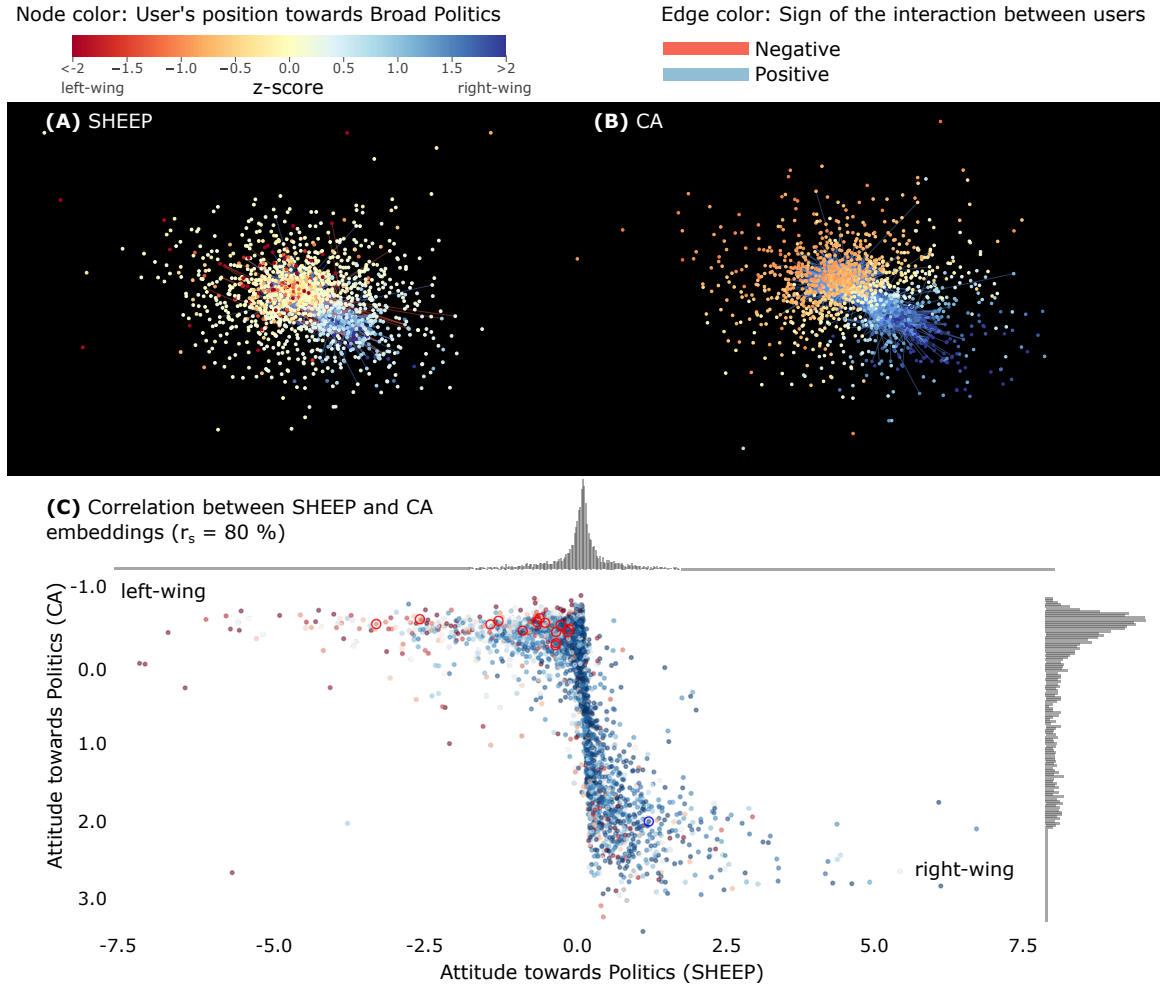


Figure 6.5: **Visualization of *user-to-user* network for Broad Politics.** Panels A and B show a Spring Layout visualization of a random sample of 3,000 users, following the conventions of Figure 6.3. The layout again reveals two main ideological factions with significant cross-faction interaction. In SHEEP (A), extreme users are embedded within the network, whereas in CA (B) such users tend to have few votes and therefore lack visible edges. The histograms display user distributions across the embedding space for each method. SHEEP places extremely negative voters within the left faction, while overall the two methods provide consistent results, with a Spearman correlation of 80%.

Broad Politics

In this section, we apply the same procedure as in the Russia–Ukraine case to the *user-to-user* network representing Broad Politics. Unlike in the previous case, the visualization of the politics network provided in Fig. 6.5 does not reveal two clear ideological factions in the user base. Instead, the distribution of users appears to follow a more continuous transition across the ideological spectrum. A representative comment from the left-extreme states: “of course, you talk like “left-wing voters,” but in the end, you end up saying that you’ll let “the right” win. Very logical, all of it. Then it’s four years of eating shit.” Similarly, an illustrative comment from the right-extreme is: “Isn’t that exactly what Sánchez [Spanish PM] wants to do to govern? Despite having fewer votes

than the PP [main left-wing party], it seems he'll be the one governing. I haven't seen any criticism about that here... Could it be that everything the left does seems fine to you, and everything the right does seems wrong?"

As before, SHEEP tends to position highly active hostile users at the ideological extremes—those with many negative links, represented as dark red nodes in Fig. 6.5A. This results in a higher correlation between degree centrality and ideological position (9% for SHEEP versus 4% for CA).

For this topic, we observe significant differences between the results of the two embedding methods, as shown in Fig. 6.5C. CA, unlike SHEEP, produces a bimodal distribution of users, dominated by a large majority of left-wing users. The SHEEP embedding, however, exhibits a unimodal distribution with long tails, particularly on the left-wing side (Fig. 6.6A). While the user distributions have different shapes, the position of users within the embedding are remarkably similar, with a Spearman correlation of 80%. The main difference occurs at the tails, where the two methods identify different users as extreme. As before, this difference is driven by negative votes; only SHEEP can separate far-left users from other left users. These individuals, who vote positively toward far-left media, are indicated by red circles in Fig. 6.5C.

As in Section 6.2.6, we perform the k-means clustering on the user interaction matrices, obtaining four clusters. Using user-media vote propensities, we identify these clusters as far-left, left-leaning, right-leaning, and far-right. As in the normalized vote matrices provided in Fig. 6.6C–D, users on the left and the right are more likely to vote on comments of similar users. The average vote matrices in Fig. 6.6E–F reveal that only far-left and far-right users vote negatively towards the opposite faction. Importantly, the asymmetry of these matrices indicates that extreme far-left users exhibit a strong tendency to downvote far-right users, whereas far-right users do not engage with left-wing users to the same extent (see bottom-right corner of Fig. 6.6E–F). This could be a platform-specific effect, perhaps related to the higher number of left-wing users on the platform.

Overall, the analysis indicates that the platform's audience is skewed toward left-leaning users. In general, users engage positively with others who share similar views, while negative voting is concentrated among the most ideologically extreme. In particular, left-wing extremists employ negative votes to target users in the opposing extreme faction. These findings highlight the importance of negative interactions for detecting and distinguishing the extreme users, and their antagonistic tendencies.

User Ideology Validation

Finally, we validate our results on the unipartite network by comparing the structural positions generated by SHEEP and CA on the *user-to-user* network with the measure of ideology obtained from user-media voting propensities, following Section 6.2.6.

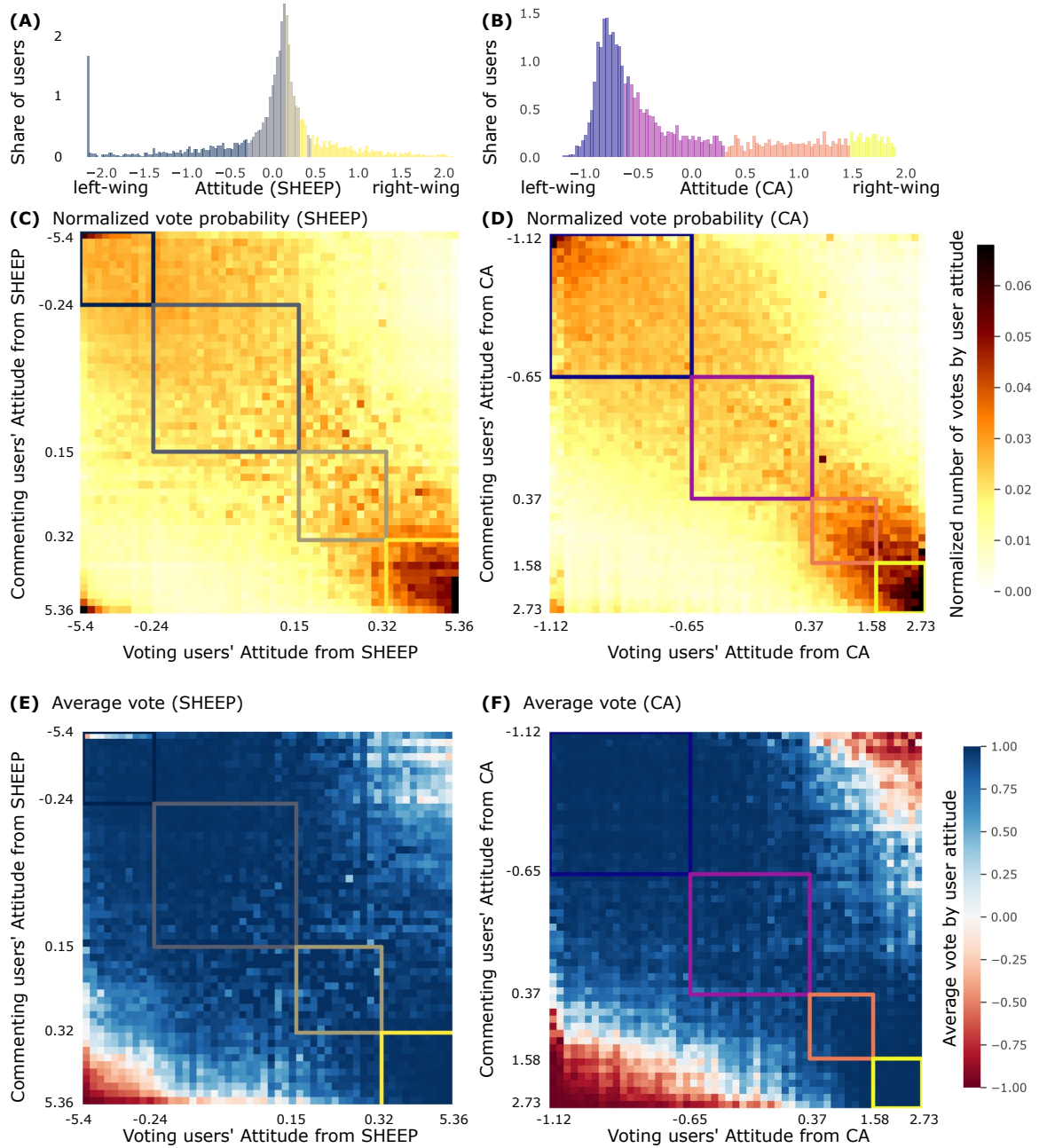


Figure 6.6: **Analysing users' views on Broad Politics.** As in Figure 6.4, Panels A and B show the distribution of users across the embedding space for SHEEP and CA, with bar colours indicating the obtained k-means clusters. Panels C and D display heatmaps of the normalized vote probability matrices for SHEEP and CA, while Panels E and F show the average vote probability matrices. Users generally vote positively, except for extreme left-wing users, who vote negatively towards the opposite extreme. Clusters identified by k-means (matching panel A/B colours) are highlighted as boxes on the heatmaps: far-left (dark blue, dark purple), left-wing (light blue, pink), right-wing (grey, orange), and far-right (yellow).

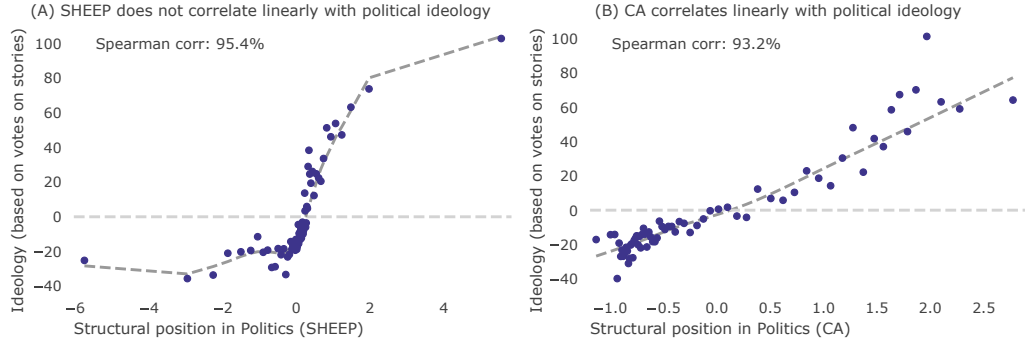


Figure 6.7: **Comparison of SHEEP and CA with ideological positioning.** The x-axis shows binned structural positions obtained from the embeddings: (A) SHEEP and (B) CA. The y-axis indicates the average vote of users in each bin toward stories in the bipartite network, weighted by the average number of votes per domain and the sign of those votes. Each point represents an ideological bin, and the dashed gray line shows a smoothed regression (lowess) to highlight trends. Positive y-values indicate a higher propensity to vote positively for right-wing outlets or a lower propensity to vote positively for left-wing outlets. Compared to SHEEP, which exhibits a non-linear pattern, CA captures political ideology in a more linear manner.

Figure 6.7 plots embedding scores against the ideologies derived from media voting behaviour. CA exhibits a clear linear correlation with left–right political ideology, effectively positioning users along the ideological spectrum. In contrast, the SHEEP embedding displays a non-linear but largely monotonic relationship with ideology, suggesting that it captures additional dimensions beyond simple ideological alignment. Extreme users identified by SHEEP cast negative votes more frequently than other users, and prior work has indicated that negative interactions may signal emotional hostility [27]. As such, SHEEP likely captures underlying elements of this hostility. For most users, this distinction is subtle, but it becomes particularly pronounced during highly polarizing discussion such as those regarding the Russia–Ukraine war, compared to broader political discussions.

Methodologically, we note that SHEEP and CA take as input two different networks; SHEEP takes a signed network, while CA takes an unsigned network. To confirm that the differences we observe using the two methods in the following analysis is indeed due to the addition of the negative ties, rather than the choice of algorithm, we reproduce the above analysis comparing SHEEP and CA, where SHEEP takes as input a null model in which we preserve the existing positive ties and add artificial negative ties to all missing edges, simulating for SHEEP what CA already assumes: that not having a connection implies some level of dissimilarity. We find for both topics that the embeddings generated by the SHEEP null model are highly correlated with those from CA, far more than with the embeddings generated by the original version of SHEEP, as in Figures 6.8 and 6.9. These results confirm that the differences we observe between the signed and unsigned networks are not artifacts of the algorithm choice but are indeed coming from the addition of negative ties, which offer a more nuanced understanding of the platform polarization.

The two methods are, in effect, complementary, revealing different patterns and structures in the underlying network. Although negative votes account for only 3% of all votes on the platform, they

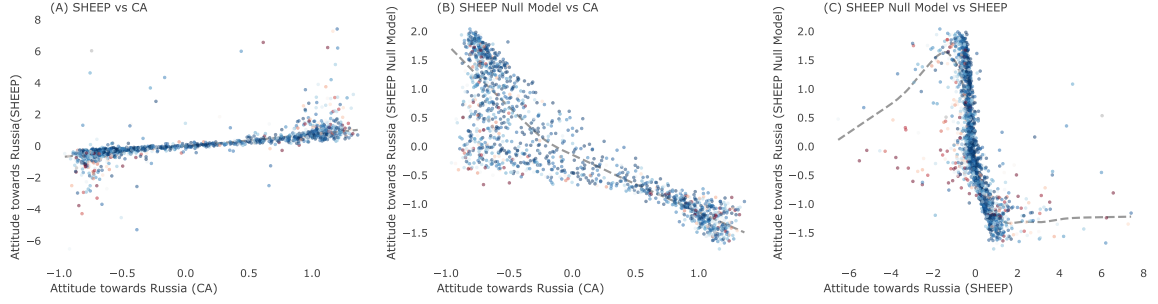


Figure 6.8: **Null model experiments on the user-to-user Russia topic network.** Comparison of user embeddings generated by CA, SHEEP, and SHEEP null model. Colour indicates the tendency of the user to vote positively (dark blue) or negatively (red). The Pearson correlation for SHEEP vs CA is 0.68, SHEEP null model vs CA is 0.83, and SHEEP null model vs SHEEP is 0.61.

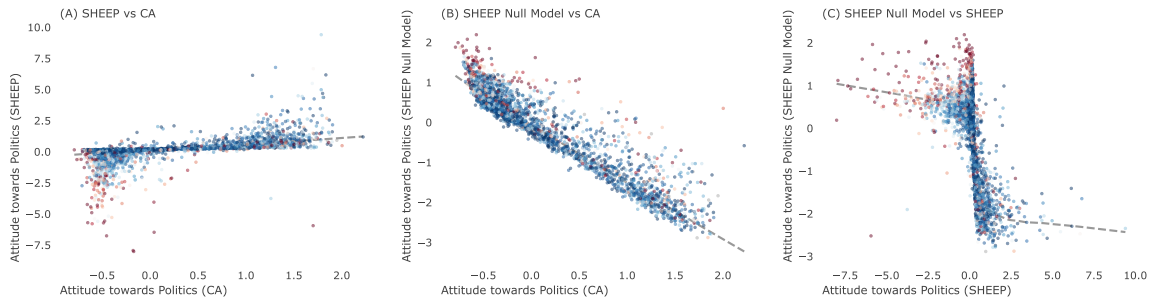


Figure 6.9: **Null model experiments on the user-to-user politics topic network.** Comparing user embeddings generated by CA, SHEEP, and SHEEP null model. Colour indicates the tendency of the user to vote positively (dark blue) or negatively (red). The Pearson correlation for SHEEP vs CA is 0.56, SHEEP null model vs CA is 0.94, and SHEEP null model vs SHEEP is 0.54.

are highly relevant for detecting antagonistic hostility in polarized discussions. Extreme users, both right- or left-leaning, often employ negative votes as a tool against the other faction, while more moderate users tend to cast positive votes, signalling support for those with similar views.

6.4 Conclusion

As social media has become one of the primary mechanisms for ideological exchange and emotional expression, it is increasingly crucial to develop methods to extract information about interactions in digital environments. Because interactions on such platforms are inherently heterogeneous, they provide an ideal stage for our work adapting traditional network science methods to diverse ties. In this final chapter, we have used data from Menéame to both quantify the value of preserving negative tie information, and to investigate the dynamics of polarization apparent on the platform.

Using a dual method approach, comparing the results of SHEEP on the signed network and Correspondence Analysis (CA) on the unsigned network, this analysis revealed ideological divisions within the user-base across two main topics, as well as antagonistic interaction patterns between extreme factions. While both methods successfully map users and news outlets along a political spectrum, as verified against independent measures of ideology, we find that CA shows closer alignment with traditional ideology measures, finding disparate groups of users within the unsigned networks.

However, the negative links prove crucial for understanding interactions between users on more granular, polarizing topics. In particular, on the Russia-Ukraine topic, we observe extreme users behaving antagonistically towards users with opposing views, and consistently supporting content from Russia-affiliated outlets. These patterns remained undetected by CA, which was unable to distinguish these users from those with traditional left-wing orientations. Thus, we show that SHEEP and the signed network reveal behaviours that remain hidden in unsigned networks, perhaps identifying users who play a disproportionate role in spreading polarizing narratives. Moreover, the signed network naturally reveals groups of extreme users who engage in antagonistic behaviour, echoing the distinction drawn in the Introduction between structural and affective polarization. These dynamics can be asymmetric; on Menéame, far-left users are more likely to employ negative votes to interact with ideological opponents, while far-right users tend to remain isolated, interacting between themselves. In general, the dual-method approach we take illustrates the complementary nature of the signed and unsigned network. Considering both methods and their interplay provides a more nuanced understanding of the platform polarization and ideological dynamics, revealing not only users' ideological stances, but also how they engage with opposing views and contribute to antagonistic activity.

While the methods employed here are broadly applicable to any signed social network, some aspects of the analysis are highly tailored to the nuances of Menéame, and may not generalize to other online platforms or social networks. A valuable direction for future research would be

to replicate this framework on additional platforms with negative signals, to allow for a broader comparison between user behaviour and political leanings across different digital contexts.

Our study advances the theoretical understanding of online polarization, outlining methods to identify extreme users and their behaviours, while simultaneously reinforcing the importance of preserving complexity in heterogeneous networks. The conclusions from the dual-method approach could be used to inform strategies for mitigating the negative effects of online social media and fostering healthier, more constructive online conversations. Ultimately, this work underscores the complex interplay between ideology, emotion, and interaction in digital spaces, and contributes to the broader literature on detecting polarization and methodologies for signed, weighted and multiplex networks.

Chapter 7

Learnings and Future Directions

This thesis presents a series of methodologies and empirical studies that advance the network science literature on social networks with heterogeneous ties. Our contributions deal with social interactions that differ in either *sentiment*, through our work on signed networks, or in *nature*, through our work on multiplex, weighted networks. We developed the *SHEEP* spectral embedding algorithm for signed networks, which can be used to recover continuous node attributes by preserving multi-scale latent structure. We also presented the *weak walk*, a random walk adapted to the signed context, which outperforms other walk-based methods at capturing weak balance structure. We demonstrate the effectiveness of these methods across synthetic networks, and empirical social networks alike.

Our empirical studies further make explicit the value of preserving tie-level complexity. We presented a model of music taste diffusion over the Spotify social network, enhancing the characterization of this spreading process by preserving the weighted and multiplex nature of the network. Similarly, through our dual-method approach to the Menéame social network, we demonstrated that preserving tie complexity, in the form of negative, weighted and topic-specific networks, allows us to probe deeper into the latent structure and interaction dynamics of the users on the platform. Together, these two works situated in empirical contexts, illustrate the impact of maintaining tie-level complexity, not merely as a methodological refinement, but as necessary to gain a deeper understanding of the dynamics and patterns in the network structure.

Though our work here seeks to address the challenges of maintaining heterogeneous links in social networks, it also introduces limitations that arise through both methodological choices and data handling strategies, which highlight important directions for future research. In social network analysis, graph sparsity can present a significant challenge, limiting the effectiveness of network-based methods. An important direction for future research is to investigate the performance of both *SHEEP* and the *weak walk* methods on sparse graphs, and to explore potential modifications. This issue could be addressed through additional post-processing, as with the choice made in Chapter 6 of performing dimensionality reduction on the resulting embeddings. Other works on embedding methods have addressed sparsity through regularization, which could also be a potential path forward [170].

Moreover, graph sparsity poses an epistemological challenge to the study of signed networks, where missing links are treated as equivalent to neutral links, failing to distinguish between genuinely neutral interactions, and interactions that simply did not occur [62]. This ambiguity is amplified for sparse graphs, which are common in social network representations of online platforms, and is further compounded when the network construction aggregates over positive and negative links, thus rendering neutral interactions as equivalent to missing ones, as in Chapter 6. Future work could investigate the downstream consequences of conflating these two types of zeros, develop models that explicitly differentiate between them, or indeed explore alternative aggregation methods during the network construction.

Overall, while our works here indeed sought to increase complexity along one dimension, tie diversity, we necessarily reduced other aspects of the system’s complexity. Nevertheless, reintroducing the additional information lost through our model simplifications, could prove a fruitful direction for future research, and could address several limitations of the current work. In particular, social networks naturally carry a temporal dimension, as interactions, especially on social media, have an associated timestamp. In both Chapter 5 and 6, our network construction discards temporal information by aggregating all interactions occurring in the specified duration. Instead, our methods could be adapted to robustly accommodate fluctuations in the number and density of edges or nodes, facilitating applications to temporal networks and enabling cross-time comparisons. In Chapter 5, temporal information could sharpen the definition of contagion, helping to disentangle causal diffusion from homophily-driven effects, while in Chapter 6 it could help assess whether negative interactions actively cause or simply reflect existing polarization, within the limits of observational data.. Temporal information may provide deeper insight into the interplay between the evolving network, the social interactions occurring overtop, and the latent patterns that emerge in time.

Similarly, the social interactions under study in this work are often inherently directed, particularly in Chapter 6 where one user votes on another user’s content. As with the temporal information, the direction of the edges was suppressed to construct undirected signed networks. Research on directed signed networks remains limited, especially as they have been found to conflict with the classical theory of structural balance [132, 115]. Nevertheless, directionality carries important behavioural information in social systems, distinguishing between those who initiate antagonistic interactions and those who receive them. Incorporating direction would enable richer analyses of asymmetries in influence and antagonism, and help determine whether such sentiments are reciprocated between ideological factions.

Taken together, these directions emphasize a wider methodological objective of developing signed and heterogeneous network models capable of handling additional layers of complexity, while remaining robust and adaptable models. This challenge simultaneously echoes one of the fundamental questions in network science, of how model selection should be optimized to be sufficiently, but not

extraneously, sophisticated. Our empirical works presented in this thesis suggest that model selection at the micro-scale can perhaps be guided by a thorough understanding of the macroscopic underlying system, which can be validated and refined using the methods presented in this thesis [124].

Recently, the increasing availability of large network datasets has driven growing interest in more advanced graph machine learning methods, particularly Graph Neural Networks (GNNs). GNNs, and particularly Graph Convolutional Networks (GCNs), learn node representations using both graph connectivity and node attributes, and have outperformed standard network science techniques at downstream tasks like link prediction and node classification on unsigned networks [211]. However, what these methods gain in predictive performance, they lose in interpretability. GCNs have been adapted to the signed network context, notably through Signed Graph Convolutional Network (SGCN), which draws from structural balance theory to define the aggregation method for the positive and negative link information [59, 136]. SGCN illustrates how classical network science methods support and inform these advanced machine learning methods, both in shaping the intuition required for the underlying model design, and in producing interpretable features that can be fed into the model itself. Such hybrid models represent an exciting direction for future research, especially for networks with heterogeneous edges, and an elegant way of combining advanced predictive models with the model selection and system interpretability afforded by network science.

Finally, although our focus is on social networks in this thesis, the methods presented are broadly applicable to any system with heterogeneous edges. In particular, signed networks arise across a diverse array of disciplines, from synchronous and asynchronous interactions between brain regions to financial systems exhibiting positive and negative co-movements [179, 118]. Such networks can be constructed from time series data or bipartite interactions, as discussed in Chapter 2. Further work could be done to ensure these methods are robust across a wide range of applications.

Each one of the future directions outlined above builds upon the current work and may provide deeper insight into the interplay between social network dynamics and heterogeneous ties, further elevating our understanding of emergent behaviour in human society.

Publications and pre-prints over the course of the DPhil

2022 Babul, Shazia’Ayn, Karel Devriendt, and Renaud Lambiotte. “Gromov centrality: A multiscale measure of network centrality using triangle inequality excess.” *Physical Review E*, 106(3), 034312. [20]

2024 Babul, Shazia’Ayn, and Renaud Lambiotte. “SHEEP, a signed Hamiltonian eigenvector embedding for proximity.” *Communications Physics*, 7(1), 8. [22]

2025 Babul, Shazia'Ayn, Yu Tian, and Renaud Lambiotte. "Strong and weak random walks on signed networks." *npj Complexity*, 2(1), 4. [21]

2025 Babul, Shazia'Ayn*, Desislava Hristova*, et al. (*joint first authors) "Link Me Baby One More Time: Social Music Discovery on Spotify." *Proceedings of the International AAAI Conference on Web and Social Media*, 19. [106]

2025 Candellone, E.*, Babul, Shazia'Ayn*, Togay, Ö., Bovet, A., & Garcia-Bernardo, J. (*joint first authors) "Negative ties highlight hidden extremes in social media polarization." *Network Science*, 13, e10. doi:10.1017/nws.2025.10006 [39]

2025 Medina, Sofia, Shazia'Ayn Babul, Timothy LaRock, Rohit Sahasrabuddhe, Renaud Lambiotte, and Nicola Pedreschi. "Detection of anomalous spatio-temporal patterns of app traffic in response to catastrophic events." *EPJ Data Science*, 14(1), 35. [146]

2025 Noulas, Anastasios, Yasin Acikmese, Charles Q. C. Li, Milan Y. Patel, Shazia'Ayn Babul, Ronald C. Cohen, Renaud Lambiotte, and Marta C. Gonzalez. "Modeling Urban Air Quality Using Taxis as Sensors." *arXiv preprint* arXiv:2506.11720. [158]

Appendix A

SHEEP Analysis and Proofs

A.1 Repelling Laplacian Spectra Proofs

Theorem 1. *For an n -complete signed graph subject to the condition that $E^+, E^- \neq \emptyset$, the smallest eigenvalue of the repelling Laplacian is $\lambda_n \geq -n$.*

Proof: Beginning with the n -complete graph of all negative edges, which has the associated repelling Laplacian of form:

$$\hat{L}_r = \begin{pmatrix} -(n-1) & 1 & \cdots & 1 \\ 1 & -(n-1) & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & -(n-1) \end{pmatrix},$$

with eigenvalues:

$$\lambda_n = \lambda_{n-1} \dots = \lambda_2 = -n, \quad (\text{A.1})$$

$$\lambda_1 = 0. \quad (\text{A.2})$$

Any other graph in the set $\{G^n\}$ can be reached by successively changing the sign of an edge (i, j) from negative to positive. Suppose we choose a graph G_1^n which has one positive edge. Without loss of generality, we can assume this edge occurs between nodes $i = 1$ and $j = 2$. For any other edge (i, j) , the proof below holds up to multiplication on the left and right by a permutation matrix P , which does not change the resulting eigenvalues, because the repelling Laplacian is a symmetric matrix with an orthogonal decomposition. The new graph has repelling Laplacian,

$$L_r^{G_1^n} = \hat{L}_r + L^{shift}, \quad (\text{A.3})$$

where:

$$L^{shift} = \begin{pmatrix} 2 & -2 & 0 & \cdots & 0 \\ -2 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The matrix L^{shift} is positive semi-definite with eigenvalue 0 of multiplicity $n - 1$ and eigenvalue $\lambda_1(L^{shift}) = 4$. Since all matrices involved in Eq. A.3 are symmetric and real, and therefore Hermitian, we can apply Weyl's inequality on the sum of two matrices to find that:

$$\lambda_n(\hat{L}_r) + 0 \leq \lambda_n(L_r^{G_1^n}), \quad (\text{A.4})$$

or equivalently,

$$-n \leq \lambda_n(L_r^{G_1^n}). \quad (\text{A.5})$$

Where we recall that λ_n is the smallest eigenvalue of the matrix. This argument can be inductively applied to any graph in the set $\{G^n\}$ with more than one positive edge, directly implying that the ground state energy of the n -complete signed graph is bounded from below by $-n$.

Theorem 2. *Consider a strongly balanced graph $\tilde{G} \in \{G^n\}$, such that the nodes of graph $\tilde{G}$ can be partitioned into two sets V_1 and V_2 , where $V_1, V_2 \neq \emptyset$. Nodes inside each set are connected with positive edges, while the edges connecting the sets are negative. The smallest eigenvalue of the repelling Laplacian is $\lambda_n = -n$.*

Proof: Since $\tilde{G}$ is a complete graph, there are $|V_1||V_2|$ negative edges. We want to find the eigenvector ν associated with the minimal eigenvalue of $L_r^{\tilde{G}}$, which we call λ_n . We also know that the Laplacian $L_r^{\tilde{G}}$ has a 0 eigenvalue corresponding to the vector $\frac{1}{\sqrt{n}}\mathbf{1}$. Since the eigenvectors of $L_r^{\tilde{G}}$ are orthogonal, and assuming $\lambda < 0$, then we have the two conditions that:

$$\sum_i \nu_i^2 = 1, \quad (\text{A.6})$$

$$\sum_i \nu_i = 0. \quad (\text{A.7})$$

In order to minimize the energy associated to the positive edges, set $\nu_i = x_1, \forall i \in V_1$ and $\nu_i = x_2, \forall i \in V_2$. From the above conditions, we have that

$$|V_1|(x_1)^2 + |V_2|(x_2)^2 = 1, \quad (\text{A.8})$$

$$|V_1|x_1 + |V_2|x_2 = 0. \quad (\text{A.9})$$

The ground state energy is:

$$E_{bal}^0 = \nu^T L_r^{\tilde{G}} \nu = -|V_1||V_2|(x_1 - x_2)^2 = \lambda_n. \quad (\text{A.10})$$

Using conditions given in Eqs. A.8 and A.9, along with the fact that $|V_1| + |V_2| = n$, we have that:

$$E_{bal}^0 = \lambda_n = -n. \quad (\text{A.11})$$

By Theorem 1, this is the minimum eigenvalue.

Theorem 3. Consider a strongly balanced graph $\tilde{G} \in \{G^n\}$, such that the nodes of graph $\tilde{G}$ can be partitioned into two sets V_1 and V_2 , where $V_1, V_2 \neq \emptyset$. Introduce frustration into the graph by exchanging the signs of one positive and one negative edge, thereby preserving the overall density of positive and negative links. When the graph is sufficiently large such that $n > 4$, and $|V_1|, |V_2| > 2$, the ground state energy is strictly increased $E_{frus}^0 > E_{bal}^0$.

Proof: Without loss of generality, assume that $|V_1| \geq |V_2|$. Next, we introduce a perturbation to the Hamiltonian associated with the switching of signs of a positive and negative edge. Assume that nodes $i = 0, 1, 2 \in V_1$ and $i = 3 \in V_2$, and we switch the signs of the edges $(0, 1) = 1$ and $(2, 3) = -1$. Then this "switching" perturbation is associated with the matrix:

$$L^{switch} = \begin{pmatrix} -2 & 2 & 0 & 0 & \cdots & 0 \\ 2 & -2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & -2 & \cdots & 0 \\ 0 & 0 & -2 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

We can easily calculate that L^{switch} has eigenvalues -4, 4, and 0 with multiplicity $n - 2$. We take $\mu_n, \mu_{n-1} \dots \mu_1$ to be the associated eigenvectors, ordered according to the increasing eigenvalue. These vectors span $\mathbb{R}^n$ because L^{switch} is symmetric. Note that $\mu_n = \frac{1}{\sqrt{2}}(1, -1, 0, 0 \dots 0)$ is associated with eigenvalue -4. Finally, the Laplacian of the frustrated graph is:

$$L_r = L_r^{\tilde{G}} + L^{switch}. \quad (\text{A.12})$$

We call this graph "frustrated" because the two terms $L_r^{\tilde{G}}$ and L^{switch} cannot be simultaneously minimized. To see this, take ν_n to be the eigenvector associated with the minimum in Eq. 3.9 for the balanced graph, which has the form $\nu_n = (x_1, x_1, x_1, x_2, \dots)$, where, as before, each node is assigned position x_1 or x_2 according to its faction identity. Here, ν_n and μ_n are orthogonal. Furthermore, L^{switch} contributes an increase in energy of $\nu_n^T L^{switch} \nu_n = +2(x_1 - x_2)^2$. We want to show that the ground state energy of the frustrated Laplacian is larger than $-n$. This requires the following three results, the first is a result of a theorem proved in [5]:

1. The spectrum of the adjacency matrix of the balanced graph is equivalent to the spectrum of the unsigned graph.
2. The spectrum of an unsigned complete graph has smallest eigenvalue of 0, and all other eigenvalues are n , with multiplicity $n - 1$
3. $L_o = L_r + 2D^-$.

For a complete graph, $D = n - 1I$, so (1) also implies that the spectrum of the opposing Laplacian is the same as the spectrum of the Laplacian of the unsigned graph. It follows that the second smallest eigenvalue of the opposing Laplacian for the balanced graph is n . The maximum possible entry of $2D^-$ is $2|V_1|$, if a node in V_2 is connected negatively to all members of the other faction, V_1 . Applying Weyl's inequality on the relation given in result 3, to find a bound on the second smallest eigenvalue of the repelling Laplacian of the balanced graph λ_{n-1} :

$$\lambda_{n-1} \geq n - 2|V_1|. \quad (\text{A.13})$$

Recall that $\lambda_n = -n$ for the balanced graph. We use Eq. A.13 to give an approximation for the "energy gap" for the balanced graph, between the ground and first energy levels, and we define $\nabla E = \lambda_{n-1} - \lambda_n \geq 2(n - |V_1|)$.

Next, let α_n be the eigenvector associated to the ground state of the frustrated Laplacian in Eq.A.12. There is a decomposition of α_n in terms of the eigenvectors of the balanced Laplacian which takes the form:

$$\alpha_n = \sum_{i=1}^n c_i \nu_i. \quad (\text{A.14})$$

We can also use the fact that ν_n and μ_n are orthogonal to define a new orthonormal basis of $\mathbb{R}^n$. This new orthonormal basis $\{x_n, x_{n-1} \dots x_1\}$ has $x_1 = \nu_n$ and $x_n = \mu_n$. We can now also express α_n in terms of this new basis:

$$\alpha_n = \sum_{i=1}^n d_i x_i. \quad (\text{A.15})$$

Using the orthogonality of the basis vectors, it follows that $d_1 = c_n$. From this, we have the following two decompositions:

$$\alpha_n = c_n \nu_n + \sum_{i=1}^{n-1} c_i \nu_i, \quad (\text{A.16})$$

$$= c_n \nu_n + \sum_{i=2}^{n-1} d_i x_i + d_n \mu_n. \quad (\text{A.17})$$

For notational purposes, set $\alpha'_n = \alpha_n - d_n \mu_n$. The ground state energy associated with the frustrated Laplacian in Eq. A.12 is:

$$E_{frus}^0 = \alpha_n^T L_r^{\tilde{G}} \alpha_n + \alpha_n^T L^{switch} \alpha_n. \quad (\text{A.18})$$

Using Eq. A.16 to simplify the first term on the right hand side, and equation Eq. A.17 to simplify the second term gives:

$$E_{frus}^0 = \sum_{i=1}^n c_i^2 \lambda_i + d_n^2 (-4) + \alpha_n'^T L^{switch} \alpha_n'. \quad (\text{A.19})$$

Because α'_n is orthogonal to μ_n by construction, it has a decomposition in terms of $\mu_{n-1}, \mu_{n-2} \dots \mu_1$, which are eigenvectors of L^{switch} with eigenvalues greater than or equal to 0. Since the third term

in Eq. A.19 must be greater than or equal to 0, E_{frus}^0 is minimized when d_n^2 is maximized. From A.17, $d_n^2 \leq 1 - c_n^2$, and

$$E_{frus}^0 \geq \sum_{i=1}^n c_i^2 \lambda_i - 4(1 - c_n^2). \quad (\text{A.20})$$

We can now use the energy gap to bound the ground state energy from below, resulting in:

$$E_{frus}^0 \geq c_n^2(-n + 4) + \sum_{i=1}^{n-1} c_i^2(-n + \nabla E) - 4. \quad (\text{A.21})$$

Imposing the conditions that $n > 4$ and $|V_1|, |V_2| > 2$, such that each faction has more than two members. Then, $\nabla E > 4$, and since $\sum_{i=1}^n c_i^2 = 1$, the strict inequality follows:

$$E_{frus}^0 > -n = E_{bal}^0. \quad (\text{A.22})$$

A.2 Bi-Polarization Measure

In this section, we present the results of the bi-polarization measure on synthetic signed networks with a clear polarization structure, generated using signed stochastic block models (SSBMs). The first graph is a two-community SSBM with 50 nodes in each community, and a link probability of 0.5. As this graph naturally has a bi-partition, it also obeys strong balance. Figure A.1 shows the SHEEP (repelling Laplacian first eigenvector) embedding of the SSBM, where negative edges are marked in red, and positive edges are marked in blue. The histogram compares the ground state energy of the strongly balanced SSBM to 1000 realizations of the null model, obtained by reshuffling the edge signs. The resulting bi-polarization measure $Z(G_{bal})$ is -12.24, correctly identifying a significant bi-polarization structure.

We present the same results for an SSBM with frustration, obtained by flipping the sign of each edge with probability 0.2. Figure A.2 shows the ground state energy and null model distribution for the frustrated graph. For the initial realization of the SSBM, the $Z(G_{0.8})$ score is -4.87, which is still significant. Since the SSBMs in Figures A.1 and A.2 have the same number of nodes and edges, the bi-polarization scores can be directly compared, indicating that the latter network exhibits lower polarization.

Finally, in Figure A.3, we examine the bi-polarization scores across several SSBM realizations with varying levels of noise and sparsity, as determined by the link probability p of the affinity matrix. We observed that as noise increases, the z-score of the SSBM decreases. The trend is less pronounced for more sparse graphs, since the effect of the noise becomes more important.

2-Community SSBM, No Noise

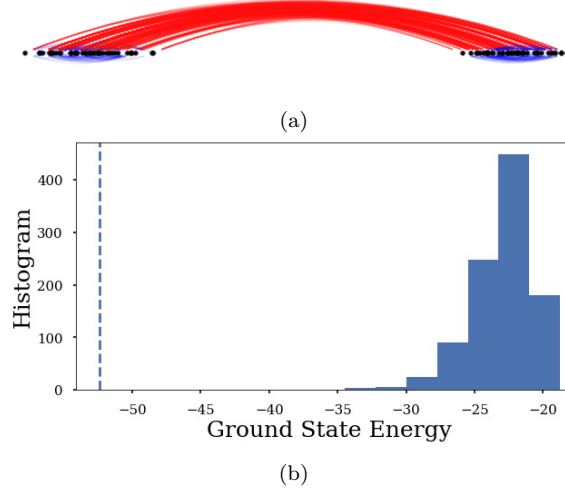


Figure A.1: **Results on strongly balanced two-community SSBM.** (a) One-dimensional SHEEP embedding. (b). Histogram of the ground state energies of the 1000 realization of the null model, with the SSBM ground state energy shown with the dotted line.

2-Community SSBM, 20% Noise

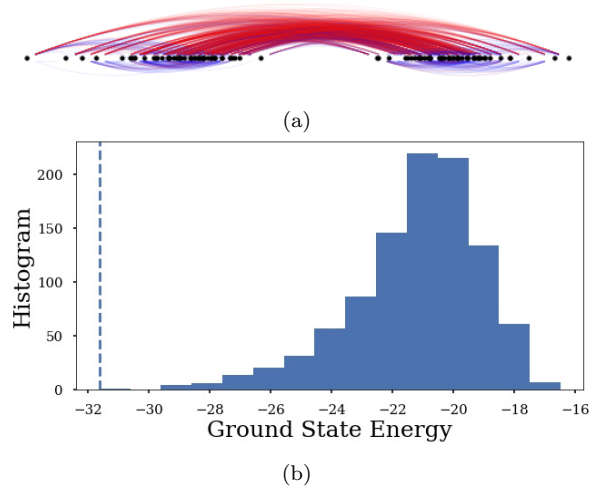


Figure A.2: **Results on frustrated two-community SSBM with 0.2 edge sign flip probability.** (a) One-dimensional SHEEP embedding. (b). Histogram of the ground state energies of the 1000 realization of the null model, with the SSBM ground state energy shown with the dotted line.

A.3 Weak Balance and Extension to Higher Dimensions

The inspiration for the higher dimensional generalized Hamiltonian comes from the spectrum of the n -complete negative graph, which is trivially cluster-able into n antagonistic factions. Intuitively, the optimal SHEEP embedding should place these n nodes at equidistant positions, due to the equal repulsive forces that act on each of them in the Hamiltonian from Eq. A.23 below.

$$H = \sum_{i,j} A_{+,ij} |\mathbf{x}_i - \mathbf{x}_j|^2 - \sum_{i,j} A_{-,ij} |\mathbf{x}_i - \mathbf{x}_j|^2. \quad (\text{A.23})$$

Recall that we imposed a spherical constraint, resulting in the Hamiltonian given in Eq A.24:

$$H(\boldsymbol{\nu}) = \boldsymbol{\nu}^T L_r \boldsymbol{\nu} = \lambda R. \quad (\text{A.24})$$

These positions can be represented by a (squared) distance matrix D with entries defined:

$$D_{ij} = a, \forall i \neq j \in \mathcal{V}, \quad (\text{A.25})$$

where $D_{ij} = |\mathbf{x}_i - \mathbf{x}_j|^2$, and $a \not\rightarrow \infty$ due to the spherical constraint. Arranging n equidistant points in Euclidean space requires $n - 1$ dimensions. We can indeed recover this $n - 1$ dimensional embedding using the repelling Laplacian as follows. The repelling Laplacian corresponding to the n -complete negative graph is of the form:

$$L_r = \begin{pmatrix} -(n-1) & 1 & \cdots & 1 \\ 1 & -(n-1) & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & -(n-1) \end{pmatrix}.$$

Using the fact that $L_r = -nI + J$, where J is the matrix of all 1s, we find that the eigenvalues of L_r are

$$\lambda_n = \lambda_{n-1} \dots = \lambda_2 = -n, \quad (\text{A.26})$$

$$\lambda_1 = 0. \quad (\text{A.27})$$

The $n - 1$ smallest eigenvalues correspond to $n - 1$ orthonormal eigenvectors $\nu_n \dots \nu_2$, which describe configurations with equivalent energy for the Hamiltonian in Eq. A.24.

For each node i , we define an $n - 1$ -dimensional vector $\mathbf{x}_i$, which is the vector built using the i -th entry of each of the first $n - 1$ eigenvectors, i.e. $\mathbf{x}_i = (\nu_{ni}, \nu_{n-1i}, \dots, \nu_{2i})$. The set of points $\{\mathbf{x}_i\}$ define an $(n - 1)$ -polytope, where each point is equidistant from the others. To see this, we define a new matrix B as $B = \frac{-1}{n} L_r$. B has eigenvalues 1 with multiplicity $n - 1$ and eigenvalue 0 with multiplicity 1. It is also symmetric with the same eigenvectors as L_r , and entries of the form:

$$B_{ii} = \frac{n-1}{n},$$

$$B_{ij} = \frac{-1}{n}, \forall i \neq j.$$

Therefore, it admits a singular value decomposition using the non-zero eigenvectors:

$$B = UU^T, \quad (\text{A.28})$$

where $U = (\nu_n, \nu_{n-1} \dots \nu_2)$. The entries of matrix B are equivalent to the inner products of the position vectors $\mathbf{x}_i$, or $B_{ij} = \mathbf{x}_i \cdot \mathbf{x}_j$. Since the relationship between the dot product and the Euclidean distance is given by:

$$|\mathbf{x}_i - \mathbf{x}_j|^2 = \mathbf{x}_i \cdot \mathbf{x}_i + \mathbf{x}_j \cdot \mathbf{x}_j - 2(\mathbf{x}_i \cdot \mathbf{x}_j), \quad (\text{A.29})$$

it follows that $|\mathbf{x}_i - \mathbf{x}_j|^2 = 2$ for all $i \neq j$. We have shown that the multiplicity of the smallest eigenvalue of the repelling Laplacian for an n -complete negative graph gives the coordinates for embedding the graph in $n - 1$ dimensions.

Let us now show that the higher dimensional generalized Hamiltonian is minimized in $k = n - 1$ dimensions. As all possible edges exist and are negative, the unmodified Hamiltonian is:

$$H(D(k)) = - \sum_{i,j} D(k)_{ij} = \sum_{\alpha=n}^{n-k} \lambda_{\alpha}. \quad (\text{A.30})$$

Let us denote $\text{vect}(D(k))$ as the vectorized distance matrix corresponding to the first k eigenvectors, without the diagonal terms which are trivially 0, so that it has length $n(n - 1)$. For $k \leq n - 1$, we have from the equation for the higher dimensional generalized Hamiltonian that:

$$\tilde{H}(D(k)) = \frac{-kn}{|\text{vect}(D(k))|}. \quad (\text{A.31})$$

By the Cauchy-Schwarz inequality,

$$\text{vect}(D(k)) \cdot \mathbf{1} \leq |\mathbf{1}| |\text{vect}(D(k))|, \quad (\text{A.32})$$

where $\text{vect}(D(k)) \cdot \mathbf{1} = \sum_{i,j} D(k)_{ij}$ and $|\mathbf{1}| = \sqrt{n(n - 1)}$. By Eq. A.30, $\sum_{i,j} D(k)_{ij} = kn$. Next, we combine Eq. A.31 with Eq. A.32 to see that:

$$\tilde{H}(D(k)) = - \frac{kn}{|\text{vect}(D(k))|} \geq -\sqrt{n(n - 1)}. \quad (\text{A.33})$$

The Cauchy-Schwarz inequality is a strict inequality, unless one vector is a scalar multiple of the other. If $\text{vect}(D(k))$ is a scalar multiple of $\mathbf{1}$, this implies that the points are equidistant, which is trivially only possible if $k \geq n - 1$ dimensions. Indeed, recall from Eq. A.25 that in $k = n - 1$ dimensions the entries of the distance matrix are given by:

$$D(n - 1)_{ij} = 2, \forall i \neq j, \quad (\text{A.34})$$

such that $\text{vect}(D(n - 1)) = 2\mathbf{1}$. It follows that:

$$\tilde{H}(D(k)) \begin{cases} > -\sqrt{n(n - 1)}, & \text{for } k < n - 1 \\ = -\sqrt{n(n - 1)}, & \text{for } k = n - 1 \end{cases},$$

or, the energy is strictly minimized in $k = n - 1$ dimensions, precisely as we needed to show.

A.3.1 Embedding dimension and number of clusters

In general, even if a signed network has n ground truth clusters, the relationship between the required dimension k and n -clusterability is complicated. In particular, it is possible to represent the distances between n factions in $k < n - 1$ dimensions if the weights of the negative forces between the factions are conducive to a lower dimensional embedding. To see this, suppose that a matrix D^* represents the set of distances between nodes that minimizes the Hamiltonian. Let $\hat{D}$ be the matrix obtained by squaring the entries of D^* . According to the Schoenberg theorem (see [32]), these distances have an embedding into Euclidean space if:

$$\forall \xi \in \mathbb{R}^n, \sum_{i=1}^N \xi_i = 0 \implies \sum_{i,j} \xi_i \hat{D}_{ij} \xi_j \leq 0. \quad (\text{A.35})$$

Suppose that $\hat{D}$ satisfies Eq. A.35. Define the $n - 1$ square matrix C^* by

$$C_{ij}^* = \frac{1}{2}(D_{in}^{*2} + D_{jn}^{*2} - D_{ij}^{*2}). \quad (\text{A.36})$$

Then, the minimal dimension required for the Euclidean embedding is given in [61] as:

$$k_{min} = \text{rank}(C^*). \quad (\text{A.37})$$

As a more illustrative example, consider a 3-complete negative graph with nodes i, j, k and edge weights defined such that $|\mathbf{x}_i - \mathbf{x}_j| = 2$, $|\mathbf{x}_i - \mathbf{x}_k| = 3$ and $|\mathbf{x}_k - \mathbf{x}_j| = 5$ are the distances the minimize the Hamiltonian. Even though there are 3 "factions", only one dimension is required for the embedding if node i is placed in between nodes j and k . Consequently, we emphasize again that the energy minimization is intended to find the dimension that produces the best proximity measure, and not the number of weak balance clusters.

A.4 Rankings on Generated Synthetic Graphs: Further Experiments

In this section, we evaluate the effectiveness of the SHEEP embedding at recovering the initial positions of 100 realizations of the well-ordered synthetic graph under varying initial node distributions and threshold values. To quantify performance, we compute the Kendall Tau (KT) correlation between the eigenvectors and the initial node distribution for SHEEP, as well as for the Laplacian and SPONGE baselines. The corresponding results are presented in Figure A.4. For the case of $n = 100$, the increased node density necessitates the use of lower threshold values. Across all tested configurations, SHEEP consistently outperforms the competing methods.

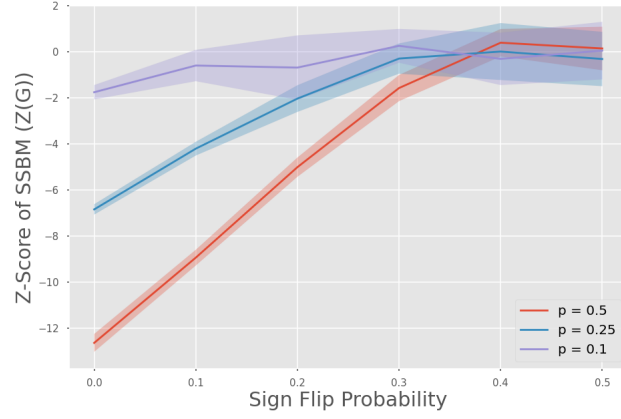


Figure A.3: $Z(G)$ for two-community SSBMs with various parameters. Score computed using 1000 realizations of the null model, the shaded area gives the standard deviation over 20 realizations of each graph.

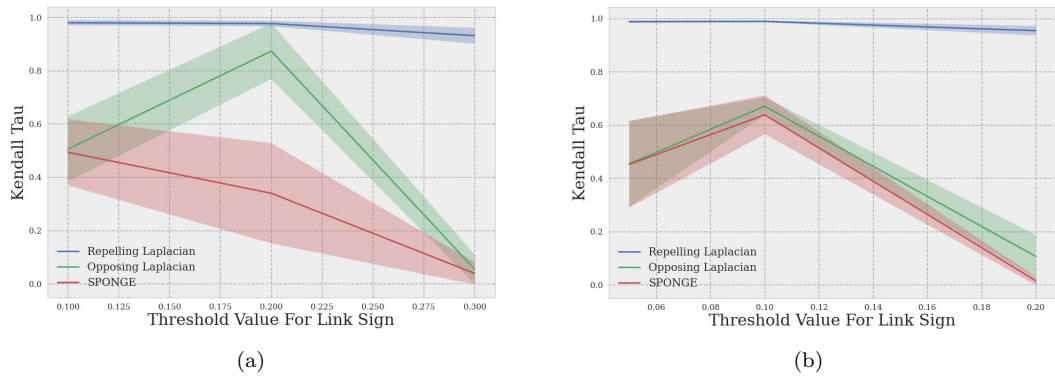


Figure A.4: **Well-Ordered Synthetic Graph Experiments.** Mean KT correlation between initial ranking and SHEEP, opposing Laplacian and SPONGE rankings over 100 realizations, for $N = 50$ nodes (a) and $N = 100$ nodes (b). Shaded area gives standard error.

A.5 Node Extremism Measure and Degree

In Chapter 3, we showed that for a random configuration model, a node’s extremism measure is correlated with its net degree, as in Figure A.5 (a). However, when the graph is not random and instead exhibits structural balance, the norm captures aspects that cannot be explained by the net degree alone. Using SSBMs with varying levels of frustration, we examine how the relationship between the norm and degree changes as graph structure becomes more or less polarized. In networks with strong bi-polarization, we find that the extremism norm is driven primarily by the node’s negative degree. To illustrate this, we consider a two-community SSBM with 100 nodes in each community and a link probability of 0.5. Figure A.5 (b) plots the positive versus negative degree of each node, with colour indicating the norm in the one-dimensional embedding. The KT correlation between the norm and the negative degree is 0.961, while the correlation with the net degree is only 0.483. In strongly bi-polarized graphs, positive edges primarily connect nodes within the same faction. As a result, the positive degree does not significantly influence the extremism measure or the embedding norm. Importantly, in that case, the positive connections of a node do not allow it to stabilise its position and bring it towards the centre. When the bi-polarization structure is perturbed by noise or other structural effects, the extremism measure is no longer determined solely by the negative degree. To demonstrate this, we again consider a two-community SSBM with 100 nodes per community and link probability 0.5, but now perturb the faction structure with a 0.7 sign-flip probability. Figure A.5 (c) shows the relationship between positive and negative degree, with color indicating the embedding norm. In this case, the KT correlation between the norm and negative degree is 0.495, while the correlation with the net degree increases to 0.523. Thus, for graphs which are not characterized by strong balance, both the positive and negative degree play a role in the ”extremism” measure that SHEEP yields, as well as the structure of the graph, indicating that this measure incorporates more than just local information.

A.5.1 Congress 114 Correlations With Degree

As in Chapter 3, the network representing Congress 114 has strong bi-polarization structure along Democrat-Republican lines [18]. As a result, the distance to the origin in the embedding cannot be directly explained by the net degree, unlike in the signed configuration model.

Figure A.6 (a) shows the distance to the origin versus the net degree for nodes in the 114th congress network. The KT correlation is lower than on the random graph experiments, and also lower than the correlation obtained with the NP scores. As Figure A.6 (b) shows, the norm in the embedding, or the ”extremism” measure, is primarily determined by the negative degree, as in the case of the non-noisy SSBM in Figure A.5. In these graphs, positive edges connect nodes within the same faction, often linking extreme members to other extreme members. However, negative degree alone does not fully explain node extremism: positive connections and the broader graph structure

Node Extremism and Net Degree

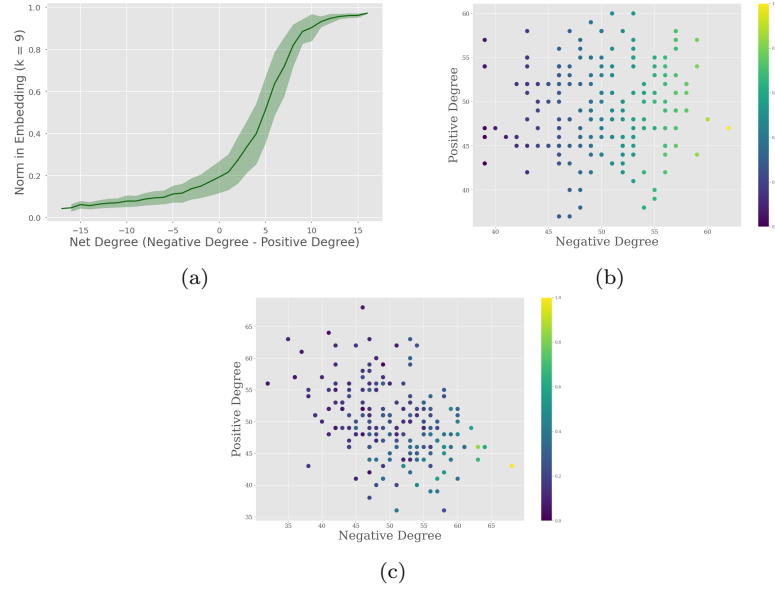


Figure A.5: **Relationship between SHEEP norm, positive and negative degree for various graphs.** a) Mean distance to origin in SHEEP embedding (the norm in $k = 9$ dimensions) versus the node's net degree over 100 realizations of the signed network configuration model, for which $k = 9$ was the optimal embedding dimension. Shaded area gives the standard deviation. Nodes having relatively more negative edges are further from the origin. b) SSBM with two communities, no noise. Positive degree versus negative degree, colour gives the norm in the one-dimensional embedding. c) SSBM with two communities, and 0.7 sign flip probability. Positive degree versus negative degree, colour gives the norm in the one-dimensional embedding.

also play a role, which SHEEP captures through the embedding. Consequently, SHEEP achieves stronger alignment with NP scores than negative degree alone. To quantify this, we compare the Kendall's Tau correlation between NP ideology scores (absolute value), the one-dimensional SHEEP embedding (absolute value), and negative degree. SHEEP shows a higher correlation with NP scores (0.413) than negative degree (0.405). Moreover, unlike negative degree—which must be compared to NP scores in absolute value—SHEEP separates the left and right members of congress directly. As discussed in Chapter 3, without taking the absolute value, the correlation between SHEEP and NP scores rises substantially to 0.698.

Congress 114: Node Extremism and Net Degree

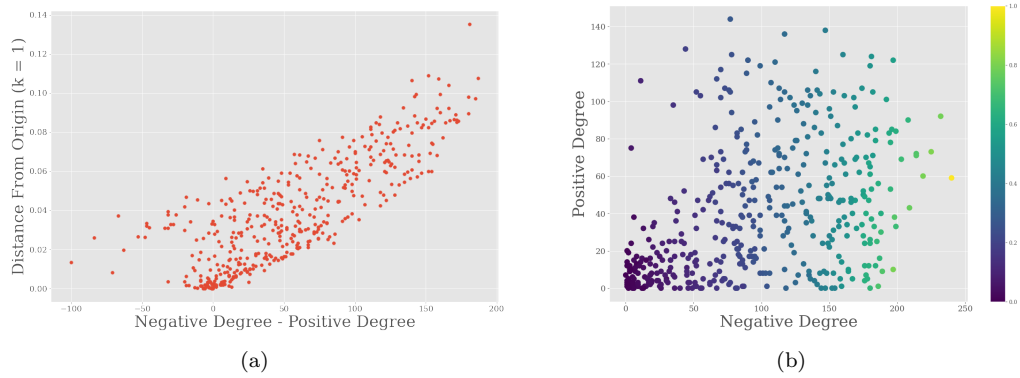


Figure A.6: **Relationship between SHEEP norm, positive and negative degree for Congress 114 graph.** a) Distance from the origin (in one dimensional SHEEP embedding) versus net degree. The KT correlation is weaker than on the random synthetic graphs (0.614), indicating that graph structure in addition to net degree plays a role in node "extremism". b) Positive degree versus negative degree, colour gives the norm in the one-dimensional embedding.

Appendix B

Menéame Data and Methods

B.1 Network Reciprocity

For the Menéame networks corresponding to the Russia-Ukraine War and Broad Politics topics, we quantify the reciprocal voting behaviour between users. For each pair of users with at least three votes in each direction, we take the number and mean of votes in each direction. These numbers are reported in Tables B.1- B.2, which illustrate that user voting behaviour is highly symmetric.

	$(-1.01, -0.899]$	$(-0.899, -0.3]$	$(-0.3, 0.3]$	$(0.3, 0.899]$	$(0.899, 1.01]$
$(-1.01, -0.899]$	116	2	2	1	0
$(-0.899, -0.3]$	2	2	1	0	1
$(-0.3, 0.3]$	2	1	0	1	1
$(0.3, 0.899]$	1	0	1	0	17
$(0.899, 1.01]$	0	1	1	17	2892

Table B.1: **Cross-tabulation of binned average vote scores (Topic: *Russia-Ukraine War*)**

	$(-1.01, -0.899]$	$(-0.899, -0.3]$	$(-0.3, 0.3]$	$(0.3, 0.899]$	$(0.899, 1.01]$
$(-1.01, -0.899]$	96	15	6	4	3
$(-0.899, -0.3]$	15	8	4	5	6
$(-0.3, 0.3]$	6	4	0	5	5
$(0.3, 0.899]$	4	5	5	24	96
$(0.899, 1.01]$	3	6	5	96	5060

Table B.2: **Cross-tabulation of binned average vote scores (Topic: *Broad Politics*)**

B.2 Topic Modelling

B.2.1 BERTopic algorithm

This section provides an overview of the BERTopic algorithm and the parameter choices that generated the topics underpinning the topic-specific networks used in the analysis. For a comprehensive description of the method, see [39]. The BERTopic procedure begins by converting the text corpus into embeddings, such that each document is represented as a vector in a high-dimensional space.

We employ SentenceTransformers [173], a variation of the BERT (Bidirectional Encoder Representations from Transformers) architecture specialized for creating sentence embeddings. As the corpus is in Spanish, we select the model called *paraphrase-spanish-distilroberta* which is pre-trained on Spanish textual data [1]. This results in a 768-dimensional vector for each document. Next, we perform dimensionality reduction via a Uniform Manifold Approximation and Projection (UMAP) projection [143]. This step reduces the high-dimensional embeddings while preserving both local and global structure in the data, as ensured by our parameter choices (see [39]). We select a $n = 75$ dimensional projection, to retain approximately 10% of the original feature space.

The reduced embeddings are clustered using Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm, setting the minimum cluster size to 50, to avoid highly granular topics [144]. We perform a sensitivity analysis to determine the parameter choices governing number of topics and outliers(see [39]).

For each resulting cluster, we identify the “topic” by extracting relevant keywords using the CountVectorizer from the *sci-kit learn* package, finding the most popular words for each topic [162]. To obtain coherent and informative topic keywords, we remove Spanish stopwords, and consider only words appearing more than 100 times. To find the most relevant word per topic, we measure term importance using class-based Term Frequency–Inverse Document Frequency (c-TF–IDF), which adapts the conventional TF–IDF weighting to the level of topics rather than individual documents. We employ the method *Maximal Marginal Relevance* to maximize the diversity of those keywords, so that each topic is identified by a maximally different set of keywords.

B.2.2 hSBM Topic Model steps

The hierarchical Stochastic Block Model (hSBM) topic modelling procedure is as follows. The text is pre-processed by removing Spanish stopwords using the NLTK stopword list, and creating tokens using the Spacy package [30, 105]. From the processed corpus, we construct a word–document bipartite graph, where nodes represent either words or documents, and weighted edges indicate the frequency of a word within a given document. We fit a community detection model to the bipartite network, which produces hierarchical subdivisions of words and documents into non-overlapping clusters. As with BERTopic, we apply the CountVectorizer and compute the c-TF–IDF. For each cluster, the ten terms with the highest c-TF–IDF scores are extracted and designated as representative keywords, allowing each cluster to be interpreted as a distinct topic.

B.2.3 Topic Modelling Comparison

In the analysis, we create robust topics by taking the overlap between the two topic modelling methods. Figure B.1 presents a confusion matrix that quantifies the degree of overlap between

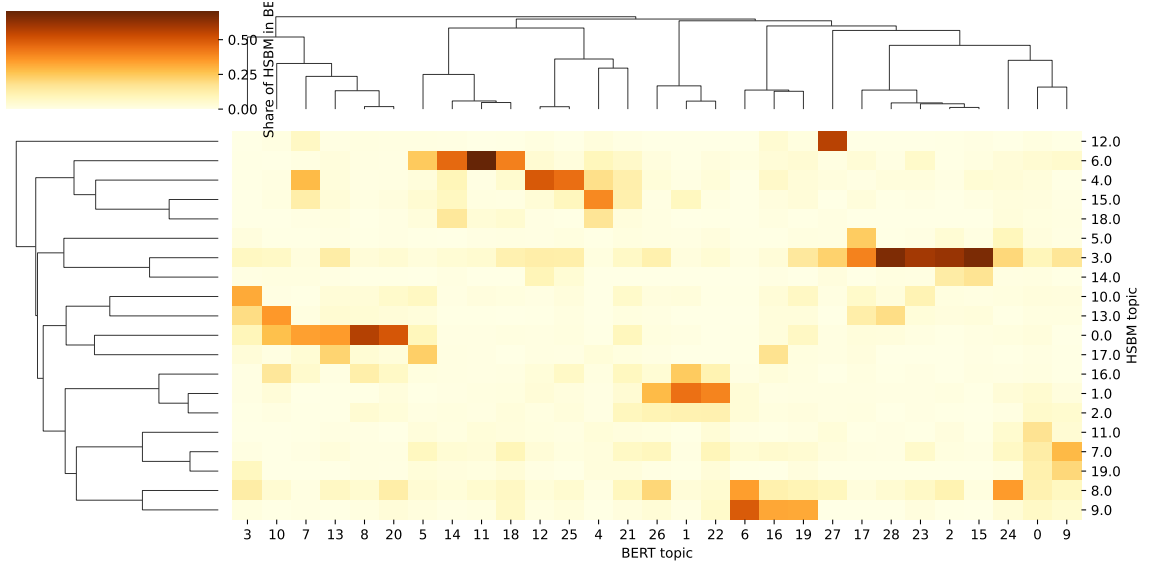


Figure B.1: **Confusion matrix with colours indicating the overlap between topics identified by BERTopic and hSBM.**

topics generated by the two methods. For a comprehensive list of the keywords associated with each topic under both approaches, see [39].

B.3 Ideology from Twitter

To derive ideology scores from Twitter data, we analyse tweets from a set of Spanish political parties and their associated accounts. The parties included are: PACMA, Más País, VOX, Izquierda Unida (IU), Podemos, Ciudadanos (CS), Partido Socialista Obrero Español (PSOE), and Partido Popular (PP). A detailed list of the accounts included is provided below:

- PACMA (@partidopacma) (Animalist Party Against Mistreatment of Animals): @lau.duart, @sanchezcastejon, @yolanda_morpe, @crisgarsalazar
- Más País (@maspais.es): @ierrejon, @Monica_Garcia_G, @EduardoFRub, @isabanes, @Rita_Maestre, @Equo, @P_GomezPerpinya, @compromis, @htejero_, @MasPais_Es, @MasMadrid_
- VOX (@vox.es): @Santi_ABASCAL, @Jorgebuxade, @Ortega.Smith, @ivanedlm, @monasterioR, @Igarraigavaz, @juan_ggallardo, @MeerRocio, @VOX_Congreso, @_patricia_rueda
- IU (@izquierdaunida): @agarzon, @sirarego, @EnriqueSantiago, @iuandalucia, @InmaNietoC, @joanmena, @iucyl, @Toni_Valero, @Congreso_Es, @ma_bustamante84, @Roser_Maestro, @iurioja, @elpce

- PODEMOS (@PODEMOS): @PabloIglesias, @IreneMontero, @ionebelarra, @isaserras, @Yolanda_Diaz_, @PabloEchenique, @PabloFdez, @MayoralRafa, @AleJacintoUrang, @Pam_Angela_, @MazelLilith, @SofCastanon, @VickyRosell, @nachoalvarez_, @juralde, @jessicaalbiach, @m_tere_perez, @JA_DelgadoRamos
- Ciudadanos (@ciudadanoscs): @InesArrimadas, @BalEdmundo, @carrizosacarlos, @GuillermoDiazCs, @begonavillacis, @FranciscoIgea, @CiutadansCs, @PatriciaGuaspB, @jordi_canyas, @MelisaRguezH, @Beatriz_Pino_, @Nmartinblanco
- PSOE (Spanish Socialist Workers' Party, @PSOE): @sanchezcastejon, @AdriLastra, @salvadorrilla, @mjmonteroc, @felipe_sicilia, @NadiaCalvino, @carmencalvo_, @abalosmeco, @isabelrguez, @Pilar_Alegria, @_JuanEspadas, @luistudanca, @santicl
- PP (Partido Popular, @ppopular): @pablocasado_, @TeoGarciaEgea, @cucagamarra, @NunezFeijoo, @IdiazAyuso, @Aglezterol, @AlmeidaPP_, @alferma1, @abeltran_ana, @alejandroTGN, @eliasbendodo, @anapastorjulian, @erodriguez_2019, @JuanMa_Moreno, @jaimeolano

We apply Correspondence Analysis (CA) to the interaction matrix between political parties and news outlets, obtaining an embedding for both the political parties and news outlets in the same embedding space. The first two dimensions are displayed in Fig. B.2. We interpret the first dimension as left-right ideology, and the second dimension, though not used in this paper, as a dimension representing a mainstream-radical divide. We also compare the ideological scores obtained for the news outlets against those provided by PoliticalWatch. Fig. B.3 compares the Twitter score against those of PoliticalWatch, revealing that the correlation is quite high.

Bibliography

- [1] Somosnlp-hackathon-2022/paraphrase-spanish-distilroberta · Hugging Face.
- [2] Emmanuel Abbe. Community detection and stochastic block models: recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.
- [3] Alan Abramowitz and Kyle Saunders. Why Can’t We All Just Get Along? The Reality of a Polarized America. *The Forum*, 3(2), July 2005.
- [4] Alan I. Abramowitz and Kyle L. Saunders. Is Polarization a Myth? *The Journal of Politics*, 70(2):542–555, April 2008.
- [5] B Devadas Acharya. Spectral criterion for cycle balance in networks. *Journal of Graph Theory*, 4(1):1–11, 1980.
- [6] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43, 2005.
- [7] Abdullah Almaatouq, Laura Radaelli, Alex Pentland, and Erez Shmueli. The role of reciprocity and directionality of friendship ties in promoting behavioral change. In *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings 9*, pages 33–41. Springer, 2016.
- [8] Claudio Altafini. Dynamics of opinion forming in structurally balanced social networks. *PloS one*, 7(6):e38135, 2012.
- [9] Pranay Anchuri and Malik Magdon-Ismail. Communities and balance in signed networks: A spectral approach. In *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 235–242. IEEE, 2012.
- [10] Ashton Anderson, Lucas Maystre, Ian Anderson, Rishabh Mehrotra, and Mounia Lalmas. Algorithmic effects on the diversity of consumption on spotify. In *Proceedings of the web conference 2020*, pages 2155–2165, 2020.

- [11] R. Kondor and J. Lafferty. Diffusion kernels on graphs and other discrete structures. In *International Conference on Machine Learning*, page 315–322, 2002.
- [12] Pablo Aragón, Vicenç Gómez, and Andreaks Kaltenbrunner. To thread or not to thread: The impact of conversation threading on online discussion. In *Proceedings of the International AAAI Conference on Web and social media*, volume 11, pages 12–21, 2017.
- [13] Sinan Aral, Lev Muchnik, and Arun Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009.
- [14] Sinan Aral and Christos Nicolaides. Exercise contagion in a global social network. *Nature communications*, 8(1):14753, 2017.
- [15] Samin Aref, Ly Dinh, Rezvaneh Rezapour, and Jana Diesner. Multilevel structural evaluation of signed directed social networks based on balance theory. *Scientific reports*, 10(1):15228, 2020.
- [16] Samin Aref and Zachary Neal. Detecting coalitions by optimally partitioning signed networks of political collaboration. 10(1):1506.
- [17] Samin Aref and Zachary P. Neal. Identifying hidden coalitions in the US House of Representatives by optimally partitioning signed networks based on generalized balance. 11(1):19939.
- [18] Samin Aref and Zachary P Neal. Identifying hidden coalitions in the us house of representatives by optimally partitioning signed networks based on generalized balance. *Scientific reports*, 11(1):19939, 2021.
- [19] Samin Aref and Mark C Wilson. Balance and frustration in signed networks. *Journal of Complex Networks*, 7(2):163–189, 2019.
- [20] Shazia’Ayn Babul, Karel Devriendt, and Renaud Lambiotte. Gromov centrality: A multiscale measure of network centrality using triangle inequality excess. *Physical Review E*, 106(3):034312, 2022.
- [21] Shazia’Ayn Babul, Yu Tian, and Renaud Lambiotte. Strong and weak random walks on signed networks. *npj Complexity*, 2(1):4, 2025.
- [22] Shazia’Ayn Babul and Renaud Lambiotte. Sheep, a signed hamiltonian eigenvector embedding for proximity. *Communications Physics*, 7(1):8, 2024.
- [23] Norman TJ Bailey et al. *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

- [24] Pablo Barberá. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis*, 23(1):76–91, January 2015.
- [25] Pablo Barberá. Social Media, Echo Chambers, and Political Polarization. In Joshua A. Tucker and Nathaniel Persily, editors, *Social Media and Democracy*, SSRC Anxieties of Democracy, pages 34–55. Cambridge University Press, Cambridge, 2020.
- [26] Pablo Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science*, 26(10):1531–1542, October 2015.
- [27] Ghazaleh Beigi, Jiliang Tang, and Huan Liu. Social Science-guided Feature Engineering: A Novel Approach to Signed Link Analysis. *ACM Transactions on Intelligent Systems and Technology*, 11(1):1–27, February 2020.
- [28] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in neural information processing systems*, 14, 2001.
- [29] J.P. Benzécri and L. Bellier. *L’analyse des données: Benzécri, J.-P. et al. L’analyse des correspondances*. L’analyse des données: leçons sur l’analyse factorielle et la reconnaissance des formes, et travaux du Laboratoire de statistique de l’Université de Paris VI. Dunod, 1973.
- [30] Steven Bird, Edward Loper, and Ewan Klein. *Natural Language Processing with Python*. O’Reilly Media Inc., 2009.
- [31] David M. Blei and John D. Lafferty. Topic Models. In *Text Mining*. Chapman and Hall/CRC.
- [32] E Bogomolny, O Bohigas, and C Schmit. Distance matrices and isometric embeddings. *arXiv preprint arXiv:0710.2063*, 2007.
- [33] Francesco Bonchi, Edoardo Galimberti, Aristides Gionis, Bruno Ordozgoiti, and Giancarlo Ruffo. Discovering polarized communities in signed networks. In *Proceedings of the 28th acm international conference on information and knowledge management*, pages 961–970, 2019.
- [34] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298, 2012.
- [35] Aaron Bramson, Patrick Grim, Daniel J. Singer, William J. Berger, Graham Sack, Steven Fisher, Carissa Flocken, and Bennett Holman. Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of Science*, 84(1):115–159, January 2017.

- [36] Ulrik Brandes, Daniel Fleischer, and Jürgen Lerner. Summarizing dynamic bipolar conflict structures. *IEEE Transactions on Visualization and Computer Graphics*, 12(6):1486–1499, 2006.
- [37] Ulrik Brandes and Jürgen Lerner. Visual analysis of controversy in user-generated encyclopedias. *Information visualization*, 7(1):34–48, 2008.
- [38] Jared C Bronski and Lee DeVille. Spectral theory for dynamics on graphs containing attractive and repulsive interactions. *SIAM Journal on Applied Mathematics*, 74(1):83–105, 2014.
- [39] E. Candellone, S. A. Babul, Ö. Togay, A. Bovet, and J. Garcia-Bernardo. Negative ties highlight hidden extremes in social media polarization. *Network Science*, 13:e10, 2025.
- [40] Jiayu Cao, Ying Fan, and Zengru Di. Frustration of signed networks: How does it affect the thermodynamic properties of a system? *arXiv preprint arXiv:1810.10481*, 2018.
- [41] Royce Carroll, Jeffrey B Lewis, James Lo, Keith T Poole, and Howard Rosenthal. Measuring bias and uncertainty in dw-nominate ideal point estimates via the parametric bootstrap. *Political analysis*, 17(3):261–275, 2009.
- [42] Dorwin Cartwright and Frank Harary. Structural balance: a generalization of heider’s theory. *Psychological review*, 63(5):277, 1956.
- [43] Giulia Cencetti, Diego Andrés Contreras, Marco Mancastroppa, and Alain Barrat. Distinguishing simple and complex contagion processes on networks. *Physical Review Letters*, 130(24):247401, 2023.
- [44] Damon Centola. The spread of behavior in an online social network experiment. *science*, 329(5996):1194–1197, 2010.
- [45] Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. *American journal of Sociology*, 113(3):702–734, 2007.
- [46] Jeff Cheeger. A lower bound for the smallest eigenvalue of the laplacian. *Problems in analysis*, 625(195-199):110, 1970.
- [47] Fenxiao Chen, Yun-Cheng Wang, Bin Wang, and C-C Jay Kuo. Graph representation learning: a survey. *APSIPA Transactions on Signal and Information Processing*, 9:e15, 2020.
- [48] Justin Cheng, Jon Kleinberg, Jure Leskovec, David Liben-Nowell, Karthik Subbian, Lada Adamic, et al. Do diffusion protocols govern cascade growth? In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12, 2018.

- [49] Kai-Yang Chiang, Joyce Jiyoung Whang, and Inderjit S Dhillon. Scalable clustering of signed networks using balance normalized cut. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 615–624, 2012.
- [50] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9):e2023301118, March 2021.
- [51] Joshua Clinton, Simon Jackman, and Douglas Rivers. The Statistical Analysis of Roll Call Data. *American Political Science Review*, 98(2):355–370, May 2004.
- [52] Michael Conover, Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer, and Alessandro Flammini. Political polarization on twitter. In *Proceedings of the international aaai conference on web and social media*, volume 5, pages 89–96, 2011.
- [53] Mihai Cucuringu, Peter Davies, Aldo Glielmo, and Hemant Tyagi. Sponge: A generalized eigenproblem for clustering signed networks. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 1088–1098. PMLR, 2019.
- [54] James A Davis. Clustering and structural balance in graphs. *Human relations*, 20(2):181–187, 1967.
- [55] Caterina De Bacco, Daniel B Larremore, and Cristopher Moore. A physical model for efficient ranking in networks. *Science advances*, 4(7):eaar8260, 2018.
- [56] EL Deacon. Climatic change in australia since 1880. *Australian Journal of Physics*, 6(2):209–218, 1953.
- [57] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, January 2016.
- [58] J. Delvenne, S. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *Proceedings of the National Academy of Sciences*, 107(29):12755–12760, 2010.
- [59] Tyler Derr, Yao Ma, and Jiliang Tang. Signed graph convolutional networks. In *2018 IEEE international conference on data mining (ICDM)*, pages 929–934. IEEE, 2018.
- [60] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, May 2019.
- [61] Michel Marie Deza, Monique Laurent, and R Weismantel. *Geometry of cuts and metrics*, volume 2. Springer, 1997.

- [62] Fernando Diaz-Diaz, Elena Candellone, Miguel A González-Casado, Emma Fraxanet, Antoine Vendeville, Irene Ferri, and Andreia Sofia Teixeira. Signed networks: theory, methods, and applications. *arXiv preprint arXiv:2511.17247*, 2025.
- [63] Paul DiMaggio, John Evans, and Bethany Bryson. Have Americans’ social attitudes become more polarized? *American Journal of Sociology*, 102(3):690–755, 1996.
- [64] Patrick Doreian and David Krackhardt. Pre-transitive balance mechanisms for signed networks. *Journal of Mathematical Sociology*, 25(1):43–67, 2001.
- [65] Patrick Doreian and Andrej Mrvar. Partitioning signed social networks. *Social Networks*, 31(1):1–11, 2009.
- [66] Ritabrata Dutta, Antonietta Mira, and Jukka-Pekka Onnela. Bayesian inference of spreading processes on networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 474(2215):20180129, 2018.
- [67] Jan Alyne Barbosa e Silva. An exploratory study about issue and attribute salience within the system of news promotion, menéame.
- [68] P. Eades. A heuristic for graph drawing. *Congressus Numerantium*, 42:149–160, 1984.
- [69] James M. Enelow and Melvin J. Hinich. *The Spatial Theory of Voting: An Introduction*. CUP Archive, April 1984.
- [70] Katharina Esau, Tariq Choucair, Samantha Vilkins, Sebastian F. K. Svegaard, Axel Bruns, Kate S. O’Connor-Farfan, and Carly Lubicz-Zaorski. Destructive polarization in digital communication contexts: A critical review and conceptual framework. *Information, Communication & Society*, 0(0):1–22.
- [71] Pouya Esmailian and Mahdi Jalili. Community Detection in Signed Networks: The Role of Negative ties in Different Scales. *Scientific Reports*, 5(1):14339, September 2015.
- [72] Max Falkenberg, Alessandro Galeazzi, Maddalena Torricelli, Niccolò Di Marco, Francesca Larosa, Madalina Sas, Amin Mekacher, Warren Pearce, Fabiana Zollo, Walter Quattrociocchi, and Andrea Baronchelli. Growing polarization around climate change on social media. *Nature Climate Change*, 12(12):1114–1121, December 2022.
- [73] Max Falkenberg, Fabiana Zollo, Walter Quattrociocchi, Jürgen Pfeffer, and Andrea Baronchelli. Affective and interactional polarization align across countries, November 2023.

- [74] Henrique Ferraz de Arruda, Felipe Maciel Cardoso, Guilherme Ferraz de Arruda, Alexis R. Hernández, Luciano da Fontoura Costa, and Yamir Moreno. Modelling how social network algorithms can influence opinion polarization. *Information Sciences*, 588:265–278, April 2022.
- [75] Clay Fink, Aurora Schmidt, Vladimir Barash, John Kelly, Christopher Cameron, and Michael Macy. Investigating the observability of complex contagion in empirical social networks. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, pages 121–130, 2016.
- [76] James Flamino, Alessandro Galeazzi, Stuart Feldman, Michael W. Macy, Brendan Cross, Zhenkun Zhou, Matteo Serafino, Alexandre Bovet, Hernán A. Makse, and Boleslaw K. Szymanski. Political polarization of news media and influencers on Twitter in the 2016 and 2020 US presidential elections. *Nature Human Behaviour*, 7(6):904–916, March 2023.
- [77] F. Fouss, A. Pirotte, J. Renders, and M. Saerens. Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 19(3):355–369, 2007.
- [78] François Fouss, Marco Saerens, and Masashi Shimbo. *Algorithms and models for network data and link analysis*. Cambridge University Press, 2016.
- [79] Alyson Fox, Thomas Manteuffel, and Geoffrey Sanders. Numerical methods for grebman’s expansion of signed graphs. *SIAM Journal on Scientific Computing*, 39(5):S945–S968, 2017.
- [80] Alyson Fox, Geoffrey Sanders, and Andrew Knyazev. Investigation of spectral clustering for signed graph matrix representations. In *2018 IEEE High Performance extreme Computing Conference (HPEC)*, pages 1–7. IEEE, 2018.
- [81] Thomas M. J. Fruchterman and Edward M. Reingold. Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11):1129–1164, 1991.
- [82] Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.
- [83] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Reducing controversy by connecting opposing views. In *Proceedings of the tenth ACM international conference on web search and data mining*, pages 81–90, 2017.
- [84] Alexander J Gates and Yong-Yeol Ahn. The impact of random models on clustering similarity. *Journal of Machine Learning Research*, 18(87):1–28, 2017.

- [85] Martin Gerlach, Tiago P. Peixoto, and Eduardo G. Altmann. A network approach to topic models. 4(7):eaaq1360.
- [86] M. Girvan and M. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [87] David F Gleich. Pagerank beyond the web. *siam REVIEW*, 57(3):321–363, 2015.
- [88] Vicenç Gómez, Hilbert J. Kappen, Nelly Litvak, and Andreas Kaltenbrunner. A likelihood-based framework for the analysis of discussion threads. *World Wide Web*, 16(5):645–675, November 2013.
- [89] Mark Granovetter. Threshold models of collective behavior. *American journal of sociology*, 83(6):1420–1443, 1978.
- [90] Michael Greenacre. *Correspondence analysis in practice*. chapman and hall/crc, 2017.
- [91] M.J. Greenacre. *Theory and Applications of Correspondence Analysis*. Academic Press, 1984.
- [92] Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*, 2022.
- [93] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.
- [94] Daniel Gruhl, Ramanathan Guha, David Liben-Nowell, and Andrew Tomkins. Information diffusion through blogspace. In *Proceedings of the 13th international conference on World Wide Web*, pages 491–501, 2004.
- [95] David Guthrie, Ben Allison, Wei Liu, Louise Guthrie, and Yorick Wilks. A closer look at skip-gram modelling. In *LREC*, volume 6, pages 1222–1225, 2006.
- [96] Per Hage and Frank Harary. Structural models in anthropology. (*No Title*), 1983.
- [97] Max Halford. Prince.
- [98] Robert A Hanneman and Mark Riddle. Introduction to social network methods, 2005.
- [99] Frank Harary. On the notion of balance of a signed graph. *Michigan Mathematical Journal*, 2(2):143–146, 1953.
- [100] Harmon, Joshua Roesslein, and Other Contributors. Tweepy. Zenodo, April 2023.

- [101] Xiaochen He, Haifeng Du, Xiangning Xu, and Wei Du. An energy function for computing structural balance in fully signed network. *IEEE Transactions on Computational Social Systems*, 7(3):696–708, 2020.
- [102] Fritz Heider. Attitudes and cognitive organization. *The Journal of psychology*, 21(1):107–112, 1946.
- [103] H. O. Hirschfeld. A connection between correlation and contingency. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(4):520–524, 1935.
- [104] Marilena Hohmann, Karel Devriendt, and Michele Coscia. Quantifying ideological polarization on a network using generalized Euclidean distance. *Science Advances*, 9(9):eabq2044, March 2023.
- [105] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. spaCy: Industrial-strength Natural Language Processing in Python. 2020.
- [106] Desislava Hristova, Antonio Lima, Renaud Lambiotte, Mariano Beguerisse-Díaz, et al. Link me baby one more time: Social music discovery on spotify. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 19, pages 146–159, 2025.
- [107] Zexi Huang, Arlei Silva, and Ambuj Singh. Pole: Polarized embedding for signed networks. *arXiv preprint arXiv:2110.09899*, 2021.
- [108] Ruben Interian, Ruslán G. Marzo, Isela Mendoza, and Celso C Ribeiro. Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods. *International Transactions in Operational Research*, 30(6):3122–3158, 2023.
- [109] Jinhong Jung, Woojeong Jin, Lee Sael, and U Kang. Personalized ranking in signed networks using signed random walk with restart. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 973–978, 2016.
- [110] Andreas Kaltenbrunner, Gustavo Gonzalez, Ricard Ruiz De Querol, and Yana Volkovich. Comparative analysis of articulated and behavioural social networks in a social news sharing website. *New Review of Hypermedia and Multimedia*, 17(3):243–266, 2011.
- [111] Tomihisa Kamada, Satoru Kawai, et al. An algorithm for drawing general undirected graphs. *Information processing letters*, 31(1):7–15, 1989.
- [112] Tatsuro Kawamoto, Masaki Ochi, and Teruyoshi Kobayashi. Consistency between ordering and clustering methods for graphs. *arXiv preprint arXiv:2208.12933*, 2022.
- [113] Anne-Marie Kermarrec and Afshin Moin. Energy models for drawing signed graphs. 2011.

- [114] Anna Keuchenijs, Petter Törnberg, and Justus Uitermark. Why it is important to consider negative ties when studying polarized debates: A signed network analysis of a dutch cultural controversy on twitter. *PloS one*, 16(8):e0256696, 2021.
- [115] Junghwan Kim, Haekyu Park, Ji-Eun Lee, and U Kang. Side: representation learning in signed directed networks. In *Proceedings of the 2018 world wide web conference*, pages 509–518, 2018.
- [116] Alec Kirkley, George T. Cantwell, and M. E. J. Newman. Balance in signed networks. *Physical Review E*, 99(1), jan 2019.
- [117] Andrew Knyazev. On spectral partitioning of signed graphs. In *2018 Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing*, pages 11–22. SIAM, 2018.
- [118] Adam Korniejczuk and Robert Ślepaczuk. Statistical arbitrage in multi-pair trading strategy based on graph clustering algorithms in us equities market. *arXiv preprint arXiv:2406.10695*, 2024.
- [119] Milan Kundera. *The Unbearable Lightness of Being*. Faber & Faber, 1999.
- [120] Jérôme Kunegis, Stephan Schmidt, Andreas Lommatzsch, Jürgen Lerner, Ernesto W De Luca, and Sahin Albayrak. Spectral analysis of signed graphs for clustering, prediction and visualization. In *Proceedings of the 2010 SIAM international conference on data mining*, pages 559–570. SIAM, 2010.
- [121] R. Lambiotte, J. Delvenne, and M. Barahona. Random walks, markov processes and the multiscale modular organization of complex networks. *IEEE Transactions on Network Science and Engineering*, 1(2):76–90, 2014.
- [122] Renaud Lambiotte. Network science: Course notes. Unpublished manuscript, University of Oxford, 2021. Course notes.
- [123] Renaud Lambiotte, Jean-Charles Delvenne, and Mauricio Barahona. Random walks, Markov processes and the multiscale modular organization of complex networks. *IEEE Transactions on Network Science and Engineering*, 1(2):76–90, 2014.
- [124] Renaud Lambiotte, Martin Rosvall, and Ingo Scholtes. From networks to optimal higher-order models of complex systems. *Nature physics*, 15(4):313–320, 2019.
- [125] Renaud Lambiotte and Michael T Schaub. *Modularity and dynamics on complex networks*. Cambridge University Press, 2021.
- [126] Amy N Langville and Carl D Meyer. *Google’s PageRank and beyond: The science of search engine rankings*. Princeton university press, 2006.

- [127] G. Lawler. *Random walk and the heat equation*. American Mathematical Society, 2010.
- [128] Yphtach Lelkes. Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, 80(S1):392–410, 2016.
- [129] Kristina Lerman and Rumi Ghosh. Information contagion: An empirical study of the spread of news on digg and twitter social networks. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 4, pages 90–97, 2010.
- [130] Jure Leskovec, Lada A Adamic, and Bernardo A Huberman. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)*, 1(1):5–es, 2007.
- [131] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *Proceedings of the 19th International Conference on World Wide Web*, pages 641–650, Raleigh North Carolina USA, April 2010. ACM.
- [132] 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.
- [133] Jure Leskovec, Kevin J Lang, Anirban Dasgupta, and Michael W Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009.
- [134] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.
- [135] Si-Yuan Liu, Jing Xiao, and Xiao-Ke Xu. Link prediction in signed social networks: From status theory to motif families. *IEEE Transactions on Network Science and Engineering*, 7(3):1724–1735, 2019.
- [136] Isabelle Lorge, Li Zhang, Xiaowen Dong, and Janet B Pierrehumbert. Stentconv: Predicting disagreement with stance detection and a signed graph convolutional network. *arXiv preprint arXiv:2403.15885*, 2024.
- [137] Regina Marchi. With facebook, blogs, and fake news, teens reject journalistic “objectivity”. *Journal of communication inquiry*, 36(3):246–262, 2012.
- [138] Mike J Mason, Guoping Fan, Kathrin Plath, Qing Zhou, and Steve Horvath. Signed weighted gene co-expression network analysis of transcriptional regulation in murine embryonic stem cells. *BMC genomics*, 10(1):327, 2009.

- [139] N. Masuda, M. Porter, and R. Lambiotte. Random walks and diffusion on networks. *Phys. Rep.*, 716–717:1–58, 2017.
- [140] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics reports*, 716:1–58, 2017.
- [141] Naoki Masuda, Mason A. Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics Reports*, 716–717:1–58, November 2017.
- [142] Heather Mattie and Jukka-Pekka Onnela. Edge overlap in weighted and directed social networks. *Network Science*, 9(2):179–193, 2021.
- [143] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints*, February 2018.
- [144] Leland McInnes, John Healy, and Steve Astels. hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11), mar 2017.
- [145] Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444, 2001.
- [146] Sofia Medina, Shazia’Ayn Babul, Timothy LaRock, Rohit Sahasrabuddhe, Renaud Lambiotte, and Nicola Pedreschi. Detection of anomalous spatio-temporal patterns of app traffic in response to catastrophic events. *EPJ Data Science*, 14(1):35, 2025.
- [147] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- [148] Lillio Mok, Samuel F Way, Lucas Maystre, and Ashton Anderson. The dynamics of exploration on spotify. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 663–674, 2022.
- [149] Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, and Sune Lehmann. Evidence of complex contagion of information in social media: An experiment using twitter bots. *PloS one*, 12(9):e0184148, 2017.
- [150] James Moody and Peter J. Mucha. Portrait of Political Party Polarization. *Network Science*, 1(1):119–121, April 2013.
- [151] Zachary P. Neal. A sign of the times? Weak and strong polarization in the US Congress, 1973–2016. *Social networks*, 60:103–112, 2020.
- [152] M. Newman. *Networks: An Introduction*. Oxford University Press, Oxford, UK, 2010.

- [153] M. E. J. Newman. Resource letter cs-1: Complex systems. *American Journal of Physics*, 79(8):800–810, July 2011.
- [154] Mark Newman. *Networks*. Oxford university press, 2018.
- [155] Mark EJ Newman. Spread of epidemic disease on networks. *Physical review E*, 66(1):016128, 2002.
- [156] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.
- [157] Timothy P Nokken and Keith T Poole. Congressional party defection in american history. *Legislative Studies Quarterly*, 29(4):545–568, 2004.
- [158] Anastasios Noulas, Yasin Acikmese, Charles QC LI, Milan Y Patel, Shazia Ayn Babul, Ronald C Cohen, Renaud Lambiotte, and Marta C Gonzalez. Modeling urban air quality using taxis as sensors. *arXiv preprint arXiv:2506.11720*, 2025.
- [159] Shira Offer. Negative Social Ties: Prevalence and Consequences. *Annual Review of Sociology*, 47(Volume 47, 2021):177–196, July 2021.
- [160] Li Ou, Yaoping Hou, and Zhuang Xiong. The net laplacian spectra of signed complete graphs. *Contemporary Mathematics*, pages 409–417, 2021.
- [161] J. Pan, H. Yang, C. Faloutsos, and P. Duygulu. Automatic multimedia cross-modal correlation discovery. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 653–658, New York, NY, USA, 2004. Association for Computing Machinery.
- [162] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [163] Tiago P. Peixoto. Hierarchical block structures and high-resolution model selection in large networks. *Phys. Rev. X*, 4:011047, Mar 2014.
- [164] Antonio F. Peralta, Pedro Ramaciotti, János Kertész, and Gerardo Iñiguez. Multidimensional political polarization in online social networks. *Phys. Rev. Res.*, 6:013170, Feb 2024.
- [165] B. Perozzi, R. Al-Rfou, and S. Skiena. DeepWalk: online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710, 2014.

- [166] Keith T. Poole. *Spatial Models of Parliamentary Voting*. Analytical Methods for Social Research. Cambridge University Press, Cambridge ; New York, 2005.
- [167] Keith T. Poole and Howard Rosenthal. A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, 29(2):357–384, 1985.
- [168] John Pougué-Biyong, Akshay Gupta, Aria Haghighi, and Ahmed El-Kishky. Learning stance embeddings from signed social graphs. In *Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining*, pages 177–185, 2023.
- [169] John Pougué-Biyong, Valentina Semanova, Alexandre Matton, Rachel Han, Aerin Kim, Renaud Lambiotte, and Doyné Farmer. Debagreement: A comment-reply dataset for (dis) agreement detection in online debates. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.
- [170] Tai Qin and Karl Rohe. Regularized spectral clustering under the degree-corrected stochastic blockmodel. *Advances in neural information processing systems*, 26, 2013.
- [171] Jiezhong Qiu, Yuxiao Dong, Hao Ma, Jian Li, Kuansan Wang, and Jie Tang. Network embedding as matrix factorization: Unifying deepwalk, line, pte, and node2vec. In *Proceedings of the eleventh ACM international conference on web search and data mining*, pages 459–467, 2018.
- [172] Kenneth E Read. Cultures of the central highlands, new guinea. *Southwestern Journal of Anthropology*, pages 1–43, 1954.
- [173] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.
- [174] Manoel Horta Ribeiro, Pedro H Calais, Virgílio AF Almeida, and Wagner Meira Jr. ” everything i disagree with is# fakenews”: Correlating political polarization and spread of misinformation. *arXiv preprint arXiv:1706.05924*, 2017.
- [175] Matthew Richardson, Rakesh Agrawal, and Pedro Domingos. Trust management for the semantic web. In *The Semantic Web-ISWC 2003: Second International Semantic Web Conference, Sanibel Island, FL, USA, October 20-23, 2003. Proceedings 2*, pages 351–368. Springer, 2003.
- [176] Dora P Rosati, Matthew H Woolhouse, Benjamin M Bolker, and David JD Earn. Modelling song popularity as a contagious process. *Proceedings of the Royal Society A*, 477(2253):20210457, 2021.

- [177] Martin Rosvall and Carl T Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the national academy of sciences*, 105(4):1118–1123, 2008.
- [178] Martin Rosvall, Jean-Charles Delvenne, Michael T Schaub, and Renaud Lambiotte. Different approaches to community detection. *Advances in network clustering and blockmodeling*, pages 105–119, 2019.
- [179] Majid Saberi, Reza Khosrowabadi, Ali Khatibi, Bratislav Misic, and Gholamreza Jafari. Topological impact of negative links on the stability of resting-state brain network. *Scientific reports*, 11(1):2176, 2021.
- [180] M. Saerens, F. Fouss, L. Yen, and P. Dupont. The principal components analysis of a graph, and its relationships to spectral clustering. In J. Boulicaut, F. Esposito, F. Giannotti, and D. Pedreschi, editors, *Machine Learning: ECML 2004*, pages 371–383, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.
- [181] Ali Salloum, Ted Hsuan Yun Chen, and Mikko Kivelä. Separating Polarization from Noise: Comparison and Normalization of Structural Polarization Measures, December 2021.
- [182] Samuel Franklin Sampson. *A novitiate in a period of change: An experimental and case study of social relationships*. Cornell University, 1968.
- [183] M. Schaub, J. Delvenne, S. Yaliraki, and M. Barahona. Markov dynamics as a zooming lens for multiscale community detection: Non clique-like communities and the field-of-view limit. *PLOS ONE*, 7(2):1–11, 02 2012.
- [184] Cosma Rohilla Shalizi and Andrew C Thomas. Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*, 40(2):211–239, 2011.
- [185] Guodong Shi, Claudio Altafini, and John S Baras. Dynamics over signed networks. *SIAM Review*, 61(2):229–257, 2019.
- [186] A. Smola and R. Kondor. Kernels and regularization on graphs. In B. Schölkopf and M. Warmuth, editors, *Learning Theory and Kernel Machines*, pages 144–158, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.
- [187] Spotify. About spotify. <https://newsroom.spotify.com/company-info>, 2023. Accessed: 2023-12-05.
- [188] Eric Sun, Itamar Rosenn, Cameron Marlow, and Thomas Lento. Gesundheit! modeling contagion through facebook news feed. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 3, pages 146–153, 2009.

- [189] Michael Szell, Renaud Lambiotte, and Stefan Thurner. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences*, 107(31):13636–13641, 2010.
- [190] Szymon Talaga, Massimo Stella, Trevor James Swanson, and Andreia Sofia Teixeira. Polarization and multiscale structural balance in signed networks. *Communications Physics*, 6(1):1–15, December 2023.
- [191] Jiliang Tang, Yi Chang, Charu Aggarwal, and Huan Liu. A survey of signed network mining in social media. *ACM Comput. Surv.*, 49(3), aug 2016.
- [192] John Ternovski and Taha Yasseri. Social complex contagion in music listenership: A natural experiment with 1.3 million participants. *Social Networks*, 61:144–152, 2020.
- [193] Yu Tian and Renaud Lambiotte. Spreading and structural balance on signed networks. *arXiv preprint arXiv:2212.10158*, 2022.
- [194] Yu Tian and Renaud Lambiotte. Spreading and structural balance on signed networks. *SIAM Journal on Applied Dynamical Systems*, 23(1):50–80, 2024.
- [195] H. Tong, C. Faloutsos, and Y. Koren. Fast direction-aware proximity for graph mining. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 747–756, New York, NY, USA, 2007. Association for Computing Machinery.
- [196] H. Tong, C. Faloutsos, and J.-Y. Pan. Randomwalk with restart: Fast solutions and applications. *Knowl. Inf. Syst.*, 14(3):327–346, 2008.
- [197] Petter Törnberg. How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences*, 119(42):e2207159119, 2022.
- [198] Vincent Traag, Patrick Doreian, and Andrej Mrvar. Partitioning signed networks. *Advances in network clustering and blockmodeling*, pages 225–249, 2019.
- [199] Vincent A Traag. Complex contagion of campaign donations. *PloS one*, 11(4):e0153539, 2016.
- [200] Vincent A Traag and Jeroen Bruggeman. Community detection in networks with positive and negative links. *Physical Review E*, 80(3):036115, 2009.
- [201] Ruo-Chun Tzeng, Bruno Ordozgoiti, and Aristides Gionis. Discovering conflicting groups in signed networks. *Advances in Neural Information Processing Systems*, 33:10974–10985, 2020.
- [202] Tyler J VanderWeele and Weihua An. Social networks and causal inference. *Handbook of causal analysis for social research*, pages 353–374, 2013.

- [203] Pablo Villegas, Tommaso Gili, Guido Caldarelli, and Andrea Gabrielli. Laplacian renormalization group for heterogeneous networks. *Nature Physics*, 19(3):445–450, 2023.
- [204] Qihan Wang, Anique Tahir, Zeyad Alghamdi, and Huan Liu. Exploring musical, lyrical, and network dimensions of music sharing among depression individuals, 2023.
- [205] Andrew Scott Waugh, Liuyi Pei, James H. Fowler, Peter J. Mucha, and Mason A. Porter. Party Polarization in Congress: A Network Science Approach. 2009.
- [206] Han Xiao, Bruno Ordozgoiti, and Aristides Gionis. Searching for polarization in signed graphs: a local spectral approach. In *Proceedings of The Web Conference 2020*, pages 362–372, 2020.
- [207] Moran Yarchi, Christian Baden, and Neta Kligler-Vilenchik. Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, 38:1–42, July 2020.
- [208] Xingchen Yu and Abel Rodríguez. Spatial voting models in circular spaces: A case study of the U.S. House of Representatives. *The Annals of Applied Statistics*, 15(4):1897–1922, December 2021.
- [209] Sixiao Zhang, Hongxu Chen, Xiao Ming, Lizhen Cui, Hongzhi Yin, and Guandong Xu. Where are we in embedding spaces? In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 2223–2231, 2021.
- [210] Yin Zhang, Rong Jin, and Zhi-Hua Zhou. Understanding bag-of-words model: a statistical framework. *International journal of machine learning and cybernetics*, 1(1):43–52, 2010.
- [211] Zeyu Zhang, Peiyao Zhao, Xin Li, Jiamou Liu, Xinrui Zhang, Junjie Huang, and Xiaofeng Zhu. Signed graph representation learning: A survey. *arXiv preprint arXiv:2402.15980*, 2024.