Title
Strong Consensus-Seeking in a Model of Social Consensus Formation

Author
Oakley, William Garrett

Publication Date
2020

UNIVERSITY OF CALIFORNIA

Los Angeles

Strong Consensus-Seeking in a Model of Social Consensus Formation

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Mathematics

by

William Garrett Oakley

2020

ABSTRACT OF THE DISSERTATION

Strong Consensus-Seeking in a Model of Social Consensus Formation

by

William Garrett Oakley
Doctor of Philosophy in Mathematics
University of California, Los Angeles, 2020
Professor Mason Alexander Porter, Chair

In experiments that date back to at least the 1930s, psychologists have observed the tendency of subjects to change their beliefs to match those of others. From this observation, researchers have hypothesized that the beliefs of populations should move towards consensus. However, despite marked similarities between the beliefs of individuals in social groups, large-scale consensus is often not observed. This disparity between psychological and sociological observations, called the "community-cleavage problem", is one of many problems that have been examined by researchers through the development and analysis of mathematical models of belief changes. These mathematical models, often in the form of models for opinion dynamics or social influence, attract much interest from researchers for explaining how the behaviors of individuals combine to produce sociological phenomena.

In this dissertation, we draw from psychological studies to develop a consensus model of social influence. Our model is distinct from most previous models in that (1) it includes both direct and indirect social influence and (2) it incorporates group interactions of any number of individuals. We study the convergence time to consensus of our model under various assumptions on how groups form. In doing so, we make some hypotheses about how the size and composition of groups of interacting individuals affect the speed of convergence to consensus.

In various applications—such as optimizing both marketing and political campaigns and disrupting the formation of echo chambers in online social networks—researchers have used mathematical models to develop strategies for influencing opinions. These strategies prescribe how to modify factors like individuals' opinions, susceptibilities to influence, and social connections. In this dissertation, we apply our model of opinion dynamics to conflict resolution, and we propose group-selection strategies for choosing groups of individuals to meet for the purpose of accelerating convergence to consensus. We show through analytical estimates, simulations, and examples that our strategies generate an improvement in convergence time over random choices of groups.

The dissertation of William Garrett Oakley is approved.

Christopher R. Anderson

Peter John Lamberson

Marcus Leigh Roper

Mason Alexander Porter, Committee Chair

University of California, Los Angeles

2020

To my Mom and Dad.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The research in this thesis was produced with funding from the Army Research Office (through MURI grant W911NF1810208).

I am indebted to my advisor, Mason Porter. You have been a constant source of professional and personal support. Thank you, specifically, for the many hours we have spent in your office discussing research; for finding funding opportunities for me; for keeping me on schedule to graduate; for encouraging me to do the theoretical work that I wanted to pursue during my PhD; for instructing me in technical writing through your countless comments on my drafts; and for supporting me when times were hard (like during the ongoing COVID-19 outbreak). So much of what I am today is a result of our interactions over the last four years. This thesis would not exist without your help.

I want to thank Jeff Brantingham for making this research possible through his encouragement and support. Your countless research ideas have been a source of inspiration throughout the past two years. Thank you for believing in this research.

I want to thank David Kempe for steering the research in this thesis towards more fertile waters than where it began. You have been an invaluable source of feedback. I do not want to think about what the research in this thesis would have looked like without your guidance.

I want to thank Yacoub Kureh for our useful discussions about the research in this thesis.

I want to thank Chris Anderson for his many seminars that I have had the privilege of participating in. I am fortunate to have experienced the atmosphere of excitement for math research that you foster. Also, I owe my knowledge of C++ to you.

I want to thank the chair of my committee, Mason Porter, and the members of my committee, Chris Anderson, PJ Lamberson and Marcus Roper, for their time and their feedback on my work.

I want to thank the many excellent mathematics professors at UCLA to whom I owe much of my knowledge of mathematics: Chris Anderson, Inwon Kim, Alan Laub, Mason Porter, Marcus Roper, Joey Teran, Luminita Vese, Wotao Yin, and others.

I want to thank Maida Bassili and Martha Contreras for their kindness and for their guidance and support on administrative matters.

I want to thank my Masters advisor, Lorena Bociu, and my undergraduate advisors, Robert H. Martin and Sandra O. Paur. I am a mathematician now because of your guidance then. My fondest memories of mathematics are from you.

I want to thank the many excellent professors I had the pleasure of learning from at NC State University—Chuck Wessell, Stephen Reynolds, Tomas Phillips, Marina Bykova, and Molly Pryzwansky, among many others.

I want to thank my many colleagues and friends that I had the good luck of meeting while at UCLA. Thank you Bonsoon Lin, Yacoub Kureh, and Jean-Michel Maldague for the many treasured memories that we've created together over the last six years. Thank you ChiChi Huang and Shyr-Shea Cheng for your close friendships. Thank you Zhiyun Lu for the fun times we've had together. Thank you Assaf Shani for being a great roommate and an even better friend. Thank you Peter Cheng for your constant friendship and for motivating me to learn more C++ and to become fitter. Thank you Jon Seigel for your constant friendship and for motivating me to become smarter. Thank you Mike Lindstrom for your friendship and for motivating me to learn more C++ through your rigorous classes that I had the pleasure to TA. Thank you Robert Hannah for your friendship and your career guidance. I also want to thank Mert Besken, Zach Boyd, Xenia Deviatkina-Loh, Brent Edmunds, Michelle Feng, Alex Lin, Denali Molitor, Maria Ntekoume, Minh Pham, Andre Pradhana, and Stephanie Wang.

I want to thank my many colleagues and friends that I was fortunate to meet while at NC State University. Thank you Marshall Markham being a close friend. Thank you Rohit Sivaprasad for your career help and the times I've been fortunate enough to be in your company over the last several years. Thank you Joe Murray for the good times we had together. The friends I made in the math department—Joey Arthur, Christian Chapman, Kelly Chapman, John Gray, Tyler Maltba, Drew Marquis, Liz Morris, Faye Paysley, David Rogers, Georgy Scholten, and Paulina Spencer—are some of the kindest, liveliest, and most interesting people that I know. I cherish my memories of time spent with you all.

I want to thank Dave Andrews for making my first software-engineering internship awesome and for his continued interest in my career.

I want to thank Aunt Joanna, Uncle Buck, Jackson, and Katherine for opening up their home and their hearts to me while growing up. I want to thank my late grandmother, Kay, and my late grandfather, Dan, for their love, support, and for my innumerable cherished memories of their farmhouse. I want to thank my late grandmother, Hilma, and my late grandfather, Clarence, for their hard work in life that sprouted into opportunities for their children and grandchildren. Thank you Aunt Connie, Uncle John, Anna, and Daniel; and thank you Aunt Amy, Uncle Ed, Karen, and Jeff. I am lucky to have y'all as family.

I want to thank Jessica for giving my life meaning and happiness. Thank you for being a source of support, encouragement, and love while I was doing research, searching for a job, and writing my thesis. It's all worth it because of you.

I want to thank my older brothers, Steven and James. You were a constant source of inspiration and motivation for me while growing up. The bond we have cannot be broken.

Lastly, I want to thank my Mom and Dad. Mom, with your overflowing love, your warmth, your profound understanding of people, your nurturance, your endless energy, your selflessness and self-sacrifice, and Dad, with your hard work, your love of knowledge, your prudence, your integrity, your conviction, your selflessness and self-sacrifice: you are my examples of who to be and how to live. You instilled in me a love of learning, a confidence in myself, and an introspectiveness that made this PhD possible. You shaped my world and set me out on my way through life. Thank you.

VITA

2010–2013	North Carolina State University - Bachelors in Mathematics
2013–2014	North Carolina State University - Masters in Mathematics
2014–2020	UCLA - PhD Student in Applied Mathematics
2015–2015	Verizon Digital Media Services - Software Engineering Intern
2020–	Applied Intuition - Software Engineer

LIST OF NOTATION

$\mathbb{R}$	Real numbers
$\mathbb{Z}$	Integers
$\mathbb{Z}_{\geq 0}$	Non-negative integers
$\mathbb{Z}_{>0}$	Positive integers
G	Undirected graph
V	Node set
E	Edge set
$\{v, w\}$	Undirected edge between nodes v and w
$\|A\|$	The cardinality of a set A
$n := \|V\|$	Number of nodes in G
$m := \|E\|$	Number of edges in G
$\mathbf{A}$	Adjacency matrix of G
$\mathbf{L}$	Un-normalized graph Laplacian of G
$\lambda_k(\mathbf{M})$	k^{th} smallest eigenvalue of a matrix $\mathbf{M}$
$\mathbb{1}_A$	Indicator function of a set A
$\mathbf{1}_n$	Vector of n ones
$\mathbf{1}_A$	Vector defined by $(\mathbf{1}_A)_i = (\mathbb{1}_A)(i)$
$\bar{y} := \frac{1}{n}\mathbf{y}^T\mathbf{1}_n$	Mean of the components of a vector $\mathbf{y}$
$\bar{y}^A := \frac{1}{\lvert A\rvert}\mathbf{y}^T\mathbf{1}_A$	Mean of those components of a vector $\mathbf{y}$ in a set A
t	Discrete time step
S	Meeting group
G^A	Graph G with edges between nodes in A added
$\mathbf{A}^A$	Adjacency matrix of graph G^A
$\mathbf{L}^A$	Un-normalized Laplacian matrix of graph G^A
$\ell := \|S\|$	Number of nodes in a meeting group S
$e(\mathbf{y}) := \lVert\mathbf{y} - \bar{y}\mathbf{1}_n\rVert_2^2$	Distance from consensus of a vector $\mathbf{y}$

$\imath_{\text{step}}(\mathbf{x}, Z)$	Per-step convergence factor of $\mathbf{x}$ and Z
$r_{\text{step}}(Z)$	Per-step convergence factor of Z
$T(\epsilon, Z)$	The ϵ-convergence time of Z
U_ℓ	The uniform random variable on the ℓ-node sets
U_d	The uniform random variable on the dyads
$s := \frac{\ell}{n}$	Relative group size
$\alpha := \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{ne(\mathbf{x})}$	Relative disagreement
$k_v := \sum_{w \in V} A_{vw}$	Degree of node $v \in V$
$\mathbf{x}(t)$	Belief vector at time t
$\mathbf{x}_0$	Vector of initial beliefs
S^t	Meeting group at time t
$\beta(\mathbf{x}, S)$	Optimal degree of consensus as a function of $\mathbf{x}$ and S
$\beta^*(t) := \beta(\mathbf{x}(t), S^t)$	Optimal degree of consensus as a function of t
$C_1^I(\mathbf{x}, S)$	In-group cost of meeting group S
$C_1^O(\mathbf{x}, S)$	Out-group cost of meeting group S
$C_1(\mathbf{x}, S) := C_1^I(\mathbf{x}, S) + C_1^O(\mathbf{x}, S)$	Combined in-group and out-group costs

CHAPTER 1

Introduction

The stage for the modern study of opinion dynamics was set by psychologist John R. P. French in [Fre56]. He introduced a mathematical model of social influence to "integrate previous [empirical] findings into a logically consistent theory from which one can derive testable hypotheses for future research" [Fre56]. He built his model from three constituent sub-models: (1) a model for how individuals change their opinions during interpersonal interaction, (2) a model of influence relationships between group members, and (3) a model for the occurence of interpersonal interactions. By fixing sub-model 1, varying sub-models 2 and 3, and observing changes in the steady state of his model, he hypothesized how the final opinions of individuals in a group depend on social structure and communication patterns. This strategy of studying variation in a model's steady state in response to variations in models of social structure and communication is still the primary method in use to examine the dynamics of opinions in social networks [FMF17, PT17, PT18].

Unfortunately, the above approach for studying social influence has a major limitation. Its focus on a steady state as representative of the final opinions of a group limits the testability of the strategy's resulting hypotheses. A steady state is an asymptotic state of a mathematical model, and the time that is required to approach a steady state can be so large that a mathematical prediction that a system reaches a steady state has limited real-world applicability. Accordingly, the equating of a steady state with the terminal opinions of a group requires justification through consideration of the convergence time to a steady state.

Examining the time that is required for the termination of an algorithm is one of the primary enterprises of computer science [Knu97],[1] but to our knowledge the study of the convergence time of social-influence models began recently (e.g. in [GS14, PTC17, PT18, MVP18, PFT19b, PFT19a]). In this dissertation, we continue this recent work on convergence times in opinion dynamics by analytically studying the convergence time of a new model of social influence. In doing so, we obtain hypotheses that are, in principle, empirically testable. In the following four sections, we elaborate on this and other conceptual contributions of our work.

Convergence-Time Analysis

In chapters 6, 7, and 8, we study how the convergence time of our model depends on social-network structure and model parameters. In particular, we show that the convergence time is linear in the number of nodes for some graphs, and for other graphs, we present evidence that the convergence time is at best quadratic. The difference between a linear and a quadratic convergence time can be the difference between an empirically observable convergence to a steady state and an empirically unobservable convergence. Therefore, in applications, the distinction between linear and quadratic convergence is critical.

As an example of how one can apply this distinction in convergence times to a sociological problem, we state the oft-quoted remark by Axelrod, "if people tend to become more alike in their beliefs, attitudes, and behavior when they interact, why do not all such differences eventually disappear?" [Axe97]. Axelrod's question of how consensus-seeking behavior in individuals can result in opinion diversity has been termed the "community-cleavage problem" [Fri15, PT17, PT18]. To address this problem, researchers have developed models of social influence with steady states that correspond to states other than a consensus [FMF17]. In contrast to this line of work, we argue that in some social-influence processes, consensus is unobserved not because of the model's asymptotic behavior, but because the convergence

[1]See Section 2.5 for a partial review of convergence-time analyses of some algorithms that are similar to social-influence models.

time to consensus is too large.

A Model of Opinion Changes that Incorporates Indirect Social Influence

Models of social influence usually assume that either (1) all individuals in a social network update their beliefs simultaneously or (2) only a subset of individuals, usually a single node or dyad (i.e. an adjacent pairs of nodes), update their beliefs (see [FMF17]).[2] One attractive feature of approach 2 in comparison to approach 1 is that the former more closely models the interpersonal interactions, such as conversations or meetings, within a group. This point was made by Friedkin and Johnsen in [FJ90], "interpersonal influences do not occur in the simultaneous way... and there are more or less complex sequences of interpersonal influences in the group".

Models that take approach 2 must account both for social influence by individuals who are directly involved in an interpersonal interaction and for social influence from people who are not directly involved. Indeed, social influence works both directly through interpersonal interactions and indirectly through group norms and expectations, imagined interactions, and many other mechanisms [FW16]. Although the importance of indirect influence has been acknowledged in the psychological literature at least since the work of Leon Festinger (e.g. [Fes50]), existing models of social influence that take approach 2 do not account for indirect social influence.

In this dissertation, we introduce a model of opinion dynamics that includes the effect of indirect social influence. In our model, individuals seek consensus not only among those with whom they are currently interacting but also with all of their social connections. We call this *strong consensus-seeking.*

[2]See Section 2.3 for examples of models that make these assumptions.

Group Belief Updates

Direct interpersonal interactions are a key medium for the transmission of social influence [Bur01]. Most research in the field of opinion dynamics that has attempted to model interpersonal interactions has focused on pairwise interactions [AO11, FMF17]. However, the problem of modeling higher-order interactions (e.g. a conversation between three or more individuals) is attracting increasing interest [BCI20]. In this dissertation, we study higher-order interactions by introducing a model of individual opinion changes in which interpersonal interactions can occur between any number of individuals.

Conflict Resolution

An important question in the field of opinion dynamics is how one can influence opinions to push them toward a specified state. This has been explored in applications to marketing and politics, such as through investigations of how to optimally use resources to change individuals' opinions or susceptibilities to maximize adoption of a product or support for a candidate [KKT03, HKK14, BBC14, LFW18, AKP18]. Additionally, the question of which individuals should be encouraged to become friends to reduce opinion clustering (e.g. in so-called echo chambers [GBK09, Gar09, FGR16, QSS16, BW17]) has been studied in applications to online social networks [GDG17, WDL18, DDM17]. However, in the conflict-resolution literature, meetings between individuals has been studied as the primary medium for changing individuals' opinions rather than through resource expenditure or relationship modification [Bur90, Kel10, Bre84, BBC92, Pet98, CWH06].

In this dissertation, we study the effect of a meeting's size and composition on the speed of progress towards consensus. A similar research thrust to ours is the study of opinion-dynamics models with polarized steady states, in which opinions are clustered around certain belief values.[3] These opinion-dynamics models with polarized steady states have the poten-

[3]See Section 2.3 for examples of these models.

tial to shed light on why conflicts develop by examining why and how beliefs of individuals move towards different extremes. However, compared to past efforts, which focused on why people move apart, we study how to bring people together.

1.1 Outline of the Dissertation

The rest of this dissertation is organized as follows.

Chapter 2

We review the literature of several fields from which we draw inspiration. First, in Section 2.1, we summarize various psychological experiments on both direct and indirect social influence that inform the modeling of social influence. In Section 2.2, we consider sociological factors that are important when modeling social structure for the purpose of studying social influence. In Section 2.3, we discuss some existing models and results in the study of social influence. We then summarize some models from the field of group decision-making in Section 2.4. Lastly, in Section 2.5, we discuss related research from computer science.

Chapter 3

We present our model of social influence. The model consists of group belief updates that are similar in form to those in the models of French, DeGroot, and Friedkin and Johnsen [Fre56, DeG74, FJ99]. However, in our model, the updated beliefs of individuals in a meeting depend on the beliefs of individuals who are not part of that meeting. We intend this feature of the model to capture indirect social influence, a phenomenon that has been examined in psychological studies (see Section 2.1).

Chapter 4

We establish some properties of our model, and we provide examples to help explain our model's behavior. In Propositions 2 and 3, we write our model in a form that is more amenable to analysis than the form that we present in Chapter 3. We end the chapter by considering several examples.

Chapter 5

We describe our general strategy for proving some of the converence results for our model. In Proposition 11, we derive a recurrence relation for the change in distance from consensus of individuals' beliefs. We use this proposition to derive deterministic and probabilistic convergence times in Propositions 15 and 16, respectively. These convergence times are a function of a value that we call a "convergence factor", which depends on how we choose the meeting groups in our model. We determine the convergence factor for a few examples, and we end the chapter with some general bounds that we use to derive convergence times in subsequent chapters.

Chapter 6

We study convergence to consensus when we choose meeting groups uniformly at random from the nodes of a graph. With this update method, we study how the size of meeting groups affects convergence time to consensus. We begin the chapter by showing, through examples and simulations, that the convergence time for small meeting groups is at best approximately quadratic in the number nodes on some graphs. In Theorem 36, we provide an upper bound for the convergence time to consensus that is approximately quadratic in the number of nodes for large meeting groups and approximately quartic in the number of nodes for small meeting groups. To end the chapter, we present an application of our results to determining the optimal size of a meeting given various assumptions on the cost of a meeting.

Chapter 7

We consider strategies for choosing sets of nodes for the purpose of quickly reaching consensus. In addition to being applicable to conflict resolution, these strategies provide insight into our model by indicating the properties of both successful and unsuccessful meeting groups. In Algorithm 1, we state a greedy algorithm for choosing meeting groups. We then give upper bounds for convergence times in Theorems 42, 44, and 48 for three different specifications of our greedy algorithm. We also consider examples and simulations to determine lower bounds for the best possible convergence time. Lastly, we use simulations to compare the performance of our strategies to choosing meeting groups uniformly-at-random.

Chapter 8

We study convergence to consensus when an meeting group is uniformly-at-random chosen from the dyads of a graph (i.e. from pairs of adjacent nodes). This is a common assumption in models of social influence, so by considering dyadic updates, we are able to compare the behavior of our model with that of other social-influence models. We begin with examples and simulations to illustrate that the convergence time of our model is significantly longer than those of similar models in the literature. The main result of this section is Theorem 53, which gives an upper bound for the convergence time to consensus.

Chapter 9

In this final chapter, we draw conclusions.

CHAPTER 2

Background

Models of social influence attempt to explain how influence during interactions between individuals causes the emergence of patterns of belief change in social networks [Eps99, FMF17]. Although social-influence models are often derived from psychological principles and are designed to study sociological phenomena, some of them are similar to models that arise in other fields, such as in group decision-making and in certain areas of computer science [DeG74, FJ11, OT11, DZK18].

In Section 2.1, we discuss observations from psychological experiments on social influence. We reference these observations in Chapter 3 when we introduce our model of social influence. In Section 2.2, we discuss some implications of certain social-network modeling choices on studying opinion dynamics. We consider existing models of social influence that are similar to our model in Section 2.3, we overview models and methods for group decision-making in Section 2.4, and we explain some approaches to the "distributed-consensus problem" in Section 2.5.

2.1 Social Influence in Psychology

Broadly, social influence describes psychological changes in individuals that result from "the real, implied, or imagined presence of others" [WCG91]. Research into social influence began in earnest in the 1950s when researchers in group dynamics studied how communication between individuals in small groups generates opinion changes [FJ11]. This focus on interpersonal interactions has since given way to a more individualistic focus to understand "the internal cognitive representations, thoughts, and motivations of isolated social actors rather than their reactions to real social situations" [FW16]. Combining conceptual insights from both the interpersonal and individualistic thrusts in social-influence research, recent work [FW16] has considered how social influence can occur without direct interpersonal interactions. Such indirect social influence phenomenon is called "implicit social influence".

While explicit social influence captures influence from the "real. . . presence of others", implicit social influence includes the ". . . implied, or imagined presence of others" [WCG91]. As Forgas and Williams wrote, explicit social influence in forms such as conformity, persuasion, and obedience are "exceedingly important effects", but "we now know that social influence often works not because people are exposed to explicit pressures, but because they are encouraged to imagine events, experience particular moods, [or] are subconsciously influenced by incidental observations" [FW16]. Therefore, a complete study of social influence should include both its explicit and implicit forms. In this section, we survey psychological research on both explicit and implicit social influence.

The cornerstone of empirical research into social influence is the observation that the beliefs of individuals in small groups tend to move towards a consensus from repeated interpersonal interactions. This was first observed by Muzafer Sherif in [She36] during his autokinetic-effect experiment. The "autokinetic effect" is the name of the phenomenon that humans perceive movement from small stationary light sources in dark, featureless environments. In his experiment, Sherif had subjects, who were sorted into groups of three, estimate the distance that a light seemed to move. For each group, Sherif held several trials in which he rotated through group members, with a single group member estimating the movement of a light source while the other two listened. During these trials, Sherif found that the group members' estimates converged to a consensus estimate. This consensus was so persistent that group members continued to perceive light movement that was aligned with their group's consensus estimate when they were retested alone 28 days after the original experiment. Sherif's findings have since been corroborated in similar experiments by others (e.g. [Bov48, RBH54]).

Sherif purposely chose the autokinetic effect as the source of beliefs to remove any dependencies (such as prior experience or bias) from his subjects' beliefs. In the real world, however, individuals generally have prior opinions, norms, and prejudices that inform their beliefs, so it is natural to ask whether the consensus formation that was observed by Sherif carries over to more persistently-held beliefs. Solomon Asch's experiments on conformity [Asc51, Asc55, Asc56] shed light on this question. In his experiments, individuals in small groups were shown a card on which three lines of different lengths were drawn (see Figure 2.1). Asch instructed all but one of the group members to say that the same one of the two shorter lines was longer, while the remaining participant ("the test subject") was given no instructions and was not made aware of the instructions that were given to the other group members. Asch conducted many trials for each group; he used different cards in each trial, but he always queried the test subject last. Over these trials and with many groups, Asch found that the test subject answered with the same wrong answer that was given by the other subjects 36.8% of the time. He also found that 75% of test subjects gave at least one wrong answer.

Asch's experiment and other similar experiments (e.g. [BS96]) illustrated that individuals can be influenced even when they are confident in their beliefs. Since Asch's experiment, researchers have conducted similar studies to Asch's to go beyond merely illustrating the existence of social influence. In Asch's experiment, there are at least two explanations for why test subjects gave the wrong answer: (1) they considered the answers from the other participants as sufficient evidence that their perception was incorrect; and (2) they desired to join the "group" of other participants by giving the wrong answer, regardless of the correctness of the answer.

Deutsch and Gerard introduced the concepts of informational and normative social influence in [DG55] to distinguish between these two possible explanations. In Deutsch and

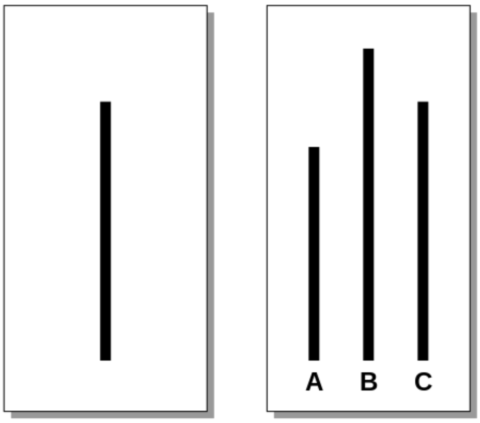


Figure 2.1: Example of a pair of cards that was shown in the Asch conformity experiments [Asc51, Asc55, Asc56]. The line on the card on the left is a reference line. [Image from [Oys].]

Gerard's conceptualization, a test subject in Asch's experiment experiences informational influence when they take the other partipants' answers as evidence of reality. By contrast, normative influence occurs when a test subject desires to be in conformity with a group by providing a certain answer. To test the effects of these different forms of influence, Deutsch and Gerard conducted several variations of Asch's experiment. In these experiments, they varied the amount of contact between test subjects and other partipants, but the information that was provided by the other partipants was kept constant [DG55]. Deutsch and Gerard found that the percentage of the time that test subjects joined other partipants in providing the wrong answer increased sharply with the amount of contact between test subjects and other participants. From this observation, they concluded that normative influence from a group increases with the degree of identification of an individual with a group. They also observed that test subjects provided the wrong answer more often when they were asked to provide an answer after the card of lines was removed from view. They thus also concluded that susceptibility to social influence is greater when there is a greater uncertainty in a belief.

To better understand how normative influence impacts individuals in the real world, researchers have measured the influence that an individual experiences from their pre-existing social groups. Blanchard et al. [BCB94] observed that students expressed more racist views after hearing racist opinions and less racist views after hearing anti-racist opinions if the opinions were expressed by students at their university. However, they did not observe this effect when students heard opinions of students from other universities. Stangor et al. [SSJ16] illustrated that this effect is not limited to being in direct personal contact with an influencer. They had students express opinions about a particular social group after telling them that other students at their university held certain racial views. Students expressed more racist beliefs if they were told other students had racist views, and conversely students expressed less racist beliefs when told that other students held unprejudiced views. Haslam et al. [HOM96] observed similar results. They asked groups of Australians to estimate the percentage of Australians with certain attributes (e.g. happy-go-lucky, pleasure loving, and sportsmanlike) and the percentage of Americans with other attributes (e.g. overly nationalistic, materialistic, and ostentatious) after either being shown fake information about the opinions of Australians about Australians and Americans or the opinions of "prejudiced people" about Australians and Americans. They found that subjects' estimated percentage for a given trait was larger if they were told that Australians believe other Australians or Americans have the trait. However, if they were told that "prejudiced people" believe that Australians or Americans have a trait, then there was no change in their estimated percentage.

The studies above give general trends on user behavior during interpersonal interaction, but individuals do not always change their beliefs. Friedkin and Johnsen [FJ11] investigated how pressure affects whether a group reaches consensus. To do so, they assumed that individuals have real-valued susceptiblities to influence between 0 and 1, where 0 means that an individual is not susceptible and 1 means that an individual has no resistance to influence, and they estimated susceptibility values by fitting parameters in their eponymous model

[FJ90]. In both high-pressure and low-pressure scenarios, they determined that susceptiblity follows a bimodal distribution with peaks at 0 and 1. However, with high pressure, 20.8% of particpants exhibited 0 susceptibility and 70.2% had a susceptibility value of 1; and with low pressure, 41.9% of participants had a susceptibility of 0 and 24.6% of participants had a susceptibility of 1. From these results, Friedkin and Johnsen concluded that pressure substantially increases the likelihood of consensus formation.

The psychological studies that we summarized above are consistent with Festinger's characterization of the force behind social influence: "...where the dependence upon physical reality is low the dependence upon social reality is correspondingly high. An opinion, a belief, an attitude is 'correct,' 'valid,' and 'proper' to the extent that it is anchored in a group of people with similar beliefs, opinions and attitudes" [Fes50]. A similar sentiment was expressed by Suk Hung Ng in [Ng16]: "[a]n influencee's resistance...is often embedded in group norms as much as in individual personality, cognitions, or emotions. Thus to influence an individual, the influencing party often has to influence the group as a whole." In general, individuals appear to seek consensus with others even when reaching consensus pushes them to compromise their own beliefs. However, the strength of influence is modulated by the stength of the influencer and influencee's relationship. With this general picture in mind, we now survey some research into quantitative aspects of social influence.

In addition to his main result on the existence of social influence, Asch [Asc51, Asc55, Asc56] tested the importance of unanimity of a majority in inducing a test subject to give an incorrect answer. Asch observed that if unanimity of the majority is broken by instructing a single individual to give the correct answer, then test subjects become much more likely to provide the correct answer than when unanimity is enforced. Taking this experiment to its logical extreme, researchers have studied whether minorities are able to change the opinions of majorities. Moscovici et al. [MLN69] showed cards with different shades of blue to groups of six particpants, and they then asked subjects to name the color on the card. The researchers instructed two participants in each trial to call every card "green". In 8.42%

of the trials, test subjects agreed that every card was green and 32% of the subjects agreed that at least one of the cards was green. This study and numerous follow-up studies have demonstrated that minorities are able to influence majorities, but the stength of a minority group's influence is less than that of a majority's [WLO94].

Many of the studies that we have cited thus far have involved discrete opinions, such as test subjects' choices of lines in Asch's studies [Asc51, Asc55, Asc56], rather than continuous opinions, such as guesses about the United States' current level of debt or about the length of the border between two countries [MTS13]. Therefore, the studies that we have summarized do not shed light on certain issues that are pertinent to continuous opinions. Friedkin and Johnsen [FJ11] investigated how continuous opinions change during interpersonal interactions by observing the initial and final opinions of individuals from group discussions on various issues. They observed that beliefs very rarely became more extreme than any of the initial beliefs; in other words, they observed that the terminal beliefs from interpersonal interactions usually lie in the convex hull of the intial beliefs.

If opinions stay in the convex hull of initial beliefs during interpersonal interactions, then where in the convex hull will the opinions lie? In particular, what is the magnitude of belief change that one should expect? McConahay et al. [MHB81] conducted a study in which they exposed college students to favorable information about a certain social group. They found that more prejudiced students changed their beliefs by a larger amount after exposure than did less prejudiced students. This observation suggests that, during some interpersonal interactions with a group of others, individuals with beliefs that are further way from the mean belief in the group undergo larger belief changes than those with beliefs that are closer to the mean belief. Mavrodiev et al. [MTS13] made a more nuanced conclusion after conducting a statistical analysis of the data from a study by Lorenz et al. [LRS11], in which individuals in groups were asked to answer a single quantitative question in each of five rounds. In some groups, individuals were provided in each round with the mean answer of another group. From their statistical analysis, Mavrodiev et al. concluded that individuals in groups that were provided with the mean guesses of other groups shifted their initial guesses by an amount that is proportional to the distance of their initial guesses from the mean [MTS13]. A similar phenomenon was reported in an experiment by Takács et al. [TFM16].

In agreement with these findings on the magnitude of belief changes, investigators have observed that if individuals are in agreement prior to an interaction, then they are likely to stay in agreement [Tho38, Bar59]. In both [Tho38] and [Bar59], subjects were placed into groups and tasked with solving several problems. When group members unanimously agreed on an answer to a problem, regardless of whether their answer was correct, the group members would assume that they were correct and move on to the next problem.

We end this section by summarizing the empirical results that we have discussed. It has been observed that social influence leads individuals to seek consensus with others even on strongly held beliefs [She36, Asc51, Asc55, Asc56, DG55]. The social influence that is experienced by individuals is both informational and normative, and normative influence (1) can be felt even without direct contact [DG55, HOM96, SSJ16], (2) is experienced more strongly from a group of individuals (assuming that the individuals have similar beliefs to one another) when an individual who is being influenced has a stronger social connection with the influencing group [DG55, BCB94, SSJ16, HOM96], (3) is felt more strongly when individuals have less confidence in their beliefs [DG55], and (4) is stronger during direct interpersonal interaction when the influencing group has a greater degree of unanimity [Asc51, MLN69, WLO94]. Additionally, individuals can be very resistant to attempts at influencing them if their beliefs are shared by many others in their social group [FI99, SGA13]. However, pressure to move towards consensus is known to be effective in motivating individuals to reach consensus [FJ11].

We have also discussed that an individual's opinion change during an interpersonal interaction wtih a group of others seems to occur within the convex hull of the initial beliefs of group members [FJ11]. More specifically, in some situations, the magnitude of a change in opinion is proportional on average to the difference between an initial opinion and the mean opinion of the group [MHB81, MTS13, TFM16]. A related observation is that opinions of individuals in a group that are initially at consensus do not change after [Tho38, Bar59]. We will use these observations in Chapter 3 during our development of our model of social influence.

2.2 Social-Network Models

A key task in mathematical sociology is to try to understand how social processes, such as opinion dynamics, are affected by social-network structure [Sco88, PG16]. To undertake this task, one must first decide how to model a social network based on the dynamics under investigation. Even if one decides to use a simple network structure like an undirected graph, in which nodes are individuals and edges represent mutual social ties, the edges of a graph can encode many possible sociological meanings [Sco88]. For example, edges in a graph may represent marriages, friendships, relationships in which individuals discuss "important matters", occassional or frequent contact between individuals, or a social connection between individuals who simply know of each other [MSC01]. For studying social influence, edges that encode certain weak ties, such as two individuals who are simply aware of each others' existence, are unlikely to be sufficient.

For studying social influence, the simplest and most direct route for defining edges is to specify that an edge between two individuals means that the individuals can influence one another. This route was taken by French [Fre56], who defined a graph in which an edge exists from individual A to individual B if A has "power" over B, where power is the "force which A can induce on B" [Fre56]. A practical problem with this approach is its abstraction: it is more difficult to collect data on power dynamics than it is to determine concrete social ties such as friendship, direct contact, and marriage. One can now obtain data on friendships or frequency of contact, for example, from online social networks [MMG07] or radio-frequency identification (RFID) sensors [CBB10] without direct questioning. However, determining whether one individual is influenced by another requires more work, such as directly asking if someone "discuss[es] important matters with" someone else [MSC01]. Accordingly, many readily available data sets represent social networks using graphs with edges that do not directly represent whether individuals can socially influence each other. A practical question, then, is if these data sets are viable resources for the study of social influence.

We argue that it is justifiable to study social-influence models on graphs with adjacencies defined by concrete social ties as long as the social ties are *strong* [Gra73]. According to [Gra73], the strength of an edge is a "combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie." It has been argued (e.g. [LEV81]) that individuals who are connected by strong ties tend to be similar in terms of ethnicity, sex and gender, age, religion, education, occupation, social class, attitudes, abilities, beliefs, aspirations, and other aspects [MSC01]. This idea that individuals who are strongly-tied are similar to each other is one notion of *homophily* [MSC01]. However, if strongly-tied individuals are similar, then our observation in Section 2.1 that social influence—specifically normative social influence—is stronger between individuals in the same social group than between individuals in different social groups implies that strongly-tied individuals experience more social influence from one another than weakly-tied individuals experience from one another. Another implication of homophily is that strongly-tied individuals are likely to have similar beliefs. This is due indirectly to homophily with respect to social groups, because individuals in the same social group—such as those who have similar religions, educations, occupations, and social classes—are likely to have similar beliefs [MSC01]. The similarity of strongly-tied individuals has also been observed directly in experiments [HL78].

Strong ties have implications beyond the viability of a graph for studying social influence, as graphs with strong ties often tend to have communities [FH16]. A community is a mesoscopic structure[1] that consists of a set of nodes in which the nodes in the set are rela-

[1]Intuitively, a mesoscopic structure exists neither at the scale of individual nodes nor at the scale of the

tively densely connected to other nodes in the set and sparsely connected to nodes outside of the set [POM09]. Communities have been observed to exist pervasively in numerous types of social networks, including "groups of hunter-gatherers, feudal structures, royal families, political and business organizations, families, villages, cities, states, nations, continents, and even virtual communities such as Facebook groups"[POM09]. The presence of community structure in graphs of strong ties is sometimes a consequence of homophily, because individuals who are similar to each other are likely to be strongly tied to one another and thereby to form a community.

To summarize, we have argued that strongly-tied individuals are likely to (1) socially influence one another, (2) have common membership in a social group, and (3) have similar beliefs. We have also argued that in graphs whose edges encode strong social ties, community structure occurs often and nodes in a community often represent individuals who are similar. In Figure 2.2, we illustrate a graph that agrees with these observations. This graph maintains its community structure with the addition of a small number of edges between red and blue nodes.

2.3 Social-Influence Modeling

Mathematical modeling of social influence began with French's seminal 1956 paper [Fre56]. French considered how a simple model of belief changes in which individuals adopt the mean belief of their social connections behaves on various directed graphs. He assigned a real-valued belief, which he denoted by $y_v^{(t)} \in \mathbb{R}$, to each node v and examined how it evolved for times $t \in \mathbb{Z}_{\geq 0}$. He specified that a node updates its belief using the rule

$$y_v^{(t+1)} = \frac{1}{|\mathcal{N}_{\text{in}}(v)|} \sum_{w \in \mathcal{N}_{\text{in}}(v)} y_w^{(t)} ,$$

entire graph; it exists somewhere in between.

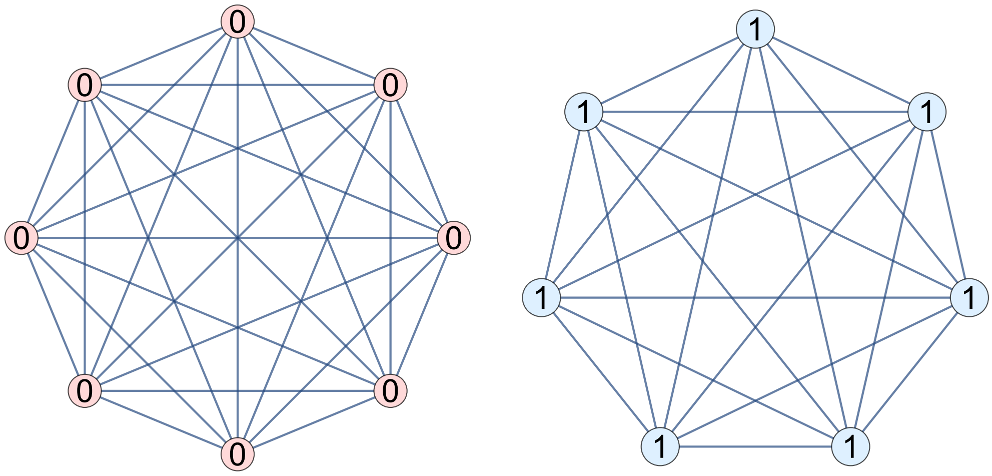


Figure 2.2: A prototypical graph from the considerations in Section 2.2. Red group members all have belief 0, and blue group members all have belief 1.

where $\mathcal{N}_{\text{in}}(v)$ denotes the set of nodes w such that there is a directed edge from w to v. With this model, French examined several directed graphs and patterns of dyadic interpersonal interactions that generate movement of beliefs towards consensus (see Figure 2.3). Additionally, French stated several theorems that provide conditions for the existence of a unique consensus in his model.

Harary analyzed French's model in [Har59] and used Markov chains to prove many of the theorems in French's paper that were stated without proof. DeGroot [DeG74] studied a model that is similar to French's model as a method that could be used by a group of experts with different initial opinions to come to a consensus opinion. Let V denote the set of individuals, and let the weight assigned by v to the opinion of w be denoted by $p_{vw} \geq 0$. DeGroot assumed that p_{vw} are non-negative real numbers that satisfy $\sum_{w \in V} p_{vw} = 1$, and he specified that each individual $v \in V$ updates their belief in each time step to be the weighted mean belief of the other individuals:

$$y_v^{(t+1)} = \sum_{w \in V} p_{vw} y_w^{(t)} .$$

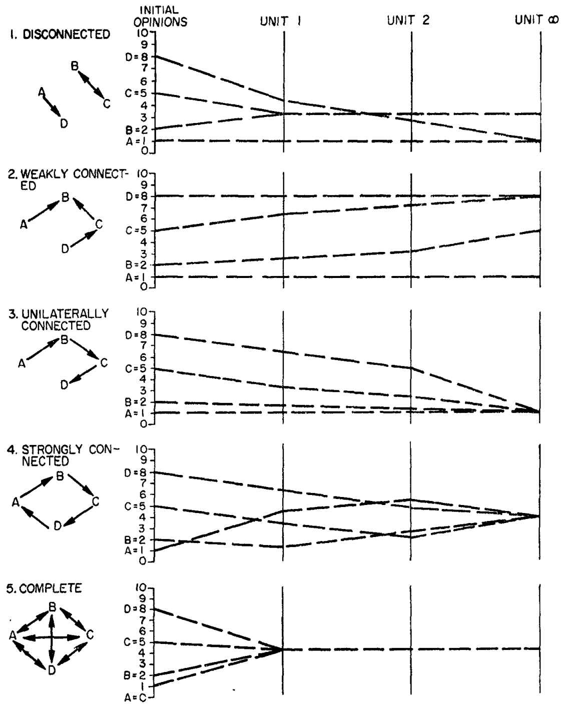


Figure 2.3: Opinion trajectories from French's model [Fre56] for various initial beliefs and graphs. For this figure [Image obtained from [Einnd]], French assumed that all nodes update their beliefs simultaneously in each time unit. Beliefs fail to converge to consensus for the disconnected and weakly-connected graphs in this figure, whereas beliefs converge to consensus for the strongly-connected graphs.

Let $V = \{1, \ldots, k\}$ for $k \in \mathbb{Z}_{>0}$, let $\mathbf{y}^{(t)}$ the vector of opinions $y_v^{(t)}$ for $v \in V$, and let $\mathbf{P}$ denote the matrix of weights $P_{vw} = p_{vw}$. In matrix form, DeGroot's model is

$$\mathbf{y}^{(t+1)} = \mathbf{P}\mathbf{y}^{(t)} . \tag{2.1}$$

DeGroot gave conditions on the matrix $\mathbf{P}$ so that his model has a unique consensus steady state. His analysis is similar to that of Harary [Har59] and is based on the theory of Markov chains.

Friedkin and Johnsen [FJ90] introduced a modified version of (2.1) that incorporates susceptibility to social influence. Let $\mathbf{W}$ be a square matrix whose entries are non-negative and whose row sums are equal to 1, and let $V = \{1, \ldots, k\}$ for $k \in \mathbb{Z}_{>0}$. For each $v, w \in V$, the value W_{vw} represents the relative interpersonal influence that is experienced by v from w. Additionally, let $\mathbf{\Lambda}$ be a diagonal matrix whose v^{th} entry $\alpha_v \in [0, 1]$ represents a susceptibility of v to influence, where a susceptiblity of 1 means that v does not resist social influence and a susceptibility of 0 means that v does not change its opinion. The model of Friedkin and Johnsen can be written as

$$\mathbf{y}^{(t+1)} = \mathbf{\Lambda} \mathbf{W} \mathbf{y}^{(t)} + (\mathbf{I} - \mathbf{\Lambda}) \mathbf{y}^{(0)} \,, \tag{2.2}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{y}^{(0)}$ is the vector of initial beliefs. The models of French and DeGroot are both special cases of (2.2). Setting $\mathbf{\Lambda}$ to be the identity matrix and setting $\mathbf{W} = \mathbf{P}$ results in the DeGroot model [DeG74]. One obtains the French model by also requiring that the entries in a given row that correspond to directed edges in a graph are equal and sum to 1. Friedkin and Johnsen [FJ11] have compared the changes predicted by their model (2.2) with belief changes that they've observed in their own psychological studies, and they have found that (2.2) is a accurate predictor of the mean and the end-of-trial beliefs of group members.[2]

The model in (2.2) has important limitations. Sociological questions like how distinct cultures, such as upper-class and lower-class cultures, develop and are maintained within the same society [Dur82, Bou84] and why polarization in politics occurs [FA08, AS08, FAP08], are not natural to study with (2.2) using the asymptotic methods that are common in opinion dynamics. Modelers have instead attempted to gain insight into these questions through the study of models with *clustered* steady states. A clustered steady state consists of sets of nodes with the same beliefs within each set but different beliefs across sets [FMF17]. A clutered steady state is called *polarized* if the steady-state beliefs are more extreme than individuals' initial beliefs [Axe97, FMF17]. Two models that have clustered steady states are the bounded-confidence models of Hegselmann and Krause [Kra00] and Deffuant et al.

[2]See Chapter 4 of [FJ11].

[DNA00].

Another limitation of (2.2) is that it is unnatural to directly model interpersonal interactions that involve a subset of individuals [FJ90]. Other models incorporate local interactions in a more natural way; these include voter models [CS73, HL75] and threshold models [Wat02, CM07, GBC18]. In the classical binary voter model [CS73, HL75], an updating node updates its belief to the belief of a randomly chosen neighbor [Red19]. In threshold models, a fraction (or number) of neighbors must meet or exceed a threshold for an individual to adopt a new opinion. Threshold models have been used often to study cascades of social influence [GBC18], whereas voter models have been used to study consensus and clustering [Red19].

2.4 Group Decision-Making

Researchers in group decision-making create strategies to promote agreement within groups of decision-makers [DG87, CEE92]. A simple example of a group-decision-making strategy is the Delphi method [DLR72, LT75], which is a process that was developed to be used by a panel of experts to reconcile differences and reach an agreement. Usually, the Delphi method consists of a round in which experts anonymously display their initial beliefs on a topic, followed by a round in which experts revise their beliefs. These two rounds are repeated until the group moves acceptably close to consensus, at which point the mean opinion of the experts is taken as the group decision. While opinion dynamics and group decision-making have many commonalities—indeed, DeGroot emphasized the similarities between his model and the Delphi technique in [DeG74]—opinion dynamics is usually descriptive or generative, whereas group decision-making usually gives prescriptions for obtaining desirable opinions.

Models of group decision-making generally consist of two processes: (1) a *consensus-reaching process* through which group members change their opinions to move towards positions that are felicitious for agreement; and (2) a *selection process*, through which group members choose agree on an opinion [HCC17, DZZ18]. The consensus-reaching process is applied first and consists of discussion rounds during which group members modify their opinions [DDM17]. After the consensus-reaching process finishes, the opinion-selection process commences and usually consists of two stages: (a) the aggregation of decision-makers' opinions into a set of acceptable alternatives and (b) the choice of an alternative from an acceptable set of decision [HCC17]. In the Delphi method, the consensus-reaching process is the repeated anonymous display of opinion by experts and subsequent revision of opinion, and the selection process is the choice of the mean opinion as the agreement.

The main body of work in group decision-making is concerned with methodologies that are run by a moderator and do not consider social structure [HL06, HCK14, KML16]. However, recent work in the field has begun to account for trust relationships, in the form of a social network, in decision-making [PMC16, WCF17]. The newest developments in the field include a focus on large-scale decision-making (situations in which at least 20 decision-makers must agree upon a decision) [DPW20] and on incorporating models of opinion dynamics [DZZ18]. The work on large-scale decision-making generally provides detailed recipes for how a moderator can break up a large group into smaller groups, apply decision-making strategies to these groups, and then manage the collection of the groups until they all move sufficiently close to the desired consensus [DPW20]. While it has been said that "[o]pinion dynamics could be used as a potential tool to develop novel consensus reaching models in group decision-making, by considering graph structures and the evolution of opinions" [DZK18], the inclusion of opinion dynamics into group decision-making is still in its infancy [HCC17].

In designing group-decision-making strategies, researchers have acknowledged that consensus is unlikely even though it is often assumed to be ideal [LL85]. This has led to the consideration of "soft consensus" in addition to "hard consensus". In a hard consensus, beliefs are identical; however, in a soft consensus, beliefs only come "close" to a consensus. The problem of measuring "closeness" was examined in the literature beginning with the work of Kacprzyk and Fedrizzi [KF86, KF88, KF89], who introduced the concept of "degree of consensus" to measure the closeness of decision-makers to a consensus. Since then, various definitions of "degree of consensus" have been developed for various applications, and each definition was proposed as a way to determine if a group is close enough to consensus to end a consensus-seeking process and to commence a selection process [HCK14].

2.5 The Distributed-Consensus Problem

The problem of bringing a network to consensus is well-studied in computer science [BCN20]. It arises, for example, in coordination problems in networks of autonomous agents and in estimation problems with noisy measurements [Lyn96, KDG03, OT11]. Distributed algorithms for these problems tend to fall into one of two categories: gossip-algorithms [BGP06] and distributed iterations [XB04]. Algorithms in either of these two categories involve nodes repeatedly passing information to their neighbors; however, algorithms in these two categories make distinct assumptions about nodal communication. In gossip algorithms, nodes are required to communicate with no more than one neighbor in each update; by contrast, in distributed-iteration algorithms, nodes use information from all of their neighbors in each update [BGP06, XB04]. If the consensus that is reached by an algorithm is the mean of the initial values at the nodes, then we say that the algorithm solves the distributed-averaging problem.

An important requirement of distributed algorithms is that they cannot use centralized nodes to facilitate consensus-reaching. This is a requirement for both the updating behavior at each node and the structure of a graph: each node must update using the same model, and there cannot be a node that is adjacent to all other nodes and whose removal disconnects the graph [BGP06]. One can implement gossip and distributed-iteration algorithms in two different ways, based on the times at which the nodes communicate. In *sychronous* implementations, each node updates its value simultaneously; in *asynchronous* implementations, each node updates its value independently of the update times of other nodes [BGP06].

Let $G = (V, E)$ denote a graph with node set V (with $n = |V|$ denoting the number of nodes in V) and edge set E, and let $\mathbf{x}(t) \in \mathbb{R}^n$ be a vector that consists of the values that are stored at the nodes at discrete time t. Let $\mathbf{W} \in \mathbb{R}^{n\times n}$ be a matrix with non-negative entries W_{vw} satisfying $W_{vw} = 0$ if $\{v, w\} \notin E$. A synchronous distributed linear iteration is given by the equation

$$\mathbf{x}(t + 1) = \mathbf{W}\mathbf{x}(t)\,. \tag{2.3}$$

If the matrix $\mathbf{W}$ is doubly stochastic, then $x_v(t) \to \bar{x}(0)$ for all $v \in V$ and $\bar{x}(t) = \bar{x}(0)$ for all $t \geq 0$, where $\bar{x}(t)$ denotes the mean of the initial values at the nodes [XB04]. Under these conditions, one obtains

$$\|\mathbf{x}(t) - \bar{x}(0)\mathbf{1}_n\|_2 \leq \left\|\mathbf{W} - \frac{\mathbf{1}_n\mathbf{1}_n^T}{n}\right\|_2^t \|\mathbf{x}(0) - \bar{x}(0)\mathbf{1}_n\|_2\,,$$

where $\mathbf{1}_n$ is the vector of n ones. Letting $\epsilon \in (0, 1)$, one can determine bounds on the convergence time of a distributed linear iteration by upper-bounding the minimum time t such that

$$\frac{\|\mathbf{x}(t) - \bar{x}(0)\mathbf{1}_n\|_2}{\|\mathbf{x}(0) - \bar{x}(0)\mathbf{1}_n\|_2} < \epsilon$$

holds for all non-zero $\mathbf{x} \in \mathbb{R}^n$ (e.g. [XB04]). Such bounds are often put in terms of the spectral radius

$$\rho\left(\mathbf{W} - \frac{\mathbf{1}_n\mathbf{1}_n^T}{n}\right) = \max\left\{|\lambda_{\min}|, |\lambda_{\max}|\right\}\,, \tag{2.4}$$

where $\lambda_{\max}$ and $\lambda_{\min}$ are the most positive and most negative eigenvalues of $\mathbf{W}$, respectively [OT11].

The ϵ-convergence time for a distributed linear iteration can be at least as bad as $O(n^3\log(\epsilon^{-1}))$ [OT11]. Even for modern computers, time complexities of n^3 are not attractive. Accordingly, in applications in which one is able to exercise a large amount of control over nodal communcation and/or graph structure, one is often interested in reducing convergence times by choosing a matrix $\mathbf{W}$. A way to do this is to minimize the spectral radius (2.4) subject to a constraint that $\mathbf{W}$ is doubly stochastic [XB04, OT11].

The worst-case time complexity of gossip algorithms is related to the worst-case time complexity of distributed linear iterations [BCN20]. For asynchronous gossip algorithms in which the mean value of $\mathbf{x}(t)$ is conserved (meaning that $\bar{x}(t) = \bar{x}(t-1)$ for all $t \in \{1, 2, \ldots\}$) the expected value of $\mathbf{x}(t)$ at time t is given by an iteration of the form (2.3). Because time is measured by the number of pairwise updates in gossip algorithms, rather than by the number of entire-network updates as in distributed averaging algorithms, convergence times that one deduces from (2.3) must be multiplied by roughly n to be translated to the gossip-algorithm setting [BGP06].

CHAPTER 3

A Strong Consensus-Seeking Model of Social Consensus-Formation

Let $G = (V, E)$ be an undirected graph with $n = |V| \geq 2$ nodes and $m = |E|$ undirected edges. We set $V = \{1, \ldots, n\}$ and denote the adjacency matrix of G by $\mathbf{A}$. We call $v, w \in V$ *neighbors* if there is an edge $\{v, w\} \in E$, and we call the set of all neighbors of v the *neighbors of* v. We denote the set of neighbors of a node v by $\mathcal{N}(v)$. The nodes in G represent individuals, and the presence of an edge $\{v, w\}$ between two nodes v and w signifies that the nodes v and w have a strong social tie.[1] From our argument in Section 2.2, if a pair of individuals has a strong social tie, then they are likely to socially influence each other. Accordingly, we make the assumption that all edges in G are pathways for social influence.

We represent the belief of each individual $v \in V$ at time t by a real number $x_v(t) \in \mathbb{R}$. We assume that beliefs change only during interpersonal interactions between two or more individuals, and we refer to these interpersonal interactions as *meetings*. We sometimes drop the dependence of $x_v(t)$ on t and write x_v when we are not concerned with time. Let $\mathbf{x}(t)$ denote the vector of beliefs of individuals at time t; the v^{th} component of $\mathbf{x}(t)$ equals $x_v(t)$. We sometimes refer to the vector of beliefs $\mathbf{x}(t)$ as a *belief state*.

Because beliefs do not change outside of discrete meeting events, our model is event-based, and we index the meeting events by $t \in \{0, 1, \ldots\}$ and call t the (discrete) meeting time. The belief $x_v(t)$ denotes the belief of individual v immediately after meeting t. We call

[1] See Section 2.2 for more on strong ties.

the set of individuals who participate in a meeting a *meeting group*, and we denote a meeting group at time t by S^t. We require that the number of nodes in S^t, which we denote by $|S^t|$, satisfies $|S^t| \in \{2, \ldots, n-1\}$. As with the belief $x_v(t)$, we sometimes drop the dependence of S^t on t and write S for a set of nodes that are participating in a meeting. In Section 2.1, we cited several psychological studies in which it was observed that individuals often seek consensus with others even for strongly held beliefs. We now formulate a model for how the nodes in S^t change their beliefs to move closer to consensus with the other nodes in S^t as a result of a meeting.

During a meeting at time t of the nodes in the set S^t, we model the belief change of node $v \in S^t$ by

$$x_v(t|\beta, S^t) = \begin{cases} x_v(t-1) + \beta\left(\bar{x}^S(t-1) - x_v(t-1)\right), & \text{if } v \in S^t \\ x_v(t-1), & \text{if } v \notin S^t, \end{cases} \tag{3.1}$$

where $\beta \in [0,1]$ and $\bar{x}(t) = \frac{1}{|S^t|}\sum_{v \in S^t} x_v(t)$ denotes the mean belief of nodes in S^t. If $\beta = 1$, then updating by (3.1) causes the nodes in S^t to reach consensus; if $\beta = 0$, then the nodes in S^t do not change their beliefs. We call β the *degree of consensus*, because it controls the amount of movement of a group towards consensus.[2] Let $\mathbb{1}_S(v)$ denote the function that is equal to 1 if $v \in S$ and that is equal to 0 otherwise. It will be convenient to write the update rule (3.1) in vector form as

$$\mathbf{x}(t|\beta, S^t) = \mathbf{x}(t-1) + \beta\left(\bar{x}^{S^t}(t-1)\mathbf{1}_{S^t} - \mathbf{x}^{S^t}(t-1)\right), \tag{3.2}$$

where $\mathbf{1}_S$ denotes the vector defined by $(\mathbf{1}_S)_v = \mathbb{1}_S(v)$ and $\mathbf{x}^{S^t}(t)$ has components

$$x_v^{S^t}(t) = x_v(t)\mathbb{1}_{S^t}(v)\,.$$

In Section 2.1, we summarized evidence from psychological experiments that indicate that (1) after group interaction, beliefs tend move to a position in the convex hull of initial

[2]We borrow the terms "soft consensus" and "degree of consensus" from the group-decision-making literature. See Section 2.4 for more on group decision-making.

beliefs, (2) the magnitude of an individual's belief change is sometimes proportional to the difference between an individual's initial belief and the mean belief of a group, and (3) influence is stronger during direct interpersonal interaction when the influencing group has a greater degree of unanimity. The model (3.1) *is* observation (2); because we have assumed that $\beta \geq 0$, model (3.1) agrees with observation (1).

The model is consistent with observation (3) in the following sense. Fix a degree of consensus β and assume that the opinions of z nodes in S^t equal 1 and that $|S| - z$ opinions of the nodes in S^t equal 0. By the model (3.1), a node v whose belief equals 0 undergoes the belief change

$$x_v(t|\beta, S) = \beta \frac{z}{|S|} \,.$$

We consider the set of individuals with belief equal to 1 to be the influencing group. For progressively larger z, which we consider to be analogous to unanimity increasing, node v experiences a progressively larger change in belief. Consequently, an increase in the number of nodes with one opinion causes an increase in the size of the belief change for the nodes of the other opinion. In this specific sense, we say that our model agrees with observation (3).

The model (3.1) requires a choice of the degree of consensus to fully specify the belief changes of nodes in a meeting. We model the degree of consensus as the cost minimizer of the total cost that is incurred by the nodes in S^t due to belief differences with their neighbors. To model this cost, we assume that the cost that is incurred by individual v due to a belief difference with w is given by $\frac{1}{2}(x_v - x_w)^2$. Therefore, the total cost that is incurred is

$$\frac{1}{2} \sum_{v,w=1}^{n} A_{vw}(x_v - x_w)^2 \,. \tag{3.3}$$

Let $\mathbf{L}$ denote the un-normalized graph Laplacian of G. It has components

$$L_{vw} = \begin{cases} k_v \,, & \text{if } w = v \\ -1 \,, & \text{if } A_{vw} = 1 \,, \end{cases}$$

where k_v denotes the degree of node v. We expand the vector–matrix product $\mathbf{x}^T\mathbf{L}\mathbf{x}$ as follows:

$$\mathbf{x}^T\mathbf{L}\mathbf{x} = \frac{1}{2}\sum_{v,w=1}^{n} A_{vw}(x_v - x_w)^2\,,$$

from which we see that the total cost (3.3) is given by $\mathbf{x}^T\mathbf{L}\mathbf{x}$.

Let G^S denote the graph G but with edges between nodes in S added, let $\mathbf{A}^S$ denote the adjacency matrix of G^S, and let $\mathbf{L}^S$ denote the un-normalized graph Laplacian of G^S. We assume that for the duration of a meeting of a group S of individuals, all pairs of individuals in the meeting seek consensus with one another. In our model, we capture this assumption by adding all edges between pairs of nodes in S to the graph G for the duration of a meeting; i.e., we consider the graph G^S. Our motivation in making this assumption are the studies that we discussed in Section 2.1 that illustrated that individuals in small groups often seek consensus with one another (e.g. [She36, Asc51, BS96, Mil63, DG55]). In particular, we cited evidence that the likelihood to reach consensus increases with pressure to make an agreement [FJ11]. Therefore, our model may be most applicable when individuals experience pressure to reach an agreement.

The total cost (3.3) at time t after a meeting with nodes S^t and with degree of consensus β is then given by

$$C(\mathbf{x}(t|\beta, S^t)) := \mathbf{x}(t|\beta, S^t)^T\mathbf{L}^{S^t}\mathbf{x}(t|\beta, S^t)\,. \tag{3.4}$$

We assume that the meeting group chooses the degree of consensus β to minimize the cost $C(\mathbf{x}(t-1|\beta, S^t))$. We denote the degree of consensus of the group by $\beta^*(t)$ and denote the vector of beliefs after a meeting at time t by $\mathbf{x}(t)$. Our model of belief changes for a meeting group S^t is

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{x}(t|\beta^*(t), S^t)\,, \\ \beta^*(t) &= \operatorname*{arg\,min}_{\beta\in[0,1]} C(\mathbf{x}(t|\beta, S^t))\,. \end{aligned} \tag{3.5}$$

We call (3.5) the *group-soft-consensus model* and call $\beta^*(t)$ the *optimal degree of consensus.*

We call an application of (3.5) at some $t \in \mathbb{Z}_{>0}$ an "update". An update at time t consists of using $\mathbf{x}(t-1)$ and S^t to compute $\mathbf{x}(t)$ from the model (3.5) and formula (3.2). We "simulate" model (3.5) by (1) setting $\mathbf{x}(0) = \mathbf{x}_0$ for some $\mathbf{x}_0 \in \mathbb{R}^n$; (2) specifying a method to select a meeting group S^t; and (3) updating $\mathbf{x}(1)$ from $\mathbf{x}(0)$ and S^1, then updating $\mathbf{x}(2)$ from $\mathbf{x}(1)$ and S^2, and so on until we have computed a desired number of updates. We can use a computer to simulate model (3.5) by calculating values of $\mathbf{x}(t)$ for some finite number of time steps, or we can consider infinite sequences $(\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \ldots)$ that we understand to be the result of using our simulation procedure for an arbitrary number of time steps. We study the "dynamics", the "behavior", or the "properties" of model (3.5) by considering the mathematical properties of sequences $(\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \ldots)$ for different choices of $\mathbf{x}_0$ and $(S^t)_{t \in \mathbb{Z}_{>0}}$.

The model (3.5) aligns with several observations from the psychological experiments on normative influence that we discussed in Section 2.1. Normative influence (1) can be felt even without direct contact [BCB94, SSJ16, HOM96, FI99, SGA13] and (2) is experienced more strongly from a group of individuals (assuming that the individuals have similar beliefs to one another) when an individual who is being influenced has a stronger social connection with the influencing group [DG55, BCB94, HOM96, FI99, SGA13, SSJ16]. The model is consistent with (1), because the optimal degree of consensus $\beta^*(t)$ takes into account the beliefs of the neighbors of nodes in S^t. We revisit the consistency of the group-soft-consensus model (3.5) with observation (2) in Chapter 4 after we have proven some properties of the model (see Remark 6).

CHAPTER 4

Properties of the Group-Soft-Consensus Model

In Chapter 3, we developed the group-soft-consensus model (3.5) to describe belief changes during group interactions.[1] In this model, interacting individuals come to a soft consensus as a result of explicit influence from other group members and implicit influence from group members' social connections. Group members choose a degree of consensus of the soft consensus by minimizing the cost that is incurred by the group from disagreements with their social connections.

We show in Chapters 6–8 that the repeated application of the group-soft-consensus model (3.5) with certain choices of meeting groups results in convergence to consensus. Using terminology from computer science (see Section 2.5), this means that the group-soft-consensus model (3.5) solves the distributed-consensus problem. More specifically, we say that the model solves the distributed-averaging problem (see Section 2.5 and [OT11]), because updates that follow (3.5) do not change the mean of the values that are stored in the nodes in a network.

Proposition 1. *The mean belief $\bar{x}(t)$ is unchanged after an update of the form* (3.5)*.*

Proof. For any set $S \subseteq V$, we have that $(\bar{x}^S \mathbf{1}_S)^T \mathbf{1}_n = \bar{x}^S |S|$ and that

$$(\mathbf{x}^S)^T \mathbf{1}_n = \sum_{v \in S} x_v = \bar{x}^S |S|\,,$$

which together imply that

$$(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)^T \mathbf{1}_n = 0\,.$$

[1]We call interactions between two or more individuals "group interactions".

Therefore, from the expression (3.2) for $\mathbf{x}(t|\beta, S)$, we obtain

$$\mathbf{x}(t|\beta, S)^T \mathbf{1}_n = \mathbf{x}(t-1)^T \mathbf{1}_n \tag{4.1}$$

for all $\beta \in [0, 1]$. In particular, (4.1) is true for $\beta = \beta^*(t)$ and $S = S^t$, so $\mathbf{x}(t) = \mathbf{x}(t-1)$ for all $t \in \{1, 2, \ldots\}$. □

We determine the optimal degree of consensus by minimizing the total cost $C(\mathbf{x}(t|\beta, S^t))$ that is defined in (3.4) with respect to β. In the following proposition, we write the total cost $C(\mathbf{x}(t)|\beta, S^t)$ in a form that will make it simpler to state the optimal degree of consensus. For notational convenience, we omit the time t from $\mathbf{x}(t|\beta, S^t)$ and write

$$\mathbf{x}(\beta, S) = \mathbf{x} + \beta\left(\bar{x}^S \mathbf{1}_S - \mathbf{x}\right).$$

Proposition 2. *Let k_v^S denote the degree of node $v \in V$ in the graph G^S. One can write the total cost as*

$$C(\mathbf{x}(\beta, S)) = \mathbf{x}^T \mathbf{L}^S \mathbf{x} + 2\beta C_1(\mathbf{x}, S) + \beta^2 C_2(\mathbf{x}, S), \tag{4.2}$$

where

$$\begin{aligned} C_1(\mathbf{x}, S) &= -\sum_{v \in S}(k_v^S + 1)(x_v - \bar{x}^S)^2 + \sum_{v \in S}\sum_{w \notin S} A_{vw}(x_w - \bar{x}^S)(x_v - \bar{x}^S), \\ C_2(\mathbf{x}, S) &= \sum_{v \in S}\left(k_v^S + 1\right)(x_v - \bar{x}^S)^2. \end{aligned} \tag{4.3}$$

To obtain the optimal degree of consensus, we set the derivative of (4.2) with respect to β equal to 0 and solve for β after making the substitutions $\mathbf{x} \to \mathbf{x}(t-1)$ and $S \to S^t$. We state the optimal degree of consensus explicitly in the following proposition, where we use the notation $\beta(\mathbf{x}, S)$ for the optimal degree of consensus when we regard it as a function $\mathbf{x}$ and S, rather than of t.

Proposition 3. *The optimal degree of consensus as a function of $\mathbf{x}$ and S is given by*

$$\beta(\mathbf{x}, S) = \begin{cases} 1, & -C_1(\mathbf{x}, S) \geq C_2(\mathbf{x}, S) \\ \frac{-C_1(\mathbf{x}, S)}{C_2(\mathbf{x}, S)}, & C_2(\mathbf{x}, S) > -C_1(\mathbf{x}, S) > 0 \\ 0, & C_1(\mathbf{x}, S) \geq 0. \end{cases} \tag{4.4}$$

As a function of t, the optimal degree of consensus satisfies $\beta^*(t) = \beta(\mathbf{x}(t-1), S^t)$.

In light of the formula (4.4) for the optimal degree of consensus from Proposition 3, we study the costs $C_1(\mathbf{x}, S)$ and $C_2(\mathbf{x}, S)$ to understand $\beta^*(t)$. In the following proposition, we write $C_1(\mathbf{x}, S)$ in terms of two costs: $C_1^I(\mathbf{x}, S)$, which represents the costs that are incurred due to changes in belief differences between group members; and $C_1^O(\mathbf{x}, S)$, which represents costs that are incurred due to changes in belief differences between group members and individuals outside of the group. We call $C_1^I(\mathbf{x}, S)$ the *in-group cost* and $C_1^O(\mathbf{x}, S)$ the *out-group cost.*

Proposition 4. *With the same setup and notation as in Proposition 2, we have*

$$C_1(\mathbf{x}, S) = C_1^I(\mathbf{x}, S) + C_1^O(\mathbf{x}, S)\,, \tag{4.5}$$

where

$$\begin{aligned} C_1^I(\mathbf{x}, S) &= -\frac{1}{2}\sum_{v,w\in S}(x_v - x_w)^2 = -|S|\sum_{v\in S}(x_v - \bar{x}^S)^2\,, \\ C_1^O(\mathbf{x}, S) &= -\sum_{v\in S}(x_v - \bar{x}^S)\sum_{w\notin S}A_{vw}(x_v - x_w)\,. \end{aligned} \tag{4.6}$$

Before we prove this proposition, we state the following lemma.

Lemma 5. *Denote the un-normalized graph Laplacian of the complete graph of n nodes by $\mathbf{L}_n$. It then follows that*

$$\|\mathbf{x} - \bar{x}\mathbf{1}_n\|^2 = \frac{1}{n}\mathbf{x}^T\mathbf{L}_n\mathbf{x}\,.$$

We now prove Proposition 4.

Proof. Using the definition of k_v^S and Lemma 5, we calculate that

$$\begin{aligned} \sum_{v\in S}(k_v^S + 1)(x_v - \bar{x}^S)^2 &= |S|\sum_{v\in S}(x_v - \bar{x}^S)^2 + \sum_{v\in S}\sum_{w\notin S}A_{vw}(x_v - \bar{x}^S)^2 \\ &= \frac{1}{2}\sum_{v,w\in S}(x_v - x_w)^2 + \sum_{v\in S}\sum_{w\notin S}A_{vw}(x_v - \bar{x}^S)^2\,. \end{aligned} \tag{4.7}$$

Inserting (4.7) into $C_1(\mathbf{x}, S)$ from (4.3) yields

$$C_1(\mathbf{x}, S) = \frac{1}{2} \sum_{v,w \in S} (x_v - x_w)^2 + \sum_{v \in S} \sum_{w \notin S} A_{vw}(x_v - \bar{x}^S) \left(x_v - \bar{x}^S - (x_w - \bar{x}^S)\right) \tag{4.8}$$
$$= \frac{1}{2} \sum_{v,w \in S} (x_v - x_w)^2 + \sum_{v \in S} (x_v - \bar{x}^S) \sum_{w \notin S} A_{vw} (x_v - x_w) .$$

This formula for $C_1(\mathbf{x}, S)$ is identical to the formula (4.5) after we use the formulas for $C_1^I(\mathbf{x}, S)$ and $C_1^O(\mathbf{x}, S)$ from (4.6). We prove the second equality for $C_1^I(\mathbf{x}, S)$ in (4.6) by applying Lemma 5. □

We revisit our assertion from Chapter 3 that the group-soft-consensus model (3.5) is consistent with certain observations about normative influence that we discussed in Section 2.1.

Remark 6. *In Chapter 3, we stated that our model (3.5) is consistent with the observation discussed in Section 2.1 that normative influence is experienced more strongly from a group of individuals (assuming that the individuals have similar beliefs to one another) when an individual who is being influenced has a stronger social connection with the influencing group. We describe why this is true when all nodes in an influencing group have identical beliefs.*

Let v^* be a node in V, and let S be a set of nodes. We assume that each node in S has the same belief y, where $y \in \mathbb{R}$. We rewrite $C_1^O(\mathbf{x}, S)$ in Proposition 4 as

$$C_1^O(\mathbf{x}, S) = - \sum_{v \in S \setminus \{v^*\}} (x_v - \bar{x}^S) \sum_{w \notin S} A_{vw}(x_v - x_w)$$
$$- (x_{v^*} - \bar{x}^S) \sum_{w \in V \setminus (S \cup \mathsf{S})} A_{v^*w}(x_{v^*} - x_w) - (x_{v^*} - \bar{x}^S) \sum_{w \in (V \setminus S) \cap \mathsf{S}} A_{v^*w}(x_{v^*} - x_w) .$$

The only term of the right-hand side that depends on the number of edges between v^* and S is the last term. Using our assumption that $x_w = y$ for all $w \in \mathsf{S}$, we rewrite this term as follows:

$$(x_{v^*} - \bar{x}^S) \sum_{w \in (V \setminus S) \cap \mathsf{S}} A_{v^*w}(x_{v^*} - x_w) = (x_{v^*} - \bar{x}^S)(x_{v^*} - y) \sum_{w \in (V \setminus S) \cap \mathsf{S}} A_{v^*w} . \tag{4.9}$$

The right-hand side of Equation (4.9) becomes increasingly positive as the number of edges between v^* and S *increases if* $(x_{v^*} - \bar{x}^S)(x_{v^*} - y) > 0$ *and becomes increasingly negative as the number of those edges increases if* $(x_{v^*} - \bar{x}^S)(x_{v^*} - y) < 0$*. From this fact, the formula (4.4) for $\beta^*(t)$ in Proposition 3, and the formula (4.5) for $C_1(\mathbf{x}, S)$ in Proposition 4, we conclude that more edges between v^* and* S *entails a greater effect of* S *on $\beta^*(t)$.*

In the next two examples, we present graphs in which the optimal degree of consensus equals 1 for all beliefs and all meeting groups that are not initially at consensus.

Example 7. *Let G be the graph of n disconnected nodes; i.e. G contains n nodes and 0 edges. Using Proposition 4, we calculate that $\beta(\mathbf{x}, S) = 1$ for all $\mathbf{x} \in \mathbb{R}^n$ and $S \subseteq V$ such that at least 2 nodes in S have different beliefs. This implies that $\beta^*(t) = 1$ for all $t \in \{1, 2, \ldots\}$ as long as the nodes in S^t do not all have the same belief.*

Example 8. *Let G be the complete graph with n nodes. Using Proposition 2, we calculate that*

$$C_1(\mathbf{x}, S) = -n \sum_{v \in S} (x_v - \bar{x}^S)^2 = -C_2(\mathbf{x}, S)\,.$$

From this equality, we obtain that $\beta^*(t) = 1$ if at least 2 nodes in S have different beliefs.

We conclude from these examples that our model behaves identically when nodes are not adjacent to any other node and when each node is adjacent to all other nodes. In a complete graph, influence cancels out so that meeting groups are not inhibited from reaching a hard consensus; in the disconnected graph of nodes, there is no external influence on a meeting.

In the following example, we present a graph and a belief state in which the only possible agreement of a meeting is a soft consensus.

Example 9. *Let G be a graph with n of nodes, where n is even, that consists of two complete graphs of size $n/2$ that are disconnected from one another. We call this graph the "graph of two equally-sized cliques". Let $\mathbf{x}_0$ be the belief state that consists of the following values: all nodes in one of the complete graphs have belief -1, and all nodes in the other complete*

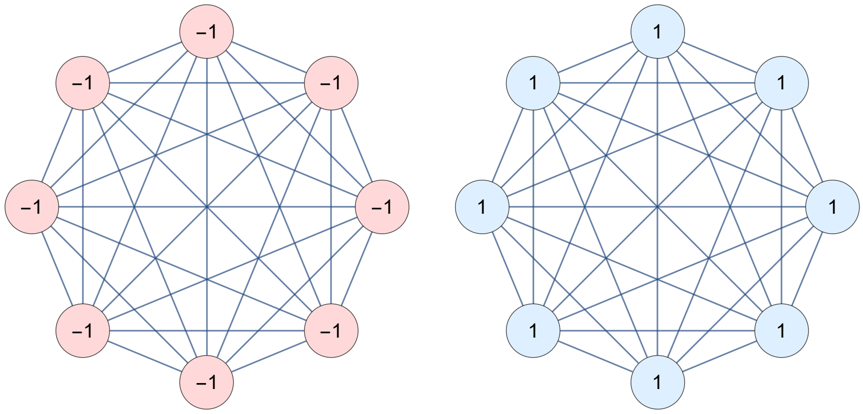


Figure 4.1: Illustration with $n = 8$ of the graph of two equally-sized cliques and the beliefs from Example 9.

graph have belief $+1$ *(see Figure 4.1). Using Proposition 4, we calculate that* $C_1^O(\mathbf{x}_0, S) = 0$, *which implies that*

$$C_1(\mathbf{x}_0, S) = -|S| \sum_{v \in S} (x_v - \bar{x}^S)^2 \,. \tag{4.10}$$

Additionally, we stated in Proposition 2 that

$$C_2(\mathbf{x}_0, S) = \sum_{v \in S} (k_v^S + 1)(x_v - \bar{x}^S)^2 \,. \tag{4.11}$$

Because $k_v^S \geq |S|$ *for any meeting group* S *that is not contained in the node set of one of the two complete graphs, we obtain* $k_v^S + 1 > |S|$ *for such meeting groups* S*. Comparing* (4.10) *and* (4.11) *and using the fact that* $k_v^S + 1 > |S|$*, we conclude that*

$$-C_1(\mathbf{x}_0, S) < C_2(\mathbf{x}_0, S) \,. \tag{4.12}$$

Therefore, $\beta(\mathbf{x}_0, S) < 1$ *for all meeting groups* S *that are not in the node set of one of the complete graphs. Additionally, if* S *is in the node set of one of the two complete graphs, then* $C_1(\mathbf{x}_0, S) = 0$*, which implies that* $\beta(\mathbf{x}_0, S) = 0$*. Therefore,* $\beta^*(1) < 1$*, so every group agreement is a soft consensus.*

Additional Proofs

Proof of Lemma 5. Because $A_{vw} = 1$ for all $i \neq j$, it follows that

$$\mathbf{x}^T \mathbf{L}_n \mathbf{x} = \frac{1}{2} \sum_{v,w=1}^{n} A_{vw}(x_v - x_w)^2 = \frac{1}{2} \sum_{v,w=1}^{n} (x_v - x_w)^2 \,.$$

Continuing this calculation by adding and subtracting $\bar{x}$ and expanding the square, we obtain

$$\mathbf{x}^T \mathbf{L}_n \mathbf{x} = \frac{1}{2} \sum_{v,w=1}^{n} (x_v - \bar{x} + \bar{x} - x_w)^2 = \frac{1}{2} \left(2 \sum_{v,w=1}^{n} (x_v - \bar{x})^2 + 2 \sum_{v,w=1}^{n} (x_v - \bar{x})(\bar{x} - x_w) \right) .$$

Using the fact that $\sum_{v=1}^{n} (x_v - \bar{x}) = 0$ and $\sum_{w=1}^{n} (x_v - \bar{x})^2 = n(x_v - \bar{x})^2$, it then follows that

$$\mathbf{x}^T \mathbf{L}_n \mathbf{x} = n \sum_{v=1}^{n} (x_v - \bar{x})^2 \,.$$

We divide this last equation by n to finish the proof. □

Proof of Proposition 2. Let $\mathbf{L}^S$ be the un-normalized graph Laplacian of the graph G^S. Using the definition of $\mathbf{x}(\beta, S)$, we obtain

$$C(\mathbf{x}(\beta, S)) - \mathbf{x}^T \mathbf{L}^S \mathbf{x} = 2\beta \mathbf{x}^T \mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) + \beta^2 (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) \mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) \,. \tag{4.13}$$

We need to show that $2\mathbf{x}^T \mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) = C_1(\mathbf{x}, S)$ and $(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) \mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) = C_2(\mathbf{x}, S)$.

We first show that $2\mathbf{x}^T \mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) = C_1(\mathbf{x}, S)$. To do so, we start by writing $\mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)$ componentwise:

$$[\mathbf{L}^S (\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)]_v = \begin{cases} (k_v^S + 1)(\bar{x}^S - x_v)\,, & v \in S \\ -\sum_{w \in S} A_{vw} (\bar{x}^S - x_w)\,, & v \notin S\,, \end{cases} \tag{4.14}$$

where we use the definitions of $\mathbf{L}^S$, $\bar{x}^S$, and $\mathbf{x}^S$ and the equality

$$\sum_{w \in S : w \neq v} (\bar{x}^S - x_w) = -(\bar{x}^S - x_v) \tag{4.15}$$

to obtain (4.14). When $v \in S$, the definition of k_v^S implies that $\sum_{w \notin S} A_{vw} = k_v^S - (|S| - 1)$. Using this equation, we obtain

$$\sum_{w \notin S} A_{vw} x_w = \sum_{w \notin S} A_{vw}(x_w - \bar{x}^S) - |S|\bar{x}^S + (k_v^S + 1)\bar{x}^S\,, \tag{4.16}$$

which, in turn, implies that

$$(k_v^S + 1)x_v - \sum_{w \notin S} A_{vw} x_w = (k_v^S + 1)(x_v - \bar{x}^S) + |S|\bar{x}^S - \sum_{w \notin S} A_{vw}(\bar{x}^S - x_w)\,. \tag{4.17}$$

Using (4.14) and

$$\sum_{v \notin S} x_v \sum_{w \in S} A_{vw}(\bar{x}^S - x_w) = \sum_{v \in S} (\bar{x}^S - x_v) \sum_{w \notin S} A_{vw} x_w\,,$$

we calculate that

$$\begin{aligned} 2\mathbf{x}^T \mathbf{L}^S(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) &= 2\sum_{v \in S}(k_v^S + 1)x_v(\bar{x}^S - x_v) - 2\sum_{v \notin S} x_v \sum_{w \in S} A_{vw}(\bar{x}^S - x_w) \\ &= 2\sum_{v \in S}\left((k_v^S + 1)x_v - \sum_{w \notin S} A_{vw} x_w\right)(\bar{x}^S - x_v)\,. \end{aligned}$$

Continuing this calculation by using (4.17), we obtain

$$\begin{aligned} 2\mathbf{x}^T \mathbf{L}^S(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) &= 2\sum_{v \in S}\left((k_v^S + 1)(x_v - \bar{x}^S) + |S|\bar{x}^S - \sum_{w \notin S} A_{vw}(\bar{x}^S - x_w)\right)(\bar{x}^S - x_v) \\ &= C_1(\mathbf{x}, S)\,. \end{aligned} \tag{4.18}$$

We now show that the expression for $C_2(\mathbf{x}, S)$ in (4.3) is the same as $(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)\mathbf{L}^S(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)$. We write $\mathbf{L}^S = \mathbf{L}|_S + \mathbf{D}_S + \mathbf{B}$, where $\mathbf{L}|_S$ denotes the un-normalized graph Laplacian of $G^S|_S$, the complete subgraph induced by S in G^S, the matrix $\mathbf{D}_S$ denotes the diagonal matrix with entries

$$[\mathbf{D}_S]_{vv} = \begin{cases} k_v^S - (|S| - 1)\,, & v \in S \\ 0\,, & v \notin S\,, \end{cases} \tag{4.19}$$

and $\mathbf{B} = \mathbf{L}^S - \mathbf{L}|_S - \mathbf{D}_S$. By Lemma 5, we obtain

$$(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)\mathbf{L}|_S(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) = |S| \sum_{v \in S} (x_v - \bar{x}^S)^2 \,. \tag{4.20}$$

The fact that $\mathbf{B}_{vw} = 0$ when $v, w \in S$ implies that

$$(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)\mathbf{B}(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) = 0 \,. \tag{4.21}$$

Combining equations (4.19), (4.20), and (4.21), we obtain the following desired result:

$$\begin{aligned}(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S)\mathbf{L}^S(\bar{x}^S \mathbf{1}_S - \mathbf{x}^S) &= |S| \sum_{v \in S} (x_v - \bar{x}^S)^2 + \sum_{v \in S} (k_v^S - |S| + 1)(x_v - \bar{x}^S)^2 \\ &= C_2(\mathbf{x}, S) \,.\end{aligned}$$

□

CHAPTER 5

Convergence to Consensus

In Chapter 1, we stressed the importance of studying convergence times to obtain empirically testable hypotheses when studying models of social influence. In this chapter, we outline our strategy for determining convergence times. We begin by defining how we measure the distance of nodes' beliefs from consensus.

Definition 10. *We denote the* distance from consensus *of the vector of beliefs* $\mathbf{x} \in \mathbb{R}^n$ *by* $e(\mathbf{x})$*, and we define it with the formula*

$$e(\mathbf{x}) := \|\mathbf{x} - \bar{x}\mathbf{1}_n\|_2^2 \,.$$

Let $\mathbf{x}(t)$ be given by the group-soft-consensus model (3.5). We derive convergence times by estimating how quickly $e(\mathbf{x}(t))$ converges to 0. As a first step, we state a recurrence relation for $e(\mathbf{x}(t))$ that describes how $e(\mathbf{x}(t))$ evolves in time.

Proposition 11. *Let* $\mathbf{x}(0) = \mathbf{x}_0$ *for some* $\mathbf{x}_0 \in \mathbb{R}^n$*. The distance from consensus satisfies the recurrence relation*

$$e(\mathbf{x}(t+1)) = \left(1 - \beta^*(t)(2 - \beta^*(t)) \frac{\sum_{v \in S^{t+1}} (x_v(t) - \bar{x}^{S^{t+1}}(t))^2}{e(\mathbf{x}(t))}\right) e(\mathbf{x}(t)) \qquad (5.1)$$

for $t \in \{1, 2, \ldots\}$*.*

Proof. Denote the un-normalized graph Laplacian of the complete graph with n nodes by $\mathbf{L}_n$. We rewrite the definition of $e(\mathbf{x})$ using Lemma 5 and obtain

$$e(\mathbf{x}) = \frac{1}{n}\mathbf{x}^T \mathbf{L}_n \mathbf{x} \,. \qquad (5.2)$$

From the definition of $C(\mathbf{x}(\beta, S))$ in (3.4), we have that

$$C(\mathbf{x}(\beta, S)) = \mathbf{x}(\beta, S)^T \mathbf{L}_n \mathbf{x}(\beta, S) \,. \tag{5.3}$$

Combining (5.2) with (5.3), we obtain

$$C(\mathbf{x}(\beta, S)) = ne(\mathbf{x}(\beta, S)) \,. \tag{5.4}$$

For the complete graph, if $w \notin S$, then $\sum_{v \in S} A_{vw}(x_v - \bar{x}^S) = 0$. Therefore,

$$\sum_{v \in S} \sum_{w \notin S} A_{vw}(x_w - \bar{x}^S)(x_v - \bar{x}^S) = 0 \,.$$

From Proposition 2 and equations (5.2) and (5.4), we obtain

$$n(e(\mathbf{x}(\beta, S)) - e(\mathbf{x})) = -n\beta(2 - \beta) \sum_{v \in S} (x_v - \bar{x}^S)^2 \,. \tag{5.5}$$

Making the substitutions $\beta \to \beta^*(t)$, $S \to S^{t+1}$, and $\mathbf{x} \to \mathbf{x}(t)$ in (5.5) and dividing by n, we see that

$$e(\mathbf{x}(t+1)) - e(\mathbf{x}(t)) = -\beta^*(t)(2 - \beta^*(t)) \sum_{v \in S^{t+1}} (x_v(t) - \bar{x}^{S^{t+1}}(t))^2 \,.$$

We put this equation into the form of the recurrence relation (5.1) by rearranging terms. □

We make the following definitions in preparation for our convergence-time analysis of the recurrence relation (5.1) in Proposition 11.

Definition 12 (See [XB04]). *Assume that $Z(\mathbf{x})$ is a function from $\mathbb{R}^n$ to the subsets of V. Define the* per-step convergence factor of $\mathbf{x}$ and Z *by*

$$r_{\text{step}}(\mathbf{x}, Z) = \frac{e(\mathbf{x}(\beta(\mathbf{x}, Z(\mathbf{x})), Z(\mathbf{x})))}{e(\mathbf{x})} \,, \tag{5.6}$$

and define the per-step convergence factor of Z *by*

$$r_{\text{step}}(Z) = \sup_{\mathbf{x} \neq \bar{x}\mathbf{1}_n} r_{\text{step}}(\mathbf{x}, Z) \,. \tag{5.7}$$

We often abuse notation slightly and use the notation $r_{\text{step}}(\mathbf{x}, S)$ and $r_{\text{step}}(S)$ to mean the per-step convergence factors of the fixed set S (i.e., $Z(\mathbf{x}) = S$).

In the following two definitions, we detail what we mean by 'convergence time'. We provide notions for the convergence time of both deterministic and probabilistic processes.

Definition 13. *Let $\epsilon \in (0, 1)$, and let $Z(\mathbf{x})$ be a deterministic function from $\mathbb{R}^n$ to the subsets of V. Let $S^t = Z(\mathbf{x}(t-1))$. We denote the* deterministic ϵ-convergence time of Z *by $T(\epsilon, Z)$ and define it by*

$$T(\epsilon, Z) = \sup_{\mathbf{x}(0)\in\mathbb{R}^n} \inf_{t\geq 0} \left(\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} < \epsilon \right).$$

Definition 14 (Definition 1 in [BGP06])**.** *Let $\epsilon \in (0, 1)$, and let $Z(\mathbf{x})$ be a random variable on the subsets of V for each $\mathbf{x} \in \mathbb{R}^n$. Let $S^t = Z(\mathbf{x}(t-1))$. We denote the* probabilistic ϵ-convergence time of Z *by $T(\epsilon, Z)$ and define it by*

$$T(\epsilon, Z) = \sup_{\mathbf{x}(0)\in\mathbb{R}^n} \inf_{t\geq 0} \left(\Pr\left(\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} \geq \epsilon \right) \leq \epsilon \right).$$

It will be clear whether Z is a deterministic function or a random variable from context, so we omit the terms 'deterministic' and 'probabilistic' when considering ϵ-convergence times.

To understand how the per-step convergence factor of Definition 12 relates to the ϵ-convergence times of Definitions 13 and 14, we use the recurrence relation (5.1) of Proposition 11. From this recurrence relation, the per-step convergence factor of Z is

$$r_{\text{step}}(Z) = \sup_{\mathbf{x}\neq\bar{x}\mathbf{1}_n} \left(1 - \beta(\mathbf{x}, Z(\mathbf{x}))(2 - \beta(\mathbf{x}, Z(\mathbf{x}))) \frac{\sum_{v\in Z(\mathbf{x})} (x_v - \bar{x}^{Z(\mathbf{x})})^2}{e(\mathbf{x})} \right). \tag{5.8}$$

Using $r_{\text{step}}(Z)$ and Proposition 11, we write

$$\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} \leq \left(r_{\text{step}}(Z)\right)^t. \tag{5.9}$$

By equation (5.9), we obtain a convergence time by determining the quantity $r_{\text{step}}(Z)$. Let $\log(\cdot)$ denote the base-e logarithm. For all $\epsilon > 0$, if

$$t > \frac{\log(\epsilon)}{\log\left(r_{\text{step}}(Z)\right)}, \tag{5.10}$$

then it follows that

$$\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} < \epsilon \,. \tag{5.11}$$

The combination of inequalities (5.10) and (5.11) yields a bound for the ϵ-convergence time of Z. We summarize this convergence-time bound in the following proposition.

Proposition 15. *Let $Z(\mathbf{x})$ denote a deterministic function from $\mathbb{R}^n$ to the subsets of V. Suppose that $\hat{r}_{\text{step}}(Z)$ is an upper bound for the per-step convergence factor $r_{\text{step}}(Z)$. For all $\epsilon \in (0,1)$, we have the following bound for the ϵ-convergence time of Z:*

$$T(\epsilon, Z) \leq \frac{\log(\epsilon)}{\log(\hat{r}_{\text{step}}(Z))} \,.$$

Proof. This is a result of Proposition 11 and the inequalities (5.9), (5.10), and (5.11). □

When examining non-deterministic Z, we consider only a uniform random variable on the sets of nodes of a fixed size. Accordingly, we state our general convergence result for non-deterministic Z when Z does not depend on $\mathbf{x}$.

Proposition 16. *Let Z denote a random variable on subsets of V. Suppose that $\mathrm{E}\left[\hat{r}_{\text{step}}(Z)\right]$ is an upper bound for the expected per-step convergence factor $\mathrm{E}\left[r_{\text{step}}(Z)\right]$. For all $\epsilon \in (0,1)$, we have the following bound for the ϵ-convergence time of Z:*

$$T(\epsilon, Z) \leq \frac{2\log(\epsilon)}{\log(\mathrm{E}\left[\hat{r}_{\text{step}}(Z)\right])} \,. \tag{5.12}$$

The convergence times in Propositions 15 and 16 depend critically on the function Z that we use to select the meeting groups S. For some choices of Z, convergence does not occur or occurs exceedingly slowly; for others, convergence is very fast.

Remark 17. *The function $x \mapsto 1/\log(1-x)$ satisfies*

$$\frac{1}{\log(1-x)} = -\frac{1}{x} + O(1) \tag{5.13}$$

as $x \to 0^+$*. Applying* (5.13) *to the bound of Proposition 15, we obtain*

$$T(\epsilon, Z) \leq \frac{\log(\epsilon^{-1})}{1 - \hat{r}_{\text{step}}(Z)} + O(1) \tag{5.14}$$

as $1 - \hat{r}_{\text{step}}(Z) \to 0^+$*. Similarly, applying* (5.13) *to the bound of Proposition 16, we obtain*

$$T(\epsilon, Z) \leq \frac{2\log(\epsilon^{-1})}{\mathrm{E}\left[1 - \hat{r}_{\text{step}}(Z)\right]} + O(1) \tag{5.15}$$

as $\mathrm{E}\left[1 - \hat{r}_{\text{step}}(Z)\right] \to 0^+$.

In the following example, we continue Examples 7 and 8 by calculating the per-step convergence factor.

Example 18 (Per-step convergence factor for Examples 7 and 8)**.** *Let* G *be the graph of* n *disconnected nodes, or let* G *be the complete graph with* n *nodes. We stated in Examples 7 and 8 that* $\beta(\mathbf{x}, S) = 1$ *for all* $\mathbf{x} \in \mathbb{R}^n$ *and* $S \subseteq V$ *that includes at least* 2 *nodes that have different beliefs. From this fact, one can show that the per-step consensus factor satisfies*

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S) = \frac{\sum_{v \in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})}$$

for all $\mathbf{x} \in \mathbb{R}^n$ *and all* $S \subseteq V$ *that includes at least* 2 *nodes that have different beliefs.*

From the result of Example 18, we conclude that the complete graph and the graph of disconnected nodes have identical convergence behavior. Additionally, for all $\mathbf{x} \in \mathbb{R}^n$ and all functions Z from $\mathbb{R}^n$ to the subsets of V, the per-step convergence factors of $\mathbf{x}$ and Z for the complete graph and the graph of disconnected nodes are minimal among undirected graphs. This is because $\mathfrak{r}_{\text{step}}(\mathbf{x}, S)$ is minimal when $\beta(\mathbf{x}, S) = 1$, which is the case for the complete graph and the graph of disconnected nodes.

Furthermore, convergence cannot be faster on a graph than it is for the complete graph and the graph of disconnected nodes. This statement follows from the equality

$$\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} = r_{\text{step}}(\mathbf{x}(t), S^t) \times r_{\text{step}}(\mathbf{x}(t-1), S^{t-1}) \times \cdots \times r_{\text{step}}(\mathbf{x}(1), S^1)\,. \tag{5.16}$$

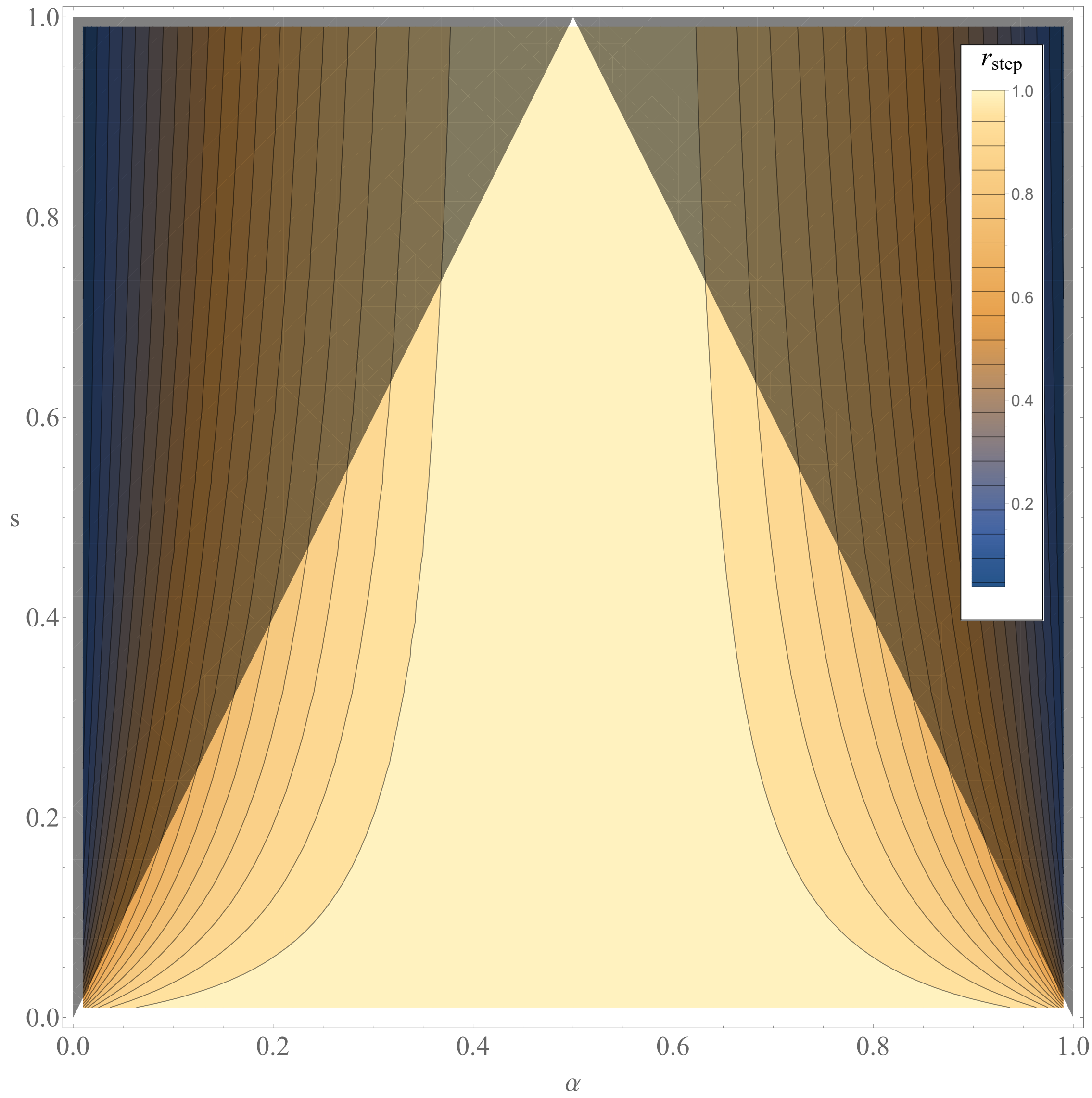


Figure 5.1: Contour plot of $(1 - r_{\text{step}}(S_p))/(1 - r_{\text{step}}(S_e))$ from Example 19. The inequality $\beta s \leq \mu$ holds in the region that is shaped like an isosceles triangle.

By (5.16), the size of $e(\mathbf{x}(t))/e(\mathbf{x}(0))$ is determined by the per-step convergence factors of $\mathbf{x}(t)$ and S^t. Because these per-step convergence factors are minimal (among undirected graphs) for the complete graph and the graph of disconnected nodes, $e(\mathbf{x}(t))/e(\mathbf{x}(0))$ is also minimal on these two graphs.

In the next example, we use our model to consider a toy scenario of how one would choose a meeting group to reduce disagreement between individuals in two disagreeing factions.

Example 19. *Let $G = (V, E)$ be a graph of n nodes that consists of two complete graphs $G_0 = (V_0, E_0)$ and $G_1 = (V_1, E_1)$ that are each of size $n/2$ and are disconnected from one another. Let $\mathbf{x}_0$ denote a belief vector whose components that correspond to nodes in G_0 are 0 and whose components that correspond to nodes in G_1 are 1 (see Figure 2.2). Denote the* relative size *of S by s, and define it by $s = |S|/n$. Additionally, denote*

$$\mu = \frac{|V_1|}{n}\,,$$
$$c = \frac{|S \cap V_1|}{|S|}\,.$$

We note that $cs \leq \mu$.

We consider the size of $\imath_{\text{step}}(\mathbf{x}, S)$ when (1) S is chosen so that the fraction of nodes from G_1 in S is equal to the fraction of nodes in V that are from G_1 (i.e. $c = \mu$) and (2) the number of nodes from G_1 in S is equal to the number of nodes from G_2 in S (i.e. $c = 1/2$). Let S_p be a subset of V that satisfies $c = \mu$, and let S_e be a subset of V that satisfies $c = 1/2$. In Figure 5.1, we plot $1 - \imath_{\text{step}}(\mathbf{x}, S_p)$ divided by $1 - \imath_{\text{step}}(\mathbf{x}, S_e)$. As indicated in the figure, this ratio is always less than or equal to 1, so

$$\imath_{\text{step}}(\mathbf{x}, S_p) \geq \imath_{\text{step}}(\mathbf{x}, S_e)\,.$$

This implies that the distance from consensus after a meeting is smaller if a meeting occurs with the same number of nodes from each clique than if a meeting occurs with a number of nodes from each clique that is proportional to the number of nodes in each clique.

Before considering more examples, we state the following lemma.

Lemma 20. *Let $\mathbf{M}$ be an $n \times n$ Hermitian, positive semi-definite matrix with real entries whose null space contains $\mathbf{1}_n$. Let $\lambda_2(\mathbf{M})$ denote the second-smallest eigenvalue of $\mathbf{M}$, and let $\lambda_{\max}(\mathbf{M})$ denote the largest eigenvalue of $\mathbf{M}$. The following inequality holds:*

$$\lambda_2(\mathbf{M})e(\mathbf{x}) \leq \mathbf{x}^T\mathbf{M}\mathbf{x} \leq \lambda_{\max}(\mathbf{M})e(\mathbf{x})\,. \tag{5.17}$$

In the following two examples, we use Proposition 16 to determine the ϵ-convergence time for certain time-homogeneous gossip algorithms (see Section 2.5 and [BGP06]).

Example 21. *In this example, we consider a "hard-consensus-only" model in which we fix β to be equal to 1 in the update rule (3.2). Let $\mathcal{U}_d$ denote the uniform random variable on the dyads $S = \{u, v\}$ in G. We calculate that*

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, \mathcal{U}_d)\right] = \frac{\mathrm{E}\left[(x_u - x_v)^2\right]}{2e(\mathbf{x})} = \frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{2me(\mathbf{x})}\,,$$

where the second equality uses the fact that $\sum_{v \in S}(x_v - \bar{x}^S)^2 = \frac{1}{2}(x_v - x_u)^2$. However, by Lemma 20,

$$\frac{\mathbf{x}^T \mathbf{L} \mathbf{x}}{2me(\mathbf{x})} \geq \frac{1}{2m}\lambda_2(L)\,,$$

where $\lambda_2(L)$ is the second-smallest eigenvalue of the un-normalized Laplacian matrix. Consequently,

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, \mathcal{U}_d)\right] \geq \frac{1}{2m}\lambda_2(L)\,. \tag{5.18}$$

Using the bound in Table 5.1 on $\lambda_2(L)$ for the path graph, we obtain that

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}_d)\right] \geq \frac{2}{n(n-1)^2}\,. \tag{5.19}$$

Our bound (5.18) allows for $\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}_d)\right]$ to be very small for large n.

In Table 5.1, we use the bound (5.18) to calculate upper bounds for the ϵ-convergence time of various graphs. For the complete graph and the star graph, the convergence time is approximately linear in the number n of nodes. By contrast, the convergence may be as poor as approximately cubic in n for the path graph and the cycle graph.

Example 22. *Let $\mathcal{U}_d'$ be a random variable that outputs a dyad by (1) selecting a node uniformly at random and (2) selecting an edge uniformly at random from the neighbors of the chosen node in step (1). For this example, we maintain the setting of Example 8, except that we consider the random variable $\mathcal{U}_d'$ instead of $\mathcal{U}_d$. We denote the expectation with respect*

Graph	Lower bound for $\lambda_2(L)$	Upper bound for $T(\epsilon, \mathcal{U}_d)$
Complete Graph (K_n)	n	$4\log(\epsilon^{-1})(n-1) + O(1)$
Path Graph (P_n)	$\frac{4}{n(n-1)}$	$\log(\epsilon^{-1})n(n-1)^2 + O(1)$
Cycle Graph (C_n)	$\frac{8}{n(n-1)}$	$\frac{1}{2}\log(\epsilon^{-1})n^2(n-1) + O(1)$
Bipartite Complete Graph ($K_{p,q}$)	$\min(p,q)$	$4\log(\epsilon^{-1})pq\min(p,q) + O(1)$
Star Graph (S_n)	1	$4\log(\epsilon^{-1})(n-1) + O(1)$

Table 5.1: We give ϵ-convergence times for the model in Example 21 for various graphs. We obtained lower bounds for the algebraic connectivity $\lambda_2(L)$ using results from [De 07]. We use (5.15) from Remark 17 to obtain the asymptotic expressions for the ϵ-convergence-time upper bounds. The limit for the asymptotic expressions is $n \to \infty$.

to the random variable $\mathcal{U}'_d$ by $\mathrm{E}[\cdot]$. The probability of choosing the edge $\{u, v\}$ is

$$P_{uv} = P_{vu} = \frac{1}{n}\left(\frac{1}{k_u} + \frac{1}{k_v}\right),$$

from which we obtain

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, \mathcal{U}'_d)\right] = \frac{\sum_{u,v \in E} P_{uv}(x_u - x_v)^2}{2e(\mathbf{x})}. \tag{5.20}$$

Letting $\hat{\mathbf{W}}$ be the matrix defined by

$$\hat{\mathbf{W}} = \mathbf{D} - \mathbf{P},$$

where $\mathbf{D}$ is a diagonal matrix with entries $\mathbf{D}_{vv} = \frac{1}{2}\sum_{u \in V}(\mathbf{P}_{uv} + \mathbf{P}_{vu})$, we write (5.20) as

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, \mathcal{U}'_d)\right] = \frac{\mathbf{x}^T \hat{\mathbf{W}} \mathbf{x}}{e(\mathbf{x})}.$$

In a similar manner as in Example 21, we have

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}'_d)\right] \geq \lambda_2(\hat{\mathbf{W}}).$$

Applying Proposition 16 then gives the following upper bound for the ϵ-convergence time:

$$T(\epsilon, \mathcal{U}'_d) \leq \frac{2\log(\epsilon)}{\log\left(1 - \lambda_2(\hat{\mathbf{W}})\right)}. \tag{5.21}$$

If we let $\mathbf{W} = I - \hat{\mathbf{W}}$*, then* $\lambda_2(\hat{\mathbf{W}}) = 1 - \lambda_2(\mathbf{W})$*. The upper bound* (5.21) *in terms of* $\lambda_2(\mathbf{W})$ *is identical to the upper bound of Theorem 3 in [BGP06].*[1]

We now present a method for determining a bound for a per-step convergence factor. Because $0 \leq \beta(\mathbf{x}, S) \leq 1$ implies that $2 - \beta(\mathbf{x}, S) \geq 1$, we obtain

$$1 - \tau_{\text{step}}(\mathbf{x}, S) \geq \beta(\mathbf{x}, S) \frac{\sum_{v \in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})} = -\frac{\beta(\mathbf{x}, S)}{|S|} \frac{C_1^I(\mathbf{x}, S)}{e(\mathbf{x})} \,.$$

Using the formula (4.4) for $\beta(\mathbf{x}, S)$ from Proposition 3, it then follows that

$$\beta(\mathbf{x}, S) \frac{\sum_{v \in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})} = \begin{cases} -\frac{1}{|S|} \frac{C_1^I(\mathbf{x},S)}{e(\mathbf{x})} \,, & -C_1(\mathbf{x}, S) \geq C_2(\mathbf{x}, S) \\ \frac{1}{|S|} \frac{C_1(\mathbf{x},S)}{C_2(\mathbf{x},S)} \frac{C_1^I(\mathbf{x},S)}{e(\mathbf{x})} \,, & 0 \leq -C_1(\mathbf{x}, S) \leq C_2(\mathbf{x}, S) \\ 0 \,, & -C_1(\mathbf{x}, S) \leq 0 \,. \end{cases} \tag{5.22}$$

We consider each of the cases in (5.22) separately. First, suppose that $0 \leq -C_1(\mathbf{x}, S) \leq C_2(\mathbf{x}, S)$. Because $C_2(\mathbf{x}, S) = \sum_{v \in S}(k_i^S + 1)(x_v - \bar{x}^S)^2$ and $k_i^S + 1 \leq n$, we obtain

$$-\frac{\beta(\mathbf{x}, S)}{|S|} \frac{C_1^I(\mathbf{x}, S)}{e(\mathbf{x})} \geq \frac{|S|}{n} \,.$$

Therefore,

$$\frac{1}{|S|} \frac{C_1(\mathbf{x}, S)}{C_2(\mathbf{x}, S)} \frac{C_1^I(\mathbf{x}, S)}{e(\mathbf{x})} \geq -\frac{1}{n} \frac{C_1(\mathbf{x}, S)}{e(\mathbf{x})} \,,$$

so we obtain the estimate

$$-\frac{\beta(\mathbf{x}, S)}{|S|} \frac{C_1^I(\mathbf{x}, S)}{e(\mathbf{x})} \geq -\frac{1}{n} \frac{C_1(\mathbf{x}, S)}{e(\mathbf{x})} \,. \tag{5.23}$$

The estimate (5.23) also holds if $-C_1(\mathbf{x}, S) \geq C_2(\mathbf{x}, S)$ and $C_1^O(\mathbf{x}, S) > 0$, because $C_1(\mathbf{x}, S) > C_1^I(\mathbf{x}, S)$ in this case. The case when (5.23) does not necessarily hold is when $-C_1(\mathbf{x}, S) \geq C_2(\mathbf{x}, S)$ and $C_1^O(\mathbf{x}, S) < 0$. In this case, using the estimate from (5.22) and the fact that $|S| < n$, we obtain

$$\beta(\mathbf{x}, S) \frac{\sum_{v \in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})} \geq -\frac{1}{n} \frac{C_1^I(\mathbf{x}, S)}{e(\mathbf{x})} \,. \tag{5.24}$$

---

[1] *Boyd et al. [BGP06] used the 2-norm, rather than the squared 2-norm, to measure the distance from consensus. Using the 2-norm results in a factor of 3 (instead of a factor of 2) in the numerator of the* ϵ*-convergence-time bound.*

To make it easier to use the bounds (5.23) and (5.24), we introduce the following notation. Let $\mathcal{S}^< = \{S \subseteq V : C_1(\mathbf{x}, S) < 0\}$ and $\mathcal{S}^- = \{S \in \mathcal{S}^< : C_1^O(\mathbf{x}, S) < 0\}$. The sets in $\mathcal{S}^<$ are the sets for which the bound (5.24) holds, and the sets in $\mathcal{S}^-$ are the sets for which the bound (5.24) holds but (5.23) does not hold. We define the function

$$\mathcal{B}(\mathbf{x}, S) := C_1(\mathbf{x}, S)\mathbb{1}(S \in \mathcal{S}^< \setminus \mathcal{S}^-) + C_1^I(\mathbf{x}, S)\mathbb{1}(S \in \mathcal{S}^-)\,.$$

Combining the estimates (5.23) and (5.24), we conclude that

$$1 - \imath_{\text{step}}(\mathbf{x}, S) \geq -\frac{1}{n}\frac{\mathcal{B}(\mathbf{x}, S)}{e(\mathbf{x})}\,. \tag{5.25}$$

Let $\mathcal{S}$ denote an arbitrary family of sets of nodes (i.e. a set of sets of nodes). In both Chapters 6 and 7, we will obtain lower bounds for the mean of $1 - \imath_{\text{step}}(\mathbf{x}, S)$ over a family of sets $\mathcal{S}$ during of our proofs of our convergence-time bounds. By (5.25), we can determine the mean of $\mathcal{B}(\mathbf{x}, S)$ over $\mathcal{S}$ to obtain a bound for the mean of $1 - \imath_{\text{step}}(\mathbf{x}, S)$ over $\mathcal{S}$. In the following proposition, we state an upper bound for the mean of $\mathcal{B}(\mathbf{x}, S)$ in terms of an upper bound for $\sum_{S\in\mathcal{S}} C_1(\mathbf{x}, S)$.

Proposition 23. *Let $\mathcal{S}$ be a family of sets of nodes, and let $\ell \in \{2, \ldots, n-1\}$ be a fixed integer. Assume that $|S| = \ell$ for all $S \in \mathcal{S}$. Additionally, let $w(S)$, where $S \in \mathcal{S}$, denote non-negative weights such that $\sum_{S\in\mathcal{S}} w(S) = 1$. If there exists a negative upper bound for the weighted sum*

$$\sum_{S\in\mathcal{S}} w(S)C_1(\mathbf{x}, S) \leq -\delta\,, \tag{5.26}$$

then

$$\sum_{S\in\mathcal{S}} w(S)\mathcal{B}(\mathbf{x}, S) \leq -\left(\sqrt{2} - 1\right)^2 \delta \min\left(1, \delta\frac{4\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right),$$

where we take $1/(\mathbf{x}^T\mathbf{L}\mathbf{x}) = +\infty$ when $\mathbf{x}^T\mathbf{L}\mathbf{x} = 0$.

To use Proposition 23, we require an upper bound for $\sum_{S\in\mathcal{S}} w(S)C_1(\mathbf{x}, S)$. The following lemma will help us obtain such an upper bound in Chapters 6 and 7.

Lemma 24. *Fix $S \subseteq V$ and $\mathbf{x} \in \mathbb{R}^n$. The following equalities hold:*

$$\sum_{v \in S} C_1^I(\mathbf{x}, S \setminus \{v\}) = (|S| - 2) C_1^I(\mathbf{x}, S)\,, \tag{5.27}$$

$$\sum_{v \in S} C_1^O(\mathbf{x}, S \setminus \{v\}) = -\frac{1}{2}\frac{|S|-2}{|S|-1} \sum_{v,w \in S} A_{vw}(x_v - x_w)^2 + \frac{|S|(|S|-2)}{|S|-1} C_1^O(\mathbf{x}, S)\,.$$

Additional Proofs

Proof of Proposition 16. Let $Z_1, \ldots, Z_t$ denote identically and independently distributed copies of the random variable Z. Denote the expectation with respect to the set of random variables $Z_1, \ldots, Z_t$ by $\mathrm{E}[e(\mathbf{x}(t))]$, and denote the conditional expectation given the random variable Z_l by $\mathrm{E}[\cdot|Z_l]$. We first prove that

$$\mathrm{E}[e(\mathbf{x}(t+1))] = \left(\mathrm{E}\left[r_{\text{step}}(Z)\right]\right)^{t+1} e(\mathbf{x}(0)) \tag{5.28}$$

by induction. Using the fact that $\mathbf{x}(0)$ is a constant and Proposition 11, we obtain

$$\mathrm{E}[e(\mathbf{x}(1))] = \mathrm{E}\left[r_{\text{step}}(\mathbf{x}(0), Z)\right] e(\mathbf{x}(0)) \le \mathrm{E}\left[r_{\text{step}}(Z)\right] e(\mathbf{x}(0))\,.$$

Assume that (5.28) holds for $t = \jmath - 1$. Because $e(\mathbf{x}(\jmath))$ is deterministic given the random variables $Z_1, \ldots, Z_\jmath$, we have that

$$\mathrm{E}\left[e(\mathbf{x}(\jmath+1))|Z_1, \ldots, Z_\jmath\right] = \mathrm{E}\left[r_{\text{step}}(\mathbf{x}(\jmath), Z)|Z_1, \ldots, Z_\jmath\right] e(\mathbf{x}(\jmath))\,.$$

However,

$$\mathrm{E}\left[r_{\text{step}}(\mathbf{x}(\jmath), Z)|Z_1, \ldots, Z_\jmath\right] \le \sup_{\substack{\mathbf{x}(\jmath) \in \mathbb{R}^n \\ \mathbf{x}(\jmath) \perp \mathbf{1}_n}} \mathrm{E}\left[r_{\text{step}}(\mathbf{x}(\jmath), Z)\right] \le \mathrm{E}\left[\sup_{\substack{\mathbf{x}(\jmath) \in \mathbb{R}^n \\ \mathbf{x}(\jmath) \perp \mathbf{1}_n}} r_{\text{step}}(\mathbf{x}(\jmath), Z)\right],$$

so using the definition of $r_{\text{step}}(Z)$ yields

$$\mathrm{E}[e(\mathbf{x}(\jmath+1))] \le \mathrm{E}\left[r_{\text{step}}(Z)\right] \mathrm{E}[e(\mathbf{x}(\jmath))]\,, \tag{5.29}$$

by the tower property of conditional expectation. Using the induction hypothesis in (5.29) then proves (5.28).

We apply Markov's inequality to obtain

$$\Pr\left(\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} > \epsilon\right) \leq \frac{\mathrm{E}[e(\mathbf{x}(t))]}{\epsilon \cdot e(\mathbf{x}(0))}\,.$$

Therefore, if

$$\frac{\mathrm{E}[e(\mathbf{x}(t))]}{\epsilon \cdot e(\mathbf{x}(0))} < \epsilon\,,$$

then it follows that

$$\Pr\left(\frac{e(\mathbf{x}(t))}{e(\mathbf{x}(0))} > \epsilon\right) < \epsilon\,. \tag{5.30}$$

By (5.28), we see that (5.30) holds if $(\mathrm{E}\left[r_{\text{step}}(Z)\right])^t < \epsilon^2$. Therefore, (5.12) holds. □

Proof of Lemma 20. We begin with the following inequality:

$$\left(\inf_{\mathbf{x}\in\mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})}\right) e(\mathbf{x}) \leq \mathbf{x}^T\mathbf{M}\mathbf{x} \leq \left(\sup_{\mathbf{x}\in\mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})}\right) e(\mathbf{x})\,. \tag{5.31}$$

Because $\mathbf{M}\mathbf{1}_n = 0$, we know that $\mathbf{M}(\mathbf{x} + c\mathbf{1}_n) = \mathbf{M}\mathbf{x}$ for all $c \in \mathbb{R}$. Therefore,

$$(\mathbf{x} + c\mathbf{1}_n)^T\mathbf{M}(\mathbf{x} + c\mathbf{1}_n) = \mathbf{x}^T\mathbf{M}\mathbf{x}\,, \tag{5.32}$$

which implies that

$$\frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})} = \frac{(\mathbf{x} - \bar{x}\mathbf{1}_n)^T\mathbf{M}(\mathbf{x} - \bar{x}\mathbf{1}_n)}{e(\mathbf{x})}\,.$$

However, $e(\mathbf{x}) = (\mathbf{x} - \bar{x}\mathbf{1}_n)^T(\mathbf{x} - \bar{x}\mathbf{1}_n)$, so

$$\frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})} = \frac{(\mathbf{x} - \bar{x}\mathbf{1}_n)^T\mathbf{M}(\mathbf{x} - \bar{x}\mathbf{1}_n)}{(\mathbf{x} - \bar{x}\mathbf{1}_n)^T(\mathbf{x} - \bar{x}\mathbf{1}_n)}\,. \tag{5.33}$$

The map $\mathbf{x} \mapsto \mathbf{x} - \bar{x}\mathbf{1}_n$ maps $\mathbb{R}^n$ onto $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \perp \mathbf{1}_n = 0\}$, the space of vectors that are orthogonal to $\mathbf{1}_n$, so applying (5.33) yields

$$\inf_{\mathbf{x}\in\mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})} = \inf_{\substack{\mathbf{x}\in\mathbb{R}^n \\ \mathbf{x}\perp\mathbf{1}_n}} \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{\mathbf{x}^T x} = \lambda_2(\mathbf{M})\,. \tag{5.34}$$

Similarly,

$$\sup_{\mathbf{x}\in\mathbb{R}^n} \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{e(\mathbf{x})} = \lambda_{\max}(\mathbf{M})\,. \tag{5.35}$$

Equation (5.17) follows from (5.31), (5.34), and (5.35). □

We now turn to proving Proposition 23. To do so, we first prove several lemmas.

Lemma 25. *The following lower bound for $C_1^O(\mathbf{x}, S)$ holds:*

$$-C_1^O(\mathbf{x}, S) \geq -\left(\frac{n - |S|}{|S|}\right)^{1/2} (-C_1^I(\mathbf{x}, S))^{1/2} (\mathbf{x}^T \mathbf{L} \mathbf{x})^{1/2}\,.$$

Proof. The proof involves two applications of the Cauchy–Schwarz inequality. Our first application of the Cauchy–Schwarz inequality is

$$\sum_{v \in S} (x_v - \bar{x}^S) \sum_{w \notin S} A_{vw}(x_v - x_w) \leq \left(\sum_{v \in S} (x_v - \bar{x}^S)^2\right)^{1/2} \left(\sum_{v \in S} \left(\sum_{w \notin S} A_{vw}(x_v - x_w)\right)^2\right)^{1/2}. \tag{5.36}$$

Our second application is to the sum $\left(\sum_{w \notin S} A_{vw}(x_v - x_w)\right)^2$ from the right-hand side of (5.36). We write

$$\left(\sum_{w \notin S} A_{vw}(x_v - x_w)\right)^2 \leq \left(\sum_{w \notin S} A_{vw}\right) \sum_{w \notin S} A_{vw}(x_v - x_w)^2 \leq (n - |S|) \sum_{w \notin S} A_{vw}(x_v - x_w)^2\,, \tag{5.37}$$

where the last inequality follows from $\sum_{w \notin S} A_{vw} \leq n - |S|$. The inequalities (5.36) and (5.37) imply that

$$-C_1^O(\mathbf{x}, S) \leq \left(\frac{n - |S|}{|S|}\right)^{1/2} \left(-C_1^I(\mathbf{x}, S)\right)^{1/2} \left(\sum_{v \in S} \sum_{w \notin S} A_{vw}(x_v - x_w)^2\right)^{1/2}. \tag{5.38}$$

Note that the sum

$$\sum_{v \in S} \sum_{w \notin S} A_{vw}(x_v - x_w)^2$$

includes at most one summand $(x_v - x_w)$ for each edge $\{i, j\}$. Therefore,

$$\sum_{v \in S} \sum_{w \notin S} A_{vw}(x_v - x_w)^2 \leq \mathbf{x}^T \mathbf{L} \mathbf{x}\,,$$

which yields the bound in the statement of the lemma when inserted into (5.38). □

Lemma 26. *Let $a_1, \ldots, a_k$ be real numbers; and let b_i and c_i be real numbers that satisfy the equation $a_i = b_i + c_i$, where $b_i \leq 0$ and $c_i \in \mathbb{R}$. Additionally, let w_i (with $i \in \{1, \ldots, k\}$) be* *non-negative weights such that* $\sum_{i=1}^{k} w_i = 1$*; and let* f *be a monotonically non-increasing, convex function such that* $f(0) = 0$ *and* $c_i \geq f(-b_i)$*. Lastly, let* $\delta > 0$*.*

We partition the set $\{i \in \{1, \ldots, k\} : a_i \leq 0\}$ *of indices into two sets,* A *and* B*, and define the functions*

$$H(i) := \begin{cases} a_i \,, & v \in A \\ b_i \,, & v \in B \\ 0 \,, & \textit{otherwise} \end{cases}$$

and

$$G(A, B) := \sum_{i=1}^{k} w_i H(i) \,.$$

The solution to the problem

$$\max_{A,B,\{a_i\},\{b_i\}} G(A, B) \tag{5.39}$$
$$\textit{such that} \qquad \sum_{i=1}^{k} w_i a_i \leq -\delta$$

is upper-bounded by the solution $\bar{b}$ *to*

$$\bar{b} + f(\bar{b}) = -\delta \,. \tag{5.40}$$

Proof. Because $b = a_i - c_i$ for all $i \in B$, we use the constraint in (5.39) to obtain

$$G(A, B) = \sum_{i \in A} w_i a_i + \sum_{i \in B} w_i b_i \leq -\delta - \sum_{i \in B} w_i c_i \,. \tag{5.41}$$

We then apply the assumption $c_i \geq f(-b_i)$ to (5.41) and use the result to obtain

$$b \leq -\delta - \sum_{i \in B} w_i f(-b_i) \,.$$

The function $-f$ is concave, so we apply Jensen's inequality to obtain

$$G(A, B) \leq -\delta - f\left(-\sum_{i \in B} w_i b_i\right) \,. \tag{5.42}$$

The expression on the right-hand side of (5.42) is largest when $B = \{1, \ldots, k\}$ because $-f$ is monotonically non-decreasing and positive. We thus obtain that

$$\sum_{i \in B} w_i b_i \leq -\delta - f\left(-\sum_{i \in B} w_i b_i\right). \tag{5.43}$$

Define $\bar{b} := \sum_{i \in B} w_i b_i$. The left-hand side of (5.43) is monotonically non-decreasing in $\bar{b}$, and the right-hand side is monotonically non-increasing in $\bar{b}$. Additionally, from the assumption that $f(0) = 0$, we see that (5.43) does not hold when we make the substitution $\bar{b} \to 0$. Therefore, the solution to

$$\bar{b} + f(\bar{b}) = -\delta \tag{5.44}$$

is the maximum value of $\bar{b}$ such that the inequality (5.43) holds. Therefore, $G(A, B) \leq \bar{b}$, where $\bar{b}$ solves (5.44). □

Lemma 27. *Suppose that $\mathcal{S}$ is a set of sets of nodes, where each set is of size ℓ. Additionally, suppose that $w(S)$ (with $S \in \mathcal{S}$) are non-negative weights such that $\sum_{S \in \mathcal{S}} w(S) = 1$. Let $\delta > 0$ satisfy*

$$\sum_{S \in \mathcal{S}} w(S) C_1(\mathbf{x}, S) \leq -\delta\,,$$

and let $X \leq 0$ solve the equation

$$X - \left(\frac{n-\ell}{\ell}\right)^{1/2} (-X)^{1/2} \left(\mathbf{x}^T \mathbf{L} \mathbf{x}\right)^{1/2} = -\delta\,.$$

It then follows that

$$\sum_{S \in \mathcal{S}} w(S) \mathcal{B}(\mathbf{x}, S) \leq X\,.$$

Proof. Once we translate notation, Lemma 27 is a direct application of Lemma 26. Let $\mathcal{S} = \{S_1, \ldots, S_k\}$; and let $A = \mathcal{S} \cap (\mathcal{S}^< \setminus \mathcal{S}^-)$, $B = \mathcal{S} \cap \mathcal{S}^-$, $a_i = X_1(\mathbf{x}, S_i)$, $b_i = X_1^I(\mathbf{x}, S_i)$, $c_i = X_1^o(\mathbf{x}, S_i)$, $w_i = w(S_i)$, and

$$f(b) = -\left(\frac{n-\ell}{\ell}\right)^{1/2} (-b)^{1/2} \left(\mathbf{x}^T \mathbf{L} \mathbf{x}\right)^{1/2}.$$

By Lemma 25, $c_i \geq f(-b_i)$ for all $i \in \{1, \ldots, k\}$.

From the assumption that

$$\sum_{i=1}^{k} w_i a_i \leq -\delta\,,$$

we see that the solution X to

$$X - \left(\frac{n-\ell}{\ell}\right)^{1/2} (-X)^{1/2} \left(\mathbf{x}^T \mathbf{L} \mathbf{x}\right)^{1/2} = -\delta$$

is an upper bound of $\sum_{S\in\mathcal{S}} w(S)\mathcal{B}(\mathbf{x}, S)$, as desired. □

Lemma 28. *Suppose that α solves the equation*

$$\alpha^2 + \alpha - \frac{\gamma}{4} = 0\,, \tag{5.45}$$

where $\gamma > 0$. It then follows that

$$\alpha \geq \left(\frac{\sqrt{2}-1}{2}\right) \min\left(\gamma, \gamma^{1/2}\right).$$

Proof. The solution to (5.45) is

$$\alpha = \frac{-1+\sqrt{1+\gamma}}{2}\,.$$

We consider two cases:

1. $\gamma \leq 1$. We have the estimate

   $$\sqrt{1+\gamma} \geq 1 + (\sqrt{2}-1)\gamma\,,$$

   which we use in the formula for α to obtain

   $$\alpha \geq \frac{\sqrt{2}-1}{2}\gamma\,.$$

2. $\gamma > 1$. We have the estimate

   $$\sqrt{1+\gamma} \geq 1 + (\sqrt{2}-1)\sqrt{\gamma}\,,$$

   which we use in the formula for α to obtain

   $$\alpha \geq \frac{\sqrt{2}-1}{2}\sqrt{\gamma}\,.$$

Because the claim holds in both cases, we've proven the lemma. □

Proof of Proposition 23. Suppose that X solves the equation

$$X - \left(\frac{n-\ell}{\ell}\right)^{1/2} (-X)^{1/2} \left(\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^{1/2} = -\delta \tag{5.46}$$

from Lemma 27.

When $\mathbf{x}^T\mathbf{L}\mathbf{x} = 0$, we obtain $X = -\delta$, as required, so we assume that $\mathbf{x}^T\mathbf{L}\mathbf{x} > 0$.

Let Y satisfy the equation $-X = \frac{n-\ell}{\ell}(\mathbf{x}^T\mathbf{L}\mathbf{x})Y^2$. In terms of Y, (5.46) is given by

$$-\frac{n-\ell}{\ell}(\mathbf{x}^T\mathbf{L}\mathbf{x})Y^2 - \frac{n-\ell}{\ell}(\mathbf{x}^T\mathbf{L}\mathbf{x})Y = -\delta\,,$$

from which it follows that

$$-Y^2 - Y = -\delta\frac{\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\,. \tag{5.47}$$

By Lemma 5.45, the solution Y to (5.47) satisfies

$$Y \geq \left(\frac{\sqrt{2}-1}{2}\right)\min\left(\delta\frac{4\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}, \left(\delta\frac{4\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right)^{1/2}\right),$$

which we rewrite as

$$Y \geq \left(\sqrt{2}-1\right)\left(\delta\frac{\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right)^{1/2}\min\left(1, \left(\delta\frac{4\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right)^{1/2}\right).$$

Because $-C = \frac{n-\ell}{\ell}(\mathbf{x}^T\mathbf{L}\mathbf{x})Y^2$, we obtain

$$\begin{aligned}-C &\geq \frac{n-\ell}{\ell}(\mathbf{x}^T\mathbf{L}\mathbf{x})\left(\sqrt{2}-1\right)^2\left(\delta\frac{\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right)\min\left(1, \delta\frac{4\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right)\\ &= \left(\sqrt{2}-1\right)^2\delta\min\left(1, \delta\frac{2\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\right).\end{aligned}$$

We now apply Lemma 27 to complete the proof of the proposition. □

Proof of Lemma 24. For fixed $i, j \in S$, exactly $|S| - 2$ distinct sets $S \setminus \{v\}$ include both nodes i and j. Therefore,

$$\sum_{v\in S} C_1^I(\mathbf{x}, S\setminus\{v\}) = -\frac{1}{2}\sum_{i,j\in S}\sum_{v\in S: i,j\in S\setminus\{v\}}(x_i - x_j)^2 = (|S|-2)C_1^I(\mathbf{x}, S)\,,$$

which is the first conclusion of the lemma.

We rewrite $\sum_{v\in S} C_1^O(\mathbf{x}, S\setminus\{v\})$ as follows:

$$\sum_{v\in S} C_1^O(\mathbf{x}, S\setminus\{v\})$$
$$= -\sum_{v\in S}\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})A_{iv}(x_i-x_v) - \sum_{v\in S}\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{ij}(x_i-x_j)\,. \quad (5.48)$$

Note that

$$\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})A_{iv}(x_i-x_v) = \sum_{i\in S}(x_i-\bar{x}^{S\setminus\{v\}})A_{iv}(x_i-x_v) \quad (5.49)$$

and

$$\sum_{v\in S}\sum_{i\in S} A_{iv}x_i(x_i-x_v) = \frac{1}{2}\sum_{v,i\in S} A_{iv}(x_i-x_v)^2\,. \quad (5.50)$$

Using (5.49) and (5.50), we obtain

$$\sum_{v\in S}\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})A_{iv}(x_i-x_v) = \frac{1}{2}\sum_{v,i\in S} A_{iv}(x_i-x_v)^2 + \sum_{v,i\in S} A_{iv}\bar{x}^{S\setminus\{v\}}(x_v-x_i)\,. \quad (5.51)$$

We then use

$$\bar{x}^{S\setminus\{v\}} = \frac{|S|}{|S|-1}\bar{x}^S - \frac{1}{|S|-1}x_v$$

to rewrite the last term in (5.51) as follows:

$$\sum_{v,i\in S} A_{iv}\bar{x}^{S\setminus\{v\}}(x_v-x_i) = -\frac{1}{|S|-1}\sum_{v,i\in S} A_{iv}x_v(x_v-x_i) = -\frac{1}{2}\frac{1}{|S|-1}\sum_{v,i\in S} A_{iv}(x_i-x_v)^2\,. \quad (5.52)$$

Combining (5.51) and (5.52) yields

$$\sum_{v\in S}\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})A_{iv}(x_i-x_v) = \frac{1}{2}\frac{|S|-2}{|S|-1}\sum_{v,i\in S} A_{iv}(x_i-x_v)^2\,. \quad (5.53)$$

We now rewrite the second term of the right-hand side of (5.50). First, we write the sum over $S\setminus\{v\}$ as follows:

$$\sum_{i\in S\setminus\{v\}} (x_i-\bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{ij}(x_i-x_j)$$
$$= \sum_{i\in S}(x_i-\bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{ij}(x_i-x_j) - (x_v-\bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{vj}(x_v-x_j)\,. \quad (5.54)$$

Because $\sum_{v\in S} \bar{x}^{S\setminus\{v\}} = |S|\bar{x}^S$, we find that

$$\sum_{v\in S}\sum_{i\in S}(x_i - \bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{ij}(x_i - x_j) = |S|C_1^O(\mathbf{x}, S)\,. \tag{5.55}$$

Additionally,

$$x_v - \bar{x}^{S\setminus\{v\}} = \frac{|S|}{|S|-1}(x_v - \bar{x}^S)\,,$$

which we use to obtain

$$\sum_{v\in S}(x_v - \bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{vj}(x_v - x_j) = \frac{|S|}{|S|-1}C_1^O(\mathbf{x}, S)\,. \tag{5.56}$$

Summing (5.54) over $v \in S$ and inserting (5.55) and (5.56) into the resulting expression, we obtain

$$\sum_{v\in S}\sum_{i\in S\setminus\{v\}}(x_i - \bar{x}^{S\setminus\{v\}})\sum_{j\notin S} A_{ij}(x_i - x_j) = \frac{|S|(|S|-2)}{|S|-1}C_1^O(\mathbf{x}, S)\,. \tag{5.57}$$

Equation (5.27) results form inserting (5.53) and (5.57) into (5.48). □

CHAPTER 6

Convergence to Consensus with Uniform-at-Random Meetings

Let U_ℓ denote the uniform random variable on the subsets of V of size ℓ. We seek to determine an upper bound for the ϵ-convergence time $T(\epsilon, U_\ell)$. As a result of Proposition 16, the ϵ-convergence time $T(\epsilon, U_\ell)$ satisfies

$$T(\epsilon, U_\ell) \leq \frac{2\log(\epsilon)}{\log(\mathrm{E}\left[\hat{r}_{\text{step}}(U_\ell)\right])},$$

where $\mathrm{E}\left[\hat{r}_{\text{step}}(U_\ell)\right]$ is an upper bound for the expected per-step convergence factor $\mathrm{E}\left[r_{\text{step}}(U_\ell)\right]$ (see Definition 12 in Chapter 5). We need to determine an upper bound $\mathrm{E}\left[\hat{r}_{\text{step}}(U_\ell)\right]$ to prove an upper bound for $T(\epsilon, U_\ell)$.

We consider some simple graphs in the following examples to gain intuition about $T(\epsilon, U_\ell)$.

Example 29. *Suppose that G is either the n-node complete graph or the graph of n disconnected nodes. In Example 18, we determined for both of these cases that*

$$1 - \imath_{\text{step}}(\mathbf{x}, S) = \frac{\sum_{v\in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})} \tag{6.1}$$

for all $S \subseteq V$. We calculate that

$$\mathrm{E}\left[\sum_{v\in S}(x_v - \bar{x}^S)^2\right] = \frac{1}{2\ell}\mathrm{E}\left[\sum_{v,w\in S}(x_v - x_w)^2\right] = \frac{1}{2\ell}\frac{\binom{n-2}{\ell-2}}{\binom{n}{\ell}}\sum_{v,w\in V}(x_v - x_w)^2 = \frac{\ell-1}{n-1}e(\mathbf{x}),$$

where we use Lemma 5 in the last equality. Therefore, for all $\mathbf{x} \in \mathbb{R}^n$, it follows that

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] = \frac{\ell-1}{n-1}, \tag{6.2}$$

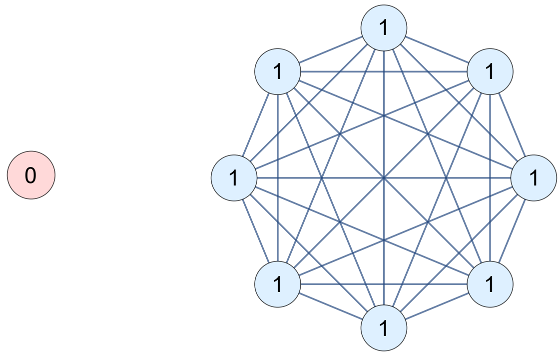


Figure 6.1: The $(n-1)$-to-1 graph for $n = 9$ and the beliefs of Example 30.

which implies that $\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] = (\ell - 1)/(n-1)$*. For all* $\epsilon \in (0,1)$*, it then follows from Proposition 16 that*

$$T(\epsilon, U_\ell) \leq \frac{2\log(\epsilon)}{\log\left(1 - \frac{\ell-1}{n-1}\right)} \tag{6.3}$$

Using (5.14) *from Remark 17, this upper bound satisfies*

$$\frac{2\log(\epsilon)}{\log\left(1 - \frac{\ell-1}{n-1}\right)} = 2\log(\epsilon^{-1})\frac{n}{\ell} + O(1)$$

as $\ell/n \to 0$*. Therefore, the* ϵ*-convergence time increases approximately as* n/ℓ*.*

The per-step convergence factor $r_{\text{step}}(\mathbf{x}, S)$ is maximized if $\beta(\mathbf{x}, S) = 1$. Because $\beta(\mathbf{x}, S) = 1$ for all $\mathbf{x} \in \mathbb{R}^n$ and $S \subseteq V$ when G is the complete graph or the graph of disconnected nodes (see Example 18), we deduce that $\mathrm{E}\left[r_{\text{step}}(U_\ell)\right]$ is maximal among undirected graphs for these two graphs. Therefore, the bound (6.3) from Example 29 is the minimal bound among undirected graphs of n nodes that one can obtain using Proposition 16. In the next example, we present a graph with a per-step convergence factor of U_ℓ that cannot be bounded by a term that is linear in ℓ/n.

Example 30. *Let* G *be a graph (see Figure 6.1) that consists of a complete graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *of* $n-1$ *nodes and a single disconnected node* v*. We call this graph the* $(n-1)$-to-1 graph*. Let* $\mathbf{x}_0 \in \mathbb{R}^n$*, and let the* w^{th} *component of* $\mathbf{x}_0$ *be equal to* 0 *if* $w \neq v$ *and be equal to* 1 *if*

$w = v$. For all sets S of nodes that do not include v, the per-step convergence factor is 0 *because the set of nodes is already at consensus. Assume that S includes v and $\ell - 1$ other nodes. We calculate that $C_1^I(\mathbf{x}_0, S) = (\ell - 1)$, $C_1^O(\mathbf{x}_0, S) = 0$, $C_2(\mathbf{x}_0, S) = \frac{\ell-1}{\ell^2}(n + \ell(\ell-1))$, $e(\mathbf{x}_0) = (n-1)/n$, and*

$$\beta^*(t) = \min\left(1, \frac{\ell^2}{n + \ell(\ell-1)}\right) = \frac{\ell^2}{n + \ell(\ell-1)}\,.$$

Therefore,

$$1 - r_{\text{step}}(\mathbf{x}_0, S) = \left(2 - \frac{\ell^2}{n + \ell(\ell-1)}\right)\frac{\ell(\ell-1)}{n + \ell(\ell-1)}\frac{n}{n-1}\,.$$

The fraction of sets of size ℓ that include v is ℓ/n. Consequently, $\mathrm{E}\left[1 - r_{\text{step}}(\mathbf{x}_0, S)\right]$ *is given by*

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathbf{x}_0, U_2)\right] = \left(2 - \frac{\ell^2}{n + \ell(\ell-1)}\right)\frac{\ell(\ell-1)}{n + \ell(\ell-1)}\frac{n}{n-1}\frac{\ell}{n}\,. \tag{6.4}$$

With $\ell = 2$, using (6.4) yields the following estimate:

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathbf{x}_0, U_2)\right] = \frac{8n}{(n-1)(n+2)^2} \le \frac{12}{n^2}\,.$$

Therefore, the per-step convergence factor satisfies

$$\mathrm{E}\left[r_{\text{step}}(\mathbf{x}_0, U_2)\right] \ge 1 - \frac{12}{n^2}\,, \tag{6.5}$$

and hence $\mathrm{E}\left[r_{\text{step}}(U_2)\right] \ge 1 - \frac{12}{n^2}$. From (6.5), it is not possible for the per-step convergence factor of U_ℓ to be linear in ℓ/n. The relatively large size of the per-step convergence factor of U_2 in (6.5) results in a relatively large bound for $T(\epsilon, U_2)$ when using Proposition 16. It follows from (6.5) that

$$T(\epsilon, U_2) \le \frac{2\log(\epsilon)}{\log(1 - \frac{12}{n^2})}\,. \tag{6.6}$$

The upper bound in (6.6) satisfies

$$\frac{2\log(\epsilon)}{\log(1 - \frac{12}{n^2})} = 24\log(\epsilon)n^2 + O(1)$$

as $n \to \infty$ by (5.15) in Remark 17.

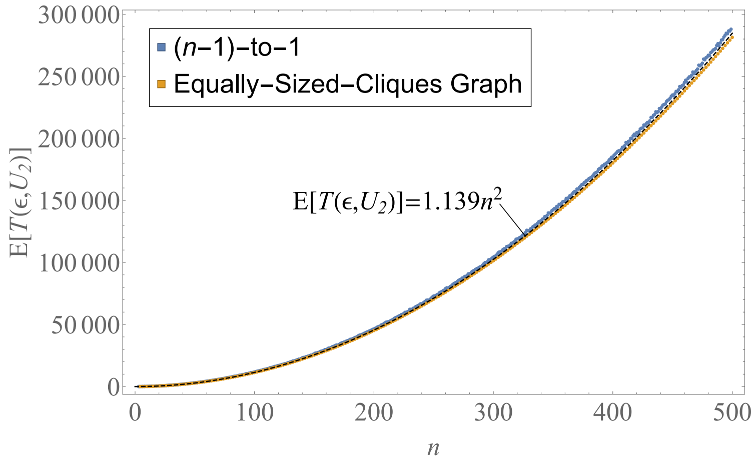


Figure 6.2: Simulated expected value of $T(\epsilon, U_\ell)$ with $\epsilon = .0001$ for the $(n-1)$-to-1 graph in Figure 6.1 and the graph of two equally-sized cliques in Figure 2.2. The black dotted curve is a quadratic that we plot to show the dependence of the ϵ-convergence times on n (the number of nodes).

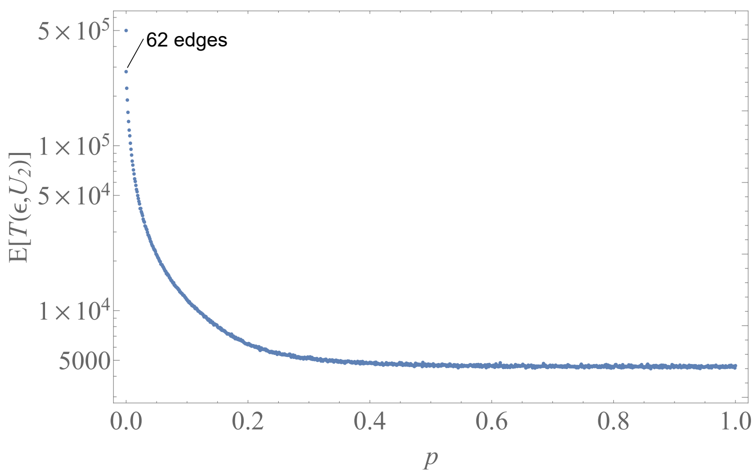


Figure 6.3: Simulated expected value of $T(\epsilon, U_\ell)$ with $\epsilon = .0001$ for graphs that we obtain by starting with two 250-node complete graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ and adding an edge with probability p for each pair of nodes $v \in V_1$ and $w \in V_2$. We label the data point that corresponds to the graph with 62 expected edges $\{v, w\}$ such that $v \in V_1$ and $w \in V_2$.

We show plots in Figure 6.2 of simulated expected ϵ-convergence times with $\ell = 2$ and $\epsilon = .0001$ for the $(n-1)$-to-1 graph of Example 30 and for the graph of two equally-sized cliques (see Figure 2.2). Our procedure for simulating expected ϵ-convergence times is to (1) record the number of meetings until $e(\mathbf{x}(t)) < \epsilon$ for 100 simulations of the group-soft-consensus model (3.5) with uniform-at-random choice of meeting-groups and then (2) take the mean of the recorded numbers of meetings. We run each simulation with an initial belief vector $\mathbf{x}_0$ for which nodes in the same connected component have the same belief, but nodes in different connected components have different beliefs. In Figure 6.2, we plot the expected ϵ-convergence time for various values of n. As shown by the curve $\mathrm{E}\left[T(\epsilon, U_2)\right] = 1.139n^2$, which we obtain by fitting a polynomial to the simulation data, the simulated expected ϵ-convergence time is approximately quadratic in n for both graphs. In Example 30, we concluded that we cannot determine an upper bound for the ϵ-convergence time of the form n^k for $k < 2$ using Proposition 16. This result and the quadratic dependence on n in the simulation in Figure 6.2 provide evidence that the worst-case ϵ-convergence time is at least $O(n^2)$ for $\ell = 2$.

We consider whether the quadratic dependence on n for graphs with more than 1 connected component that we found in Example 30 and in our simulations in Figure 6.2 also occurs in graphs with a single connected component. Let n be a positive, even integer. Let $G(p)$ be a graph that we obtain by taking two complete graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, each of $n/2$ nodes, and independently adding an edge $\{v, w\}$ with probability p between each pair of nodes $v \in V_1$ and $w \in V_2$. In Figure 6.3, we show a simulated expected value of the ϵ-convergence time for various values of p. For each value of p, we use the same procedure to simulate expected ϵ-convergence times that we used to obtain the simulated expected ϵ-convergence times in Figure 6.2, except that we randomly generate a new graph $G(p)$ at the beginning of the 100 simulations. For each simulation, we use the values $\epsilon = .0001$ and $n = 500$. Our simulations yield an expected ϵ-convergence time of 281,398 for the graph of two equally-sized cliques; whereas for $G(p)$, with p selected so that the expected number of edges in $G(p)$ is equal to 62, they yield expected ϵ-convergence time of 223,280 (see the data point labeled "62 edges" in Figure 6.2). By contrast, we find an expected ϵ-convergence time of 4,635 for the complete graph ($G(p)$ with $p = 1$) with our simulations. Even with p chosen so that hundreds of edges are expected between G_1 and G_2, we observe simulated expected ϵ-convergence time of more than 100,000.

We now prove an upper bound for the ϵ-convergence time. To begin, we use (5.25) to bound $\mathrm{E}\left[\imath_{\text{step}}(\mathbf{x}, U_\ell)\right]$ as follows:

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] \geq -\frac{\mathrm{E}\left[\mathcal{B}(\mathbf{x}, U_\ell)\right]}{n \cdot e(\mathbf{x})}. \tag{6.7}$$

If we can determine $\mathrm{E}\left[C_1(\mathbf{x}, U_\ell)\right]$, then we can use Proposition 23 to determine an upper bound of $\mathrm{E}\left[\mathcal{B}(\mathbf{x}, U_\ell)\right]$ by a negative number. As a consequence of (6.7), we then obtain a lower bound for $\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right]$. We begin by calculating $\mathrm{E}\left[C_1(\mathbf{x}, U_\ell)\right]$.

Lemma 31. *The following equalities hold:*

$$\mathrm{E}\left[C_1^I(\mathbf{x}, U_\ell)\right] = -\frac{\ell(\ell-1)}{(n-1)} e(\mathbf{x}), \tag{6.8}$$

$$\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] = -\frac{(n-\ell)(\ell-1)}{n(n-1)} \mathbf{x}^T \mathbf{L} \mathbf{x}. \tag{6.9}$$

Combining the two equalities in Lemma 31, we obtain

$$\mathrm{E}\left[C_1(\mathbf{x}, U_\ell)\right] = -\frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n} \mathbf{x}^T \mathbf{L} \mathbf{x}\right). \tag{6.10}$$

By (6.10), if we set

$$\delta = \frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n} \mathbf{x}^T \mathbf{L} \mathbf{x}\right),$$

then (5.26) from Proposition 23 holds. Therefore, by Proposition 23, we obtain the following bound:

$$\begin{aligned} \mathrm{E}\left[\mathcal{B}(\mathbf{x}, S)\right] \leq & -\left(\sqrt{2}-1\right)^2 \frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n} \mathbf{x}^T \mathbf{L} \mathbf{x}\right) \\ & \times \min\left(1, \frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n} \mathbf{x}^T \mathbf{L} \mathbf{x}\right) \frac{4\ell}{n-\ell} \frac{1}{\mathbf{x}^T \mathbf{L} \mathbf{x}}\right). \end{aligned} \tag{6.11}$$

Let $\ell = \ell(n, \mathbf{x})$ be the solution of the equation

$$\frac{4\ell(\ell-1)}{(n-\ell)(n-1)}\left(\ell\frac{e(\mathbf{x})}{\mathbf{x}^T\mathbf{L}\mathbf{x}}+\frac{n-\ell}{n}\right)=1 \tag{6.12}$$

for $\mathbf{x}^T\mathbf{L}\mathbf{x} > 0$, and define $\ell(n, \mathbf{x}) = 0$ for $\mathbf{x}^T\mathbf{L}\mathbf{x} = 0$. We use $\ell(n, \mathbf{x})$ to obtain

$$\mathrm{E}\left[\mathcal{B}(\mathbf{x}, U_\ell)\right] \leq \begin{cases} -\left(\sqrt{2}-1\right)^2 \frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n}\mathbf{x}^T\mathbf{L}\mathbf{x}\right), & n-1 \geq \ell \geq \ell(n,\mathbf{x}) \\ -4\left(\sqrt{2}-1\right)^2 \left(\frac{\ell-1}{n-1}\right)^2 \frac{\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n}\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2, & \ell(n,\mathbf{x}) > \ell \geq 2\,. \end{cases} \tag{6.13}$$

from (6.11).

Inserting (6.13) into (5.27), we obtain

$$\begin{aligned} &\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] \\ &\geq \frac{1}{n}\begin{cases} \left(\sqrt{2}-1\right)^2 \frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n}\mathbf{x}^T\mathbf{L}\mathbf{x}\right), & n-1 \geq \ell \geq \ell(n,\mathbf{x}) \\ 4\left(\sqrt{2}-1\right)^2 \left(\frac{\ell-1}{n-1}\right)^2 \frac{\ell}{n-\ell}\frac{1}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\left(\ell e(\mathbf{x}) + \frac{n-\ell}{n}\mathbf{x}^T\mathbf{L}\mathbf{x}\right)^2, & \ell(n,\mathbf{x}) > \ell \geq 2\,. \end{cases} \end{aligned} \tag{6.14}$$

We study the function $\ell(n, \mathbf{x})$ to understand the bound (6.14) further. Denote $s := \ell/n$ and $\alpha := \mathbf{x}^T\mathbf{L}\mathbf{x}/(ne(\mathbf{x}))$. Note that $2 \leq \ell \leq n-1$ implies that $2/n \leq s \leq (n-1)/n$ and that $0 \leq \mathbf{x}^T\mathbf{L}\mathbf{x} \leq ne(\mathbf{x})$ implies that $0 \leq \alpha \leq 1$. We write the equation for $\ell(n, \mathbf{x})$ in (6.12) in terms of s, α, and n to obtain

$$\frac{4s(sn-1)}{(1-s)(n-1)}\left(\frac{s}{\alpha}+(1-s)\right)=1\,. \tag{6.15}$$

Let $s(n, \alpha)$ denote the solution for s to (6.15) for $n \in \{3, 4, \ldots\}$ and $\alpha \in (0, 1]$, and define $s(n, 0) = 0$. For $\alpha \in (0, 1]$, let $s(\infty, \alpha)$ denote the solution for s of the equation

$$\frac{4s^2}{1-s}\left(\frac{s}{\alpha}+(1-s)\right)=1\,,$$

which we obtain by letting $n \to \infty$ in (6.15). For $\alpha = 0$, define $s(\infty, \alpha) = 0$. By explicitly calculating derivatives, we see that the function $s(n, \alpha)$ is monotonically non-increasing in n and monotonically non-decreasing in α for $(n, \alpha) \in \{(n, \alpha) : n \geq 3, \alpha \in [0, 1]\}$. Additionally,

$s(\infty, \alpha)$ is monotonically non-decreasing in α for all $\alpha \in [0,1]$ and $s(\infty,\alpha) < s(n,\alpha)$ for all $n \in \{3, 4, \ldots\}$. By rewriting the bound (6.14) in terms of s and α, obtain

$$\begin{aligned} &\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] \\ &\quad \geq \begin{cases} \left(\sqrt{2}-1\right)^2 \frac{s-\frac{1}{n}}{1-\frac{1}{n}} \left(s + (1-s)\alpha\right), & (n-1)/n \geq s \geq s(n,\alpha) \\ 4\left(\sqrt{2}-1\right)^2 \left(\frac{s-\frac{1}{n}}{1-\frac{1}{n}}\right)^2 \frac{s}{\alpha} \left(s + (1-s)\alpha\right) \left(\frac{s}{1-s} + \alpha\right), & s(n,\alpha) \geq s \geq 2/n\,. \end{cases} \end{aligned} \tag{6.16}$$

To understand the behavior of the bound (6.16), we segment (α, s)-space into three regions (see Figure 6.4):

- Region I $= \{(\alpha, s) : \alpha \in (0,1], s \geq s(3,\alpha)\}$
- Region II $= \{(\alpha, s) : \alpha \in (0,1], s(3,\alpha) > s > s(\infty,\alpha)\}$
- Region III $= \{(\alpha, s) : \alpha \in (0,1], s(\infty,\alpha) \geq s\}$

Because $s(n,\alpha)$ is monotonically non-increasing in n, we have that $s(3,\alpha) \geq s(n,\alpha)$ for all $n \in \{3,4,\ldots\}$ and $\alpha \in (0,1]$. It follows that $s \geq s(n,\alpha)$ for all $n \in \{3,4,\ldots\}$ in region I. Therefore, the top case of (6.16) holds in region I. In region II, the value of n determines which of the two cases of (6.16) holds. Because $s(n,\alpha) > s(\infty,\alpha)$ for all $\alpha \in (0,1]$, we have that $s(n,\alpha) > s$. Therefore, $s(\infty,\alpha) \geq s$ in region III, so the bottom case of (6.16) holds in region III.

The bound (6.16) depends on n. We obtain a bound that is independent of n by considering the limit $n \to \infty$ of (6.16).

Proposition 32. *For all $n \geq 3$, the expectation of the per-step convergence factor of $\mathbf{x}$ and U_ℓ satisfies the following bound:*

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] + \frac{c}{n} \geq \begin{cases} \left(\sqrt{2}-1\right)^2 s\left(s + (1-s)\alpha\right), & 1 > s \geq s(\infty,\alpha) \\ 4\left(\sqrt{2}-1\right)^2 \frac{s^3}{\alpha} \left(s + (1-s)\alpha\right) \left(\frac{s}{1-s} + \alpha\right), & s(\infty,\alpha) > s > 0\,, \end{cases} \tag{6.17}$$

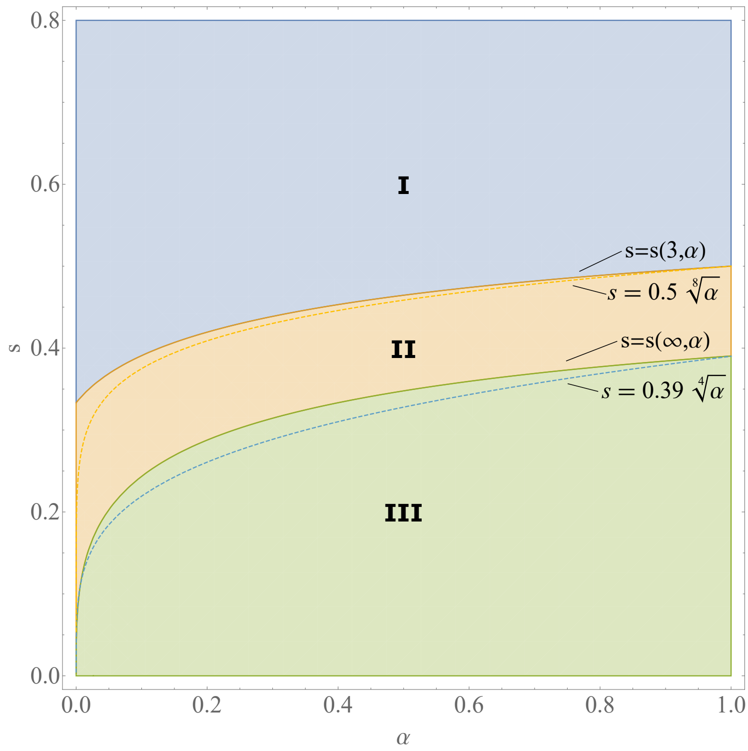


Figure 6.4: Regions I–III that we use to understand the bound (6.16) (see the main text for details). We include the curves $s = 0.5\sqrt[8]{\alpha}$ and $s = 0.39\sqrt[4]{\alpha}$ to illustrate the shape of the curves $s = s(3, \alpha)$ and $s = s(\infty, \alpha)$, respectively.

where $c > 0$ is a constant and $s(\infty, \alpha)$ is the solution for s of the equation

$$\frac{4s^2}{1-s}\left(\frac{s}{\alpha} + (1-s)\right) = 1\,,$$

if $\alpha \in (0, 1]$ and $s(\infty, \alpha) = 0$ if $\alpha = 0$.

The bound (6.17) from Proposition 32 simplifies greatly when G is the complete graph.

Example 33. *Suppose that G is the complete graph with n nodes. Applying Lemma 20 and using the fact that $\lambda_2(\mathbf{L}_n) = \lambda_{\max}(\mathbf{L}_n) = n$, we obtain $\alpha = 1$. Therefore, by Proposition 32, the expectation of the per-step convergence factor of U_ℓ is bounded for some constant $c > 0$*

as

$$\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] + \frac{c}{n} \geq \begin{cases} \left(\sqrt{2}-1\right)^2 s\,, & 1 > s \geq s(\infty,1) \\ 4\left(\sqrt{2}-1\right)^2 \frac{s^3}{1-s}\,, & s(\infty,1) > s > 0\,. \end{cases} \tag{6.18}$$

The dependence on s of (6.18) is the same as (6.2) from Example 29 when $s \geq s(\infty,1)$. However, the bound (6.2) is a much better bound for $s < s(\infty,1)$. Therefore, the bound (6.17) from Proposition 32 is not sharp.

We can use Proposition 32 to derive an upper bound for $T(\epsilon, U_\ell)$. However, we can obtain a better upper bound by combining Proposition 32 with another bound that we now derive.

Recall from Chapter 5 that $\mathcal{S}^< := \{S \subseteq V : C_1(\mathbf{x}, S) < 0\}$ and $\mathcal{S}^- := \{S \in \mathcal{S}^< : C_1^O(\mathbf{x}, S) < 0\}$. Let

$$\mathrm{E}^<\left[f(U_\ell)\right] := \mathrm{E}\left[f(U_\ell)\mathbb{1}(U_\ell \in \mathcal{S}^<)\right]$$

and

$$\mathrm{E}^-\left[f(U_\ell)\right] := \mathrm{E}\left[f(U_\ell)\mathbb{1}(U_\ell \in \mathcal{S}^-)\right].$$

Applying the inequalities

$$\mathcal{B}(\mathbf{x}, S) \geq C_1(\mathbf{x}, S)\mathbb{1}\left(S \in \mathcal{S}^< \setminus \mathcal{S}^-\right)$$

and $\mathrm{E}^<\left[-C_1(\mathbf{x}, U_\ell)\right] \geq \mathrm{E}\left[-C_1(\mathbf{x}, U_\ell)\right]$ to the bound (5.25), we obtain

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] \geq \frac{1}{n}\frac{\mathrm{E}\left[-C_1(\mathbf{x}, U_\ell)\right] - \mathrm{E}^-\left[-C_1^O(\mathbf{x}, U_\ell)\right]}{e(\mathbf{x})}\,. \tag{6.19}$$

Alternatively, we can use the bound

$$\mathcal{B}(\mathbf{x}, S) \geq C_1^I(\mathbf{x}, S)\mathbb{1}\left(S \in \mathcal{S}^-\right)$$

in (5.25) to obtain

$$\mathrm{E}\left[1 - \imath_{\text{step}}(\mathbf{x}, U_\ell)\right] \geq \frac{1}{n}\frac{\mathrm{E}^-\left[C_1^I(\mathbf{x}, U_\ell)\right]}{e(\mathbf{x})}\,. \tag{6.20}$$

Using the two bounds (6.19) and (6.20), we deduce the bound for the per-step convergence factor of U_ℓ that we state in the following proposition.

Proposition 34. *The expected value of* $r_{\text{step}}(U_\ell)$ *satisfies the bound*

$$\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] \geq \gamma_r^2 \frac{\ell^2(\ell-1)^2}{n^2(n-1)^2}\,,$$

where

$$\gamma_r = \left(1 + \sqrt{\frac{\ell(\ell-1)}{n(n-1)}}\right)^{-1}\,.$$

Corollary 35. *The per-step convergence factor of* U_ℓ *satisfies*

$$\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] \geq \frac{1}{4}s^4 - \frac{c}{n} \tag{6.21}$$

for some constant $c > 0$.

Proof. We note that

$$\gamma_r = \left(1 + \sqrt{s\frac{s+1/n}{1+1/n}}\right)^{-1} \geq \frac{1}{2}$$

and that there exists a constant $c > 0$ such that

$$\frac{\ell^2(\ell-1)^2}{n^2(n-1)^2} = s^2\frac{(s-1/n)^2}{(1-1/n)^2} \geq s^2(s-1/n)^2 \geq s^4 - \frac{4c}{n}\,.$$

We combine these these two inequalities to complete the proof. □

Using the bound (6.17) from Proposition 32, we obtain the following bound:

$$\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] + \frac{c}{n} \geq \begin{cases} \left(\sqrt{2}-1\right)^2 s^2\,, & 1 > s \geq s(\infty,1) \\ 4\left(\sqrt{2}-1\right)^2 \frac{s^5}{1-s}\,, & s(\infty,1) > s > 0 \end{cases}$$

for a constant $c > 0$. Combining this lower bound with the lower bound in Corollary 35, we determine that

$$\mathrm{E}\left[1 - r_{\text{step}}(U_\ell)\right] + \frac{c}{n} \geq \begin{cases} \left(\sqrt{2}-1\right)^2 s^2\,, & 1 > s \geq s(\infty,1) \\ \frac{1}{4}s^4, & s(\infty,1) > s > 0 \end{cases} \tag{6.22}$$

for a constant $c > 0$. Using (6.22), we obtain the ϵ-convergence time bound in the following theorem.

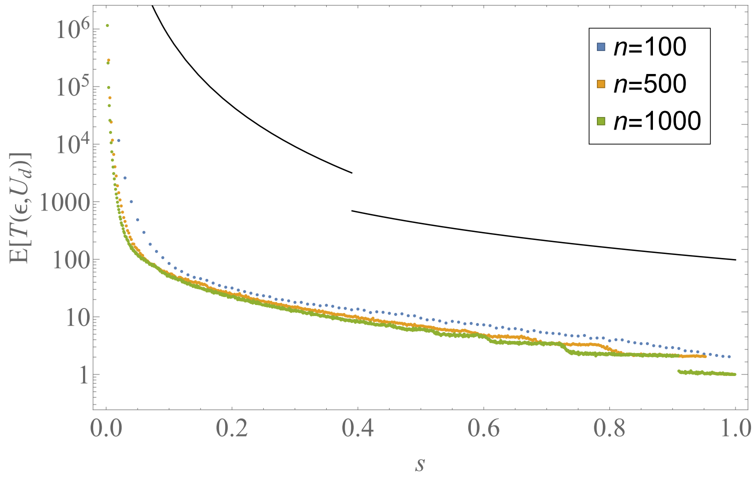


Figure 6.5: Simulated expected value of the ϵ-convergence time with $\epsilon = 0.0001$ for the $(n-1)$-to-1 graph for $n = 100$, $n = 500$, and $n = 1000$. In each simulation, the component of the initial vector $\mathbf{x}_0$ that corresponds to the isolated node is equal to 1 and all other components are equal to 0. The black curve is the bound from Theorem 36. We hypothesize that the irregularity in the curves of simulation data near $s = 1$ arises from finite-size effects.

Theorem 36. *There exists a constant $c > 0$ such that for all $n \geq 3$ and all $\epsilon \in (0, 1)$, the ϵ-convergence time $T(\epsilon, U_\ell)$ is bounded as*

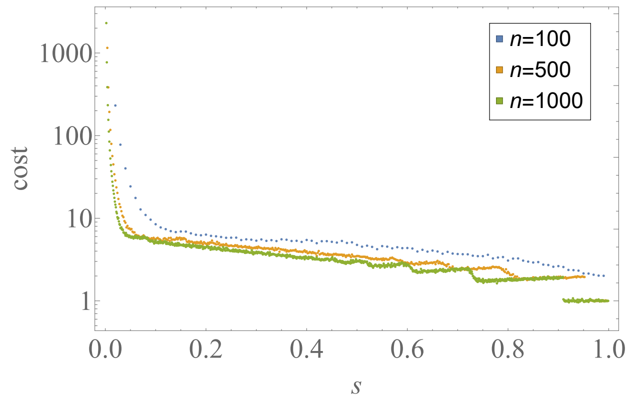
$$T(\epsilon, U_\ell) \leq \begin{cases} \frac{2\log(\epsilon)}{\log\left(1+\frac{c}{n}-(\sqrt{2}-1)^2 s^2\right)}\,, & 1 > s \geq s(\infty, 1) \\ \frac{2\log(\epsilon)}{\log\left(1+\frac{c}{n}-\frac{1}{4}s^4\right)}\,, & s(\infty, 1) > s > 0\,, \end{cases}$$

where

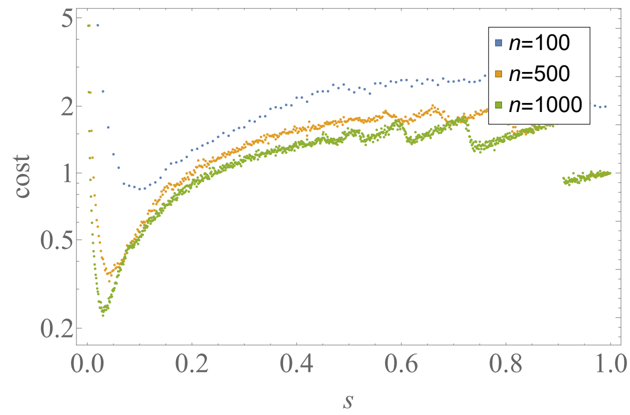
$$s(\infty, 1) = \frac{1}{8}\left(\sqrt{17}-1\right).$$

Proof. We obtain this bound by using (6.22) in Proposition 16. □

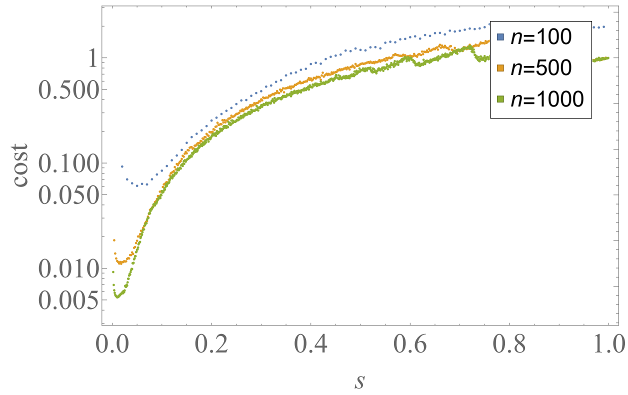
In Figure 6.5, we plot simulated expected values of ϵ-convergence time as a function of s for the $(n-1)$-to-1 graph for $n = 100$, $n = 500$, and $n = 1000$. Our procedure for obtaining these simulated exected values is the same as for the simulated expected values in Figure 6.2. We also plot the bound of Theorem 36.



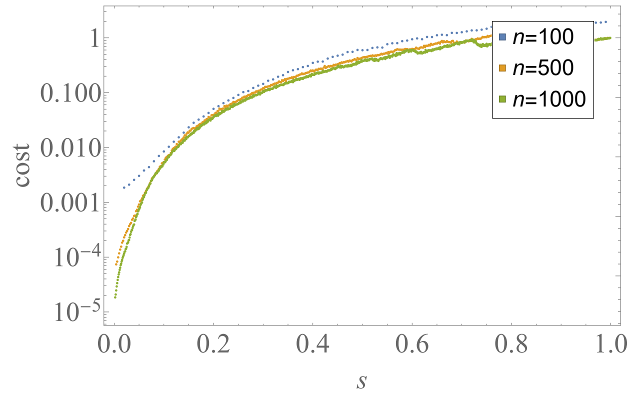
(a) $\text{cost}(S) = (|S|/n)$



(b) $\text{cost}(S) = (|S|/n)^2$



(c) $\text{cost}(S) = (|S|/n)^3$



(d) $\text{cost}(S) = (|S|/n)^4$

Figure 6.6: Simulated expected value of the ϵ-convergence costs for $\text{cost}(S) = (|S|/n)^k$ with $k \in \{1, 2, 3, 4\}$. We carry out simulations on the $(n-1)$-to-1 graph (see Figure 6.1) for $n = 100$, $n = 500$, and $n = 1000$. In each simulation, the component of the initial vector $\mathbf{x}_0$ that corresponds to the isolated node is equal to 1 and all other components are equal to 0. We use the value $\epsilon = .0001$.

We end this chapter with a toy application of our model to determining the optimal size of a meeting in an organization. Researchers in organizational science have identified the size of a meeting as a key determinant of meeting success [DS82, DM94, RN01]. On the one hand, Michael Doyle and David Straus [DS82] proposed that meeting sizes of two to seven are optimal for problem solving and decision-making meetings. On the other hand, Jeanine Drew and the 3M Management Team [DM94] suggested that five or fewer is optimal for the same types of meetings. To complement this prior empirical evidence on optimal meeting sizes, we use simulations of the group-soft-consensus model (3.5) to investigate which meeting sizes are optimal. Assume that the cost of a meeting takes the form $\text{cost}(S) = (|S|/n)^k$ for $k \in \mathbb{Z}_{>0}$. We define the *ϵ-convergence cost* as the total cost of the meetings that is required for the distance from consenus to become no larger than ϵ. Assuming that each meeting is of size ℓ, the ϵ-convergence cost is equal to $\left(\frac{\ell}{n}\right)^k T(\epsilon, U_\ell)$. In Figure 6.6, we plot a simulated expected value of the ϵ-convergence cost for $k \in \{1, 2, 3, 4\}$. Our simulation procedure is the same as for Figure 6.2. For $k = 1$, our simulations suggest that larger meetings are better. For $k = 2$ and $k = 3$, our simulations suggest an optimal meeting-size that is small but larger than 2. For $k = 4$, our simulations suggest that dyadic meetings are optimal.

Additional Proofs

Proof of Lemma 31. For each node pair $v, w \in V$, there exist $\binom{n-2}{\ell-2}$ subsets S of size ℓ in V that include both v and w. Using the definition of expected value, we obtain

$$\mathrm{E}\left[C_1^I(\mathbf{x}, U_\ell)\right] = -\frac{1}{\binom{n}{\ell}}\binom{n-2}{\ell-2}\frac{1}{2}\sum_{i,j=1}^{n}(x_v - x_w)^2 = -\frac{1}{2}\frac{\ell(\ell-1)}{n(n-1)}\sum_{i,j=1}^{n}(x_v - x_w)^2 \,.$$

This proves (6.8).

We prove (6.9) by deriving and solving a recurrence relation. The expected value $\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right]$ is

$$\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] = -\frac{1}{\binom{n}{\ell}}\sum_{\substack{S\subseteq V\\|S|=\ell}}\sum_{v\in S}(x_v - \bar{x}^S)\sum_{w\notin S}A_{vw}(x_v - x_w)\,.$$

We rewrite this expectation in terms of the sum over sets $S \subseteq V$ of size $\ell + 1$ to obtain

$$\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] = -\frac{1}{\binom{n}{\ell}}\frac{1}{n-\ell}\sum_{\substack{S\subseteq V\\|S|=\ell+1}}\sum_{v\in S}\sum_{v\in S\setminus\{v\}}(x_v - \bar{x}^{S\setminus\{v\}})\sum_{w\notin S\setminus\{v\}}A_{vw}(x_v - x_w)\,,$$

where the factor $\frac{1}{n-\ell}$ occurs because there are $n-\ell$ subsets of V of size $\ell+1$ that include a fixed set of size ℓ. Using the definition of $C_1^O(\mathbf{x}, S \setminus \{v\})$, we thus obtain

$$\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] = \frac{1}{\binom{n}{\ell}} \frac{1}{n-\ell} \sum_{\substack{S \subseteq V \\ |S|=\ell+1}} \sum_{v \in S} C_1^O(\mathbf{x}, S \setminus \{v\}) \,. \tag{6.23}$$

Using Lemma 24 in (6.23), we obtain

$$\begin{aligned} &\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] \\ &= \frac{1}{\binom{n}{\ell}} \frac{1}{n-\ell} \sum_{\substack{S \subseteq V \\ |S|=\ell+1}} \left(-\frac{\ell-1}{\ell} \sum_{v,w \in S} A_{vw}(x_v - x_w)^2 + \frac{(\ell+1)(\ell-1)}{\ell} C_1^O(\mathbf{x}, S) \right) . \end{aligned} \tag{6.24}$$

We write the first term on the right-hand side of (6.24) as

$$-\frac{1}{\binom{n}{\ell}} \frac{1}{n-\ell} \sum_{\substack{S \subseteq V \\ |S|=\ell+1}} \sum_{v,w \in S} A_{vw}(x_v - x_w)^2 = -\frac{\ell-1}{n(n-1)} \sum_{i,j=1}^{n} A_{vw}(x_v - x_w)^2 \,, \tag{6.25}$$

where we use the fact that there are $\binom{n-2}{\ell-1}$ sets of size $\ell+1$ that include both nodes v and w and that $\binom{n-2}{\ell-1} / \left(\binom{n}{\ell}(n-\ell)\right) = (\ell-1)/(n(n-1))$.

We write the second term on the right-hand side of (6.24) as $\mathrm{E}\left[C_1^O(\mathbf{x}, U_{\ell+1})\right]$ to obtain

$$\frac{1}{\binom{n}{\ell}} \frac{1}{n-\ell} \frac{(\ell+1)(\ell-1)}{\ell} \sum_{\substack{S \subseteq V \\ |S|=\ell+1}} C_1^O(\mathbf{x}, S) = \frac{\ell-1}{\ell} \mathrm{E}\left[C_1^O(\mathbf{x}, U_{\ell+1})\right] , \tag{6.26}$$

where we simplify using the fact that $(\ell+1)(\ell-1)\binom{n}{\ell+1} / \left(\binom{n}{\ell}(n-\ell)\right) = \frac{\ell-1}{\ell}$. We use (6.25) and (6.26) to write (6.24) as

$$\mathrm{E}\left[C_1^O(\mathbf{x}, U_\ell)\right] = -\frac{\ell-1}{n(n-1)} \sum_{i,j=1}^{n} A_{vw}(x_v - x_w)^2 + \frac{\ell-1}{\ell} \mathrm{E}\left[C_1^O(\mathbf{x}, U_{\ell+1})\right] . \tag{6.27}$$

We obtain (6.9) by solving the recurrence relation (6.27). □

Proof of Proposition 32. Let

$$F(y, \alpha, s) = \begin{cases} \left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y} \left(s + (1-s)\alpha\right) , & 1-y \geq s \geq s(y,\alpha) \\ 4\left(\sqrt{2}-1\right)^2 \left(\frac{s-y}{1-y}\right)^2 \frac{s}{\alpha} \left(s + (1-s)\alpha\right) \left(\frac{s}{1-s} + \alpha\right) , & s(y,\alpha) > s \geq 2y \,, \end{cases} \tag{6.28}$$

where $y \in [0, 1/3]$ and $s = s(y, \alpha)$ is the solution of

$$\frac{4s(s-y)}{(1-s)(1-y)} \left(\frac{s}{\alpha} + (1-s)\right) = 1 \,. \tag{6.29}$$

Let

$$F(\alpha, s) = \begin{cases} \left(\sqrt{2}-1\right)^2 s\left(s + (1-s)\alpha\right), & 1 > s \geq s(\infty, \alpha) \\ 4\left(\sqrt{2}-1\right)^2 \frac{s^3}{\alpha}\left(s + (1-s)\alpha\right)\left(\frac{s}{1-s} + \alpha\right), & s(\infty, \alpha) \geq s > 0 \,, \end{cases} \tag{6.30}$$

where $s = s(\infty, \alpha)$ is the solution of

$$\frac{4s^2}{1-s} \left(\frac{s}{\alpha} + (1-s)\right) = 1 \,.$$

Denote

$$R := \{(y, \alpha, s) : y \in (0, 1/3],\ \alpha \in (0, 1],\ s \in [2y, 1-y]\} \,.$$

We prove that there exists a constant $c > 0$ such that

$$F(y, s, \alpha) - F(s, \alpha) \leq cy \tag{6.31}$$

for all $(y, \alpha, s) \in R$.

For the rest of this proof, it is useful to introduce some notation. Let $s_\infty^{-1}(s)$ denote the inverse function of $s = s(\infty, \alpha)$, and let $s_3^{-1}(s)$ denote the inverse function of $s = s(3, \alpha)$. Define

$$\bar{s}_\infty^{-1}(s) := \begin{cases} s_\infty^{-1}(s) \,, & s(\infty, 1) > s \geq 0 \\ 1 \,, & s(3, 1) \geq s \geq s(\infty, 1) \,, \end{cases}$$

and define

$$R_1 := \{(y, \alpha, s) : y \in (0, 1/3] \,, \alpha \in (0, 1],\ \max\{2y, s(3, \alpha)\} \leq s \leq 1-y\} \,,$$

$$R_2 := \left\{(y, \alpha, s) : y \in (0, 1/3],\ \alpha \in [s_3^{-1}(s), \bar{s}_\infty^{-1}(s)],\ s \in [\max\{s(\infty, \alpha), 2y\} \,, s(3, \alpha)]\right\} \,,$$

$$R_3 := \{(y, \alpha, s) : y \in (0, 1/3] \,, \alpha \in [s_\infty^{-1}(s), 1],\ s \in [2y, s(\infty, \alpha)]\} \,.$$

The regions R_1, R_2, and R_3 correspond roughly to regions I, II, and III, respectively, from Figure 6.4. We also define

$$\mathcal{R}_1 := \{(y,\alpha,s) : y \in (0,1/3]\,, \alpha \in (0,1],\ \max\{2y, s(y,\alpha)\} \le s \le 1-y\}\,,$$
$$\mathcal{R}_2 := \{(y,\alpha,s) : y \in (0,1/3]\,, \alpha \in [s^{-1}(y,s),1],\ s \in [2y, s(y,\alpha)]\}\,.$$

The top case in (6.30) holds for $(y,\alpha,s) \in \mathcal{R}_1$, and the bottom case in (6.30) holds for $(y,\alpha,s) \in \mathcal{R}_2$. Note that $R_1 \subseteq \mathcal{R}_1$ and $R_3 \subseteq \mathcal{R}_2$ and that $R_1 \cup R_2 \cup R_3 = R$ and $\mathcal{R}_1 \cup \mathcal{R}_2 = R$.

We show that we can obtain a bound of the form (6.31) for a constant $c_1 > 0$ for region R_1, a bound of the form (6.31) for a constant $c_2 > 0$ for region R_2, and a bound of the form (6.31) for a constant $c_3 > 0$ for region R_3. It then follows from $R_1 \cup R_2 \cup R_3 = R$ that taking $c = \max\{c_1, c_2, c_3\}$, yields (6.31).

For $(y,\alpha,s) \in \mathcal{R}_1$, we have that

$$F(y,\alpha,s) = \left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s + (1-s)\alpha)\,. \tag{6.32}$$

We calculate that

$$\frac{\partial}{\partial y}F(y,\alpha,s) = \frac{\left(\sqrt{2}-1\right)^2 (\alpha(1-s)+s)(s-y)}{(1-y)^2} - \frac{\left(\sqrt{2}-1\right)^2 (\alpha(1-s)+s)}{1-y} \tag{6.33}$$

for $(y,\alpha,s) \in \mathcal{R}_1$. This function is continuous in $\mathrm{cl}(R)$ (where $\mathrm{cl}(A)$ denotes the closure of a set A) and thus has a bounded partial derivative with respect to y in R. We let

$$c_1 := \sup_{(y,\alpha,s)\in R} \frac{\partial}{\partial y}F(y,\alpha,s)\,.$$

Because $R_1 \subseteq \mathcal{R}_1$, it follows that

$$F(y,\alpha,s) - F(\alpha,s) \le c_1 y \tag{6.34}$$

for all $(y,\alpha,s) \in R_1$.

For $(y,\alpha,s) \in \mathcal{R}_2$, we have that

$$F(y,\alpha,s) = 4\left(\sqrt{2}-1\right)^2 \left(\frac{s-y}{1-y}\right)^2 \frac{s}{\alpha}\,(s+(1-s)\alpha)\left(\frac{s}{1-s}+\alpha\right). \tag{6.35}$$

We calculate that

$$\frac{\partial}{\partial y}F(y,\alpha,s) = 2(\sqrt{2}-1)^2\frac{1}{(1-y)^2}(s+(1-s)\alpha) \times \frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right) \quad (6.36)$$

for $(y,\alpha,s) \in \mathcal{R}_2$. Note that $s \leq s(y,\alpha)$ implies that

$$\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right) \leq 1\,. \quad (6.37)$$

Using (6.37) and the fact that the function

$$(y,\alpha,s) \mapsto 2(\sqrt{2}-1)^2\frac{(s-1)}{(1-y)^2}(s+(1-s)\alpha)$$

is continuous in $\mathrm{cl}(\mathcal{R}_2)$, we obtain from (6.36) that $F(y,\alpha,s)$ is bounded in $\mathcal{R}_2$. Let

$$c_3 := \sup_{(y,\alpha,s)\in\mathcal{R}_2}\frac{\partial}{\partial y}F(y,\alpha,s)\,.$$

Because $R_3 \subseteq \mathcal{R}_2$, it then follows that

$$F(y,\alpha,s) - F(\alpha,s) \leq c_3 y \quad (6.38)$$

for all $(y,\alpha,s) \in R_3$.

If $(y,\alpha,s) \in R_2 \cap \mathcal{R}_1$, then

$$F(y,\alpha,s) - F(\alpha,s) \leq c_1 y$$

by (6.34).

Suppose that $(y,\alpha,s) \in R_2 \cap \mathcal{R}_2$. It then follows that

$$F(y,\alpha,s) = 4\left(\sqrt{2}-1\right)^2\left(\frac{s-y}{1-y}\right)^2\frac{s}{\alpha}\left(s+(1-s)\alpha\right)\left(\frac{s}{1-s}+\alpha\right)$$

and

$$F(\alpha,s) = \left(\sqrt{2}-1\right)^2 s\left(s+(1-s)\alpha\right).$$

We show that there exists a constant $c_2 > 0$ such that

$$4\left(\sqrt{2}-1\right)^2\left(\frac{s-y}{1-y}\right)^2\frac{s}{\alpha}\left(s+(1-s)\alpha\right)\left(\frac{s}{1-s}+\alpha\right) - \left(\sqrt{2}-1\right)^2 s\left(s+(1-s)\alpha\right) \leq c_2 y\,. \quad (6.39)$$

We subtract and add the expression

$$\left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha) \tag{6.40}$$

from the left-hand side of (6.39) and separately consider the resulting terms

$$\left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha) - \left(\sqrt{2}-1\right)^2 s\,(s+(1-s)\alpha) \tag{6.41}$$

and

$$4\left(\sqrt{2}-1\right)^2 \left(\frac{s-y}{1-y}\right)^2 \frac{s}{\alpha}\,(s+(1-s)\alpha)\left(\frac{s}{1-s}+\alpha\right) - \left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha)\,. \tag{6.42}$$

Because the bound (6.34) holds for all $(y,\alpha,s)\in R$ and $R_s\cap\mathcal{R}_2\subseteq R$, we obtain

$$\left(\sqrt{2}-1\right)^2 s\,(s+(1-s)\alpha) - \left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha) \le c_1 y\,. \tag{6.43}$$

We factor (6.42) as

$$\left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha)\left(1-\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)\right). \tag{6.44}$$

The function

$$(y,\alpha,s)\mapsto \left(\sqrt{2}-1\right)^2 \frac{s-y}{1-y}\,(s+(1-s)\alpha)$$

is continuous in $\mathrm{cl}(R)$ and is thus bounded in R by a positive constant c_4. Therefore, (6.44) is bounded in R_2 by

$$c_4\left(1-\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)\right). \tag{6.45}$$

Because $s=s(y,\alpha)$ solves (6.29), it follows that

$$\begin{aligned}1&-\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)\\&=\frac{4s(y,\alpha)(s(y,\alpha)-y)}{(1-s(y,\alpha))(1-y)}\left(\frac{s(y,\alpha)}{\alpha}+(1-s(y,\alpha))\right)-\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right).\end{aligned}$$

We calculate that

$$\frac{\partial}{\partial y}\left(\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)\right)=\frac{4s(\alpha(s-1)-s)}{\alpha(y-1)^2}\,.$$

This derivative is bounded in $\mathcal{R}_2$. We let

$$c_5=\sup_{(y,\alpha,s)\in\mathcal{R}_2}\frac{4s(\alpha(s-1)-s)}{\alpha(y-1)^2}\,,$$

and we write

$$\frac{4s(y,\alpha)(s(y,\alpha)-y)}{(1-s(y,\alpha))(1-y)}\left(\frac{s(y,\alpha)}{\alpha}+(1-s(y,\alpha))\right)-\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)$$
$$\leq c_5\left(s(y,\alpha)-s\right) \quad (6.46)$$

for all $(y,\alpha,s)\in R_2$.

We write the partial derivative of $s(y,\alpha)$ with respect to y implicitly as

$$\frac{\partial}{\partial y}s(y,\alpha)=-\frac{\frac{\partial}{\partial y}\left(\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)-1\right)}{\frac{\partial}{\partial s}\left(\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{r}{\alpha}+(1-s)\right)-1\right)}\,. \quad (6.47)$$

The numerator of (6.47) is

$$\frac{\partial}{\partial y}\left(\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)-1\right)=\frac{4s\left(\frac{s}{\alpha}-s+1\right)(s-y)}{(1-s)(1-y)^2}+\frac{4s\left(\frac{s}{\alpha}-s+1\right)}{(1-s)(1-y)}\,, \quad (6.48)$$

and the denominator of (6.47) is

$$\frac{\partial}{\partial s}\left(\frac{4s(s-y)}{(1-s)(1-y)}\left(\frac{s}{\alpha}+(1-s)\right)-1\right)$$
$$=\frac{4\left(\frac{1}{\alpha}-1\right)s(s-y)}{(1-s)(1-y)}+\frac{4s\left(\frac{s}{\alpha}-s+1\right)(s-y)}{(1-s)^2(1-y)}+\frac{4s\left(\frac{s}{\alpha}-s+1\right)}{(1-s)(1-y)}+\frac{4\left(\frac{s}{\alpha}-s+1\right)(s-y)}{(1-s)(1-y)}\,. \quad (6.49)$$

The left-hand side of (6.48) is bounded in $\mathcal{R}_2$, and the left-hand side of (6.49) is non-zero in $\mathcal{R}_2$. By the formula (6.47), $\frac{\partial}{\partial y}s(y,\alpha)$ is bounded in R_2. Let

$$c_6=\sup_{(y,\alpha,s)\in R_2}\frac{\partial}{\partial y}s(y,\alpha)\,,$$

and write

$$s(y,\alpha) - s \le s(y,\alpha) - s(0,\alpha) \le c_6 y\,. \tag{6.50}$$

We now let $c_3 = c_1 + c_4c_5c_6$ and apply (6.43), (6.45), (6.46), and (6.50) to obtain

$$F(y,\alpha,s) - F(\alpha,s) \le c_2 y \tag{6.51}$$

for all $(y,\alpha,s) \in R_2$. Let $c = \max\{c_1, c_2, c_3\}$. It follows from (6.34), (6.38), and (6.51) that

$$F(y,\alpha,s) - F(\alpha,s) \le cy \tag{6.52}$$

for all $(y,\alpha,s) \in R$. We have a value of c such that (6.52) holds, so we have completed the proof. □

Proof of Proposition 34. Applying Lemma 25 to (6.20), we obtain the inequality

$$\mathrm{E}\left[1 - \imath_{\mathrm{step}}(\mathbf{x}, U_\ell)\right] \ge \frac{\ell}{n-\ell} \frac{\mathrm{E}^-\left[C_1^O(\mathbf{x},S)\right]^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\,. \tag{6.53}$$

Because $\mathrm{E}\left[1 - \imath_{\mathrm{step}}(\mathbf{x}, U_\ell)\right]$ satisfies both the bound (6.19) and the bound (6.53), it also satisfies the inequality

$$\begin{aligned} &\mathrm{E}\left[1 - \imath_{\mathrm{step}}(\mathbf{x}, U)\right] \ge \frac{1}{e(\mathbf{x})} \\ &\times \min\left(\frac{\ell}{n-\ell}\frac{\mathrm{E}^-\left[C_1^O(\mathbf{x},S)\right]^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}}, \frac{\ell(\ell-1)}{n-1}e(\mathbf{x}) + \frac{(n-\ell)(\ell-1)}{n(n-1)}\mathbf{x}^T\mathbf{L}\mathbf{x} - \mathrm{E}^-\left[-C_1^O(\mathbf{x},S)\right]\right)\,. \end{aligned} \tag{6.54}$$

We proceed by cases to prove a lower bound using (6.54).

1. $\mathbf{x}^T\mathbf{L}\mathbf{x} \ge \gamma_r^2 \frac{\ell n}{n-\ell} e(\mathbf{x})$. By Lemma 31, we obtain

$$\mathrm{E}^-\left[C_1^O(\mathbf{x},S)\right] \le -\frac{(n-\ell)(\ell-1)}{n(n-1)}\mathbf{x}^T\mathbf{L}\mathbf{x}\,.$$

   Squaring both sides of this inequality and rearranging terms, we find that

$$\frac{\ell}{n-\ell}\frac{\mathrm{E}^-\left[C_1^O(\mathbf{x},S)\right]^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}} \ge \frac{\ell(l-1)^2(n-\ell)}{n^2(n-1)^2}\mathbf{x}^T\mathbf{L}\mathbf{x}\,.$$

We apply $\mathbf{x}^T\mathbf{L}\mathbf{x} \geq \gamma_r^2 \frac{\ell n}{n-\ell} e(\mathbf{x})$ to this last inequality and conclude that

$$\frac{\ell}{n-\ell} \frac{\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right]^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}} \geq \gamma_r^2 \frac{\ell^2(\ell-1)^2}{n(n-1)^2} e(\mathbf{x})\,. \tag{6.55}$$

2. $\mathbf{x}^T\mathbf{L}\mathbf{x} < \gamma_r^2 \frac{\ell n}{n-\ell} e(\mathbf{x})$ and $\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right] \geq 2\gamma_r \frac{\ell-1}{n-1}\sqrt{\frac{\ell(n-\ell)}{n}} e(\mathbf{x})^{1/2}(\mathbf{x}^T\mathbf{L}\mathbf{x})^{1/2}$. In this case, we use the lower bound for $\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right]$ to obtain

$$\frac{\ell}{n-\ell} \frac{\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right]^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}} \geq 4\gamma_r^2 \frac{\ell^2(\ell-1)^2}{n(n-1)^2} e(\mathbf{x})\,. \tag{6.56}$$

3. $\mathbf{x}^T\mathbf{L}\mathbf{x} < \gamma_r^2 \frac{\ell n}{n-\ell} e(\mathbf{x})$ and $\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right] < 2\gamma_r \frac{\ell-1}{n-1}\sqrt{\frac{\ell(n-\ell)}{n}} e(\mathbf{x})^{1/2}(\mathbf{x}^T\mathbf{L}\mathbf{x})^{1/2}$. We begin by writing the expression

$$\frac{\ell(l-1)}{n-1} e(\mathbf{x}) + \frac{(n-\ell)(\ell-1)}{n(n-1)} \mathbf{x}^T\mathbf{L}\mathbf{x} - \mathrm{E}^- \left[-C_1^O(\mathbf{x}, S)\right] \tag{6.57}$$

from (6.54) as

$$\left(\sqrt{\frac{\ell(\ell-1)}{n-1}}\; e(\mathbf{x})^{1/2} - \sqrt{\frac{(n-\ell)(\ell-1)}{n(n-1)}} (\mathbf{x}^T\mathbf{L}\mathbf{x})^{1/2}\right)^2 + \left(2\frac{\ell-1}{n-1}\sqrt{\frac{\ell(n-\ell)}{n}} e(\mathbf{x})^{1/2}(\mathbf{x}^T\mathbf{L}\mathbf{x})^{1/2} - \mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right]\right). \tag{6.58}$$

We use the upper bound for $\mathrm{E}^- \left[C_1^O(\mathbf{x}, S)\right]$ and the fact that $\gamma_r < 1$ to lower-bound (6.58) by

$$\left(\sqrt{\frac{\ell(\ell-1)}{n-1}} e(\mathbf{x})^{1/2} - \sqrt{\frac{(n-\ell)(\ell-1)}{n(n-1)}} (\mathbf{x}^T\mathbf{L}\mathbf{x})^{1/2}\right)^2 . \tag{6.59}$$

Using our assumption that $\mathbf{x}^T\mathbf{L}\mathbf{x} \leq \gamma_r^2 \frac{\ell n}{n-\ell} e(\mathbf{x})$, we lower-bound (6.59) by

$$\frac{\ell(\ell-1)}{n-1} (1-\gamma_r)^2 e(\mathbf{x})\,.$$

Because γ_r satisfies the equation $(1-\gamma_r)^2 = \frac{\ell(\ell-1)}{n(n-1)}\gamma_r^2$, we lower-bound (6.57) as follows:

$$\frac{\ell(l-1)}{n-1} e(\mathbf{x}) + \frac{(n-\ell)(\ell-1)}{n(n-1)} \mathbf{x}^T\mathbf{L}\mathbf{x} - \mathrm{E}^- \left[-C_1^O(\mathbf{x}, S)\right] \geq \gamma_r^2 \frac{\ell^2(\ell-1)^2}{n(n-1)^2} e(\mathbf{x})\,. \tag{6.60}$$

From the inequalities (6.55), (6.56), and (6.60), we see that the maximum in (6.54) is bounded below by

$$\frac{\ell^2(\ell-1)^2}{n(n-1)^2}\gamma_r^2 e(\mathbf{x}) \,.$$

This completes the proof. □

CHAPTER 7

Convergence to Consensus with Strategies for Choosing Meeting Groups

In this chapter, we study the ϵ-convergence time of the group-soft-consensus model (3.5) for various strategies for choosing meeting groups. Our motivation is twofold: we seek (1) to evaluate the effectiveness of our strategies for bringing a population to consensus quickly and (2) to better understand the behavior of our model by identifying the meeting groups that result in significant progress towards consensus. We call our general strategy the *group-interventionist strategy*. This strategy is a greedy algorithm that works by removing nodes from the set V of nodes until we obtain a set of a desired size. We present the details in Algorithm 1.

Algorithm 1 has as input a function $C(\mathbf{x}, S)$ and a set size $\ell \in \{2, 3, \dots, n-1\}$. Starting with the set V of nodes, the algorithm picks a set $S_{n-1} \subseteq V$ of size $n-1$ such that S_{n-1} maximizes $C(\mathbf{x}, S)$ among all sets S of size $n-1$. It then does the same procedure with S_{n-1} to obtain a set S_{n-2} of size $n-2$ that is contained in S_{n-1} and maximizes $C(\mathbf{x}, S)$ among all sets S of size $n-1$ that are contained in S_{n-1}. Continuing in this fashion, Algorithm 1 removes nodes until we obtain a set of size $S_{\ell+1}$, at which point it chooses a set S_ℓ of size ℓ in $S_{\ell+1}$ that maximizes $\beta(\mathbf{x}, S)C_1^I(\mathbf{x}, S)$ among sets S of size ℓ that are contained in $S_{\ell+1}$. We use the notation $\mathcal{A}(C, \ell)$ to denote Algorithm 1 with inputs C and ℓ. Note that $\mathcal{A}(C, \ell)$ is a function of $\mathbf{x} \in \mathbb{R}^n$. We denote the set S_ℓ that $\mathcal{A}(C, \ell)$ returns with input vector $\mathbf{x} \in \mathbb{R}^n$ by $S_\ell = \mathcal{A}_{\mathbf{x}}(C, \ell)$.

The lower bound in (5.25) depends on the size of the in-group and out-group costs for a

Algorithm 1: InterventionistGroup

Input: Function $C(\mathbf{x}, S)$, meeting group size $\ell \in \{2, \ldots, n-1\}$, and vector $\mathbf{x} \in \mathbb{R}^n$

Output: Meeting group set S_ℓ of size ℓ

```
S_n ⟵ V
for k = n to ℓ + 2 do
    v_{k−1} ⟵ arg min_{w∈S_k} C(x, S_k \ {w})
    S_{k−1} ⟵ S_k \ {v_{k−1}}
end
v_ℓ ⟵ arg min_{v∈S_{ℓ+1}} β(x, S_{ℓ+1} \ {v}) C_1^I(x, S_{ℓ+1} \ {v})
S_ℓ ⟵ S_{ℓ+1} \ {v_ℓ}
return S_ℓ
```

choice of meeting group S. This observation leads us to consider three different functions C in Algorithm 1: the in-group cost $C(\mathbf{x}, S) = C_1^I(\mathbf{x}, S)$, the out-group cost $C(\mathbf{x}, S) = C_1^O(\mathbf{x}, S)$, and the combined cost $C_1(\mathbf{x}, S)$. We state bounds on the ϵ-convergence times for these functions in Theorems 42, 44, and 48 for $C = C_1^I$, $C = C_1$, and $C = C_1^O$, respectively. To begin, we state the following lemma, which provides bounds on the mean over the set $S_{\ell+1}$ that we obtain from Algorithm 1 for each of the choices of C.

Lemma 37. *The following statements hold:*

(1) *In Algorithm 1 with $C = C_1$, we have that*

$$\frac{1}{\ell+1} \sum_{v \in S_{\ell+1}} C_1(\mathbf{x}, S_{\ell+1} \setminus \{v\}) \leq -\frac{(\ell-1)\ell}{n-1} e(\mathbf{x}) - \frac{\ell-1}{(n-1)n} \mathbf{x}^T \mathbf{L} \mathbf{x}\,.$$

(2) *In Algorithm 1 with $C = C_1^I$, we have that*

$$\frac{1}{\ell+1} \sum_{v \in S_{\ell+1}} C_1^I(\mathbf{x}, S_{\ell+1} \setminus \{v\}) \leq -\frac{(\ell-1)\ell}{n-1} e(\mathbf{x})\,.$$

(3) *In Algorithm 1 with $C = C_1^I$, we have that*

$$\frac{1}{\ell+1} \sum_{v \in S_{\ell+1}} C_1^O(\mathbf{x}, S_{\ell+1} \setminus \{v\}) \leq -\frac{\ell-1}{(n-1)n} \mathbf{x}^T \mathbf{L} \mathbf{x} .$$

In the following proposition, we describe the sets S_k, for $k \geq \ell+1$, that are selected by Algorithm 1 with $C = C_1^I$. We show that the node that is removed from the set S_k is the node whose belief is closest to the mean belief of the nodes in S_k. Additionally, we prove that if a node is in all of the sets $S_k, \ldots, S_n$, then its belief must either be always less than, always greater than, or always equal to the mean beliefs in the sets S_k for $k \in \{\ell+1, \ldots, n\}$.

Proposition 38. *In Algorithm 1 with $C = C_1^I$, the nodes v_k, with $k \in \{\ell+1, \ldots, n-1\}$, satisfy*

$$v_k = \underset{v \in S_{k+1}}{\arg\min} \left(x_v - \bar{x}^{S_{k+1}}\right)^2 . \tag{7.1}$$

Furthermore, for $v \in S_k \cap S_{k+1}$, it follows that $x_v < \bar{x}^{S_k}$ if and only if $x_v < \bar{x}^{S_{k+1}}$ and $x_v = \bar{x}^{S_k}$ if and only if $x_v = \bar{x}^{S_{k+1}}$.

With this characterization, we compute the per-step convergence factor $r_{\text{step}}(\mathcal{A}(C_1^I, \ell))$ (see Definition 12 in Chapter 5) for the graph of two equally-sized cliques (see Example 9 for the definition of this graph).

Example 39. *We consider the same two-clique graph G and initial vector $\mathbf{x}_0$ as in Example 9. Denote the complete graph in which each node has an initial belief -1 by G_{-1} and the clique in which each node has initial belief 1 by G_1. When ℓ is even, we use Proposition 38 to determine that the set $S_\ell = \mathcal{A}_{\mathbf{x}_0}(C_1^I, \ell)$ consists of exactly $\ell/2$ nodes from G_{-1} and $\ell/2$ nodes from G_1. Using this, we calculate that $C_1^I(\mathbf{x}_0, S_\ell) = \ell^2$, $C_1^O(\mathbf{x}_0, S_\ell) = 0$, $C_2(\mathbf{x}_0, S_\ell) = \frac{\ell}{2}(n+\ell)$, and $e(\mathbf{x}_0) = n$. It then follows that the optimal degree of consensus is*

$$\beta^*(t) = \min\left(1, \frac{2\ell}{n+\ell}\right) = \frac{2\ell}{n+\ell}$$

and thus that the per-step convergence factor of $\mathbf{x}_0$ and $\mathcal{A}(C_1^I, \ell)$ is

$$1 - r_{\text{step}}(\mathbf{x}_0, \mathcal{A}(C_1^I, \ell)) = \left(2 - \frac{2\ell}{n+\ell}\right) \frac{2\ell}{n+\ell} \frac{\ell}{n} = 4\frac{\ell^2}{(n+\ell)^2} = 4\frac{s^2}{(1+s)^2} . \tag{7.2}$$

Using (7.2)*, we obtain*

$$1 - r_{\text{step}}(\mathcal{A}(C_1^I, \ell)) \leq 4\frac{s^2}{(1+s)^2} \,. \tag{7.3}$$

With this bound for the per-step convergence factor of $\mathcal{A}(C_1^I, \ell)$*, the best upper bound for* $T(\epsilon, \mathcal{A}(C_1^I, \ell))$ *that depends only on* s *and* ϵ *that we can obtain using Proposition 15 is*

$$\frac{\log(\epsilon)}{\log\left(1 - 4\frac{s^2}{(1+s)^2}\right)} = 4\log(\epsilon^{-1})\frac{(1+s)^2}{s^2} + O(1)\,, \tag{7.4}$$

where the asymptotic expression holds as $s \to 0$ *and is a result of* (5.14) *in Remark 17.*

In the following proposition, we further characterize the behavior of Algorithm 1 with C_1^I.

Proposition 40. *Consider Algorithm 1 with inputs* C_1^I *and* ℓ*. Let* $w \in V$*. For all* $k \in \{\ell+1, \ldots, n-1\}$ *and* $v \in S_k$*, the following holds:*

(1) *If* $x_v < \bar{x}^{S_k}$ *and* w *satisfies* $x_w < x_v$*, then* $w \in S_k$*.*

(2) *If* $x_v > \bar{x}^{S_k}$ *and* w *satisfies* $x_w > x_v$*, then* $w \in S_k$*.*

A consequence of Proposition 40 is that all of the sets S_k have a non-negative out-group cost. Let $k \in \{\ell+1, \ldots, n-1\}$, let $v \in S_k$, and let $w \in V \setminus S_k$. By Proposition 40, if $x_v > \bar{x}^{S_k}$, it follows that $x_w \leq x_v$. Consequently, if $x_v \geq \bar{x}^{S_k}$, we have that

$$\left(x_v - \bar{x}^{S_k}\right)(x_v - x_w) \geq 0\,. \tag{7.5}$$

Similarly, if $x_v < \bar{x}^{S_k}$, it follows that $x_w \geq x_v$. Therefore, if $x_v < \bar{x}^{S_k}$, it is again the case that (7.5) holds. Therefore, the out-group cost is non-positive:

$$C_1^O(\mathbf{x}, S_k) = -\sum_{v \in S_k}\sum_{w \notin S_k} A_{vw}\left(x_v - \bar{x}^{S_k}\right)(x_v - x_w) \leq 0\,. \tag{7.6}$$

We use this inequality to prove the bound for the per-step convergence factor in the following proposition.

Proposition 41. *For all $n \geq 3$, the following lower bound for the convergence factor holds:*

$$1 - r_{\text{step}}(\mathcal{A}(C_1^I, \ell)) \geq \frac{1}{2}s^2 \,. \tag{7.7}$$

Proof. Let $S_{\ell+1}, S_{\ell+2}, \ldots, S_{n-1}$ be generated by Algorithm 1 with inputs C_1^I and ℓ, let

$$v_\ell := \operatorname*{arg\,min}_{w \in S_{\ell+1}} C_1^I(x, S_{\ell+1} \setminus \{w\}) \,,$$

and let $S'_\ell = S_{\ell+1} \setminus \{v_\ell\}$. From statement (2) in Lemma 37, it follows that the set S'_ℓ satisfies

$$C_1^I(\mathbf{x}, S'_\ell) \leq -\frac{(\ell-1)\ell}{n-1} e(\mathbf{x}) \,.$$

Therefore, by (7.6), the out-group cost $C_1^I(\mathbf{x}, S'_\ell)$ is non-positive. Therefore,

$$\mathcal{B}(\mathbf{x}, S'_\ell) \leq C_1^I(\mathbf{x}, S'_\ell) \leq -\frac{(\ell-1)\ell}{n-1} e(\mathbf{x}) \,.$$

We use this bound for $\mathcal{B}(\mathbf{x}, S'_\ell)$ in the bound (5.25) to obtain

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S'_\ell) \geq \frac{(\ell-1)\ell}{(n-1)n} \,.$$

Because $1 - r_{\text{step}}(S_\ell) \geq 1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S'_\ell)$, we conclude that

$$1 - r_{\text{step}}(S_\ell) \geq \frac{(\ell-1)\ell}{(n-1)n} \,. \tag{7.8}$$

Changing variables to $s = \ell/n$, we write the lower bound in (7.8) as

$$\frac{(\ell-1)\ell}{(n-1)n} = s\left(s - \frac{1}{n}\right)\left(1 - \frac{1}{n}\right)^{-1} \,. \tag{7.9}$$

Because $(1 - 1/n)^{-1} \geq 1 + 1/n$, it then follows that

$$s\left(s - \frac{1}{n}\right)\left(1 - \frac{1}{n}\right)^{-1} = s\left(s - \frac{1}{n}\right)\left(1 + \frac{1}{n}\right) = s^2 - \frac{s}{n} + \frac{s^2}{n} - \frac{s}{n^2} \,. \tag{7.10}$$

From $s \geq 2/n$, it follows that $s^2/2 \geq s/n$ and $s^2/n \geq 2s/n^2$. Using these two inequalities, we obtain

$$s^2 - \frac{s}{n} + \frac{s^2}{n} - \frac{s}{n^2} \geq \frac{1}{2}s^2 + \frac{s}{n^2} \geq \frac{1}{2}s^2 \,. \tag{7.11}$$

Combining the bounds (7.9), (7.10), and (7.11) yields

$$\frac{(\ell-1)\ell}{(n-1)n} \geq \frac{1}{2}s^2\,. \tag{7.12}$$

We insert (7.12) into (7.8) to obtain (7.7). □

Combining the bound (7.3) from Example 39 with the bound (7.7) from Proposition 41, we determine that

$$\frac{1}{2}s^2 \leq 1 - r_{\text{step}}(\mathcal{A}(C_1^I, \ell)) \leq 4\frac{s^2}{(1+s)^2}\,.$$

Because $1/(1+s)^2 \leq 1$, it follows that

$$s^2 \leq 1 - r_{\text{step}}(\mathcal{A}(C_1^I, \ell)) \leq 4s^2\,. \tag{7.13}$$

We state our main result for $\mathcal{A}(C_1^I, \ell)$.

Theorem 42. *For all $n \geq 3$ and all $\epsilon \in (0, 1)$, the ϵ-convergence time satisfies the bound*

$$T(\epsilon, (\mathcal{A}(C_1^I, \ell))) \leq \frac{\log(\epsilon)}{\log(1 - s^2)}\,. \tag{7.14}$$

Proof. This bound results from using Proposition 41 in Proposition 15. □

As a result of (7.13), the best upper bound for $T(\epsilon, \mathcal{A}(C_1^I, \ell))$ that depends only on s and ϵ that one can possibly obtain from Proposition 15 is

$$\frac{\log(\epsilon)}{\log(1 - 4s^2)}\,. \tag{7.15}$$

Comparing (7.15) with (7.14) from Theorem 42, we conclude that the quadratic dependence on s in (7.14) is optimal among the bounds that are dependent only on s and ϵ that can obtained from Proposition 15.

Using (5.14) from Remark 17 on the bound (7.14) of Theorem 42, we obtain

$$T(\epsilon, S_\ell) \leq \log(\epsilon^{-1})s^{-2} + O(1)$$

as $s \to 0$. Because $s = \ell/n$, the bound (7.14) of Theorem 42 is approximately quadratic in n for fixed ℓ.

We now upper-bound the ϵ-convergence time $T(\epsilon, \mathcal{A}(C_1, \ell))$. Our analysis closely follows our strategy to prove Proposition 32 in Chapter 6, so we omit some details. Using the notation of Proposition 23, we let $\mathcal{S}$ be the set of sets S of size ℓ that are contained in the set $S_{\ell+1}$ that we generate from Algorithm 1 with inputs C_1 and ℓ. Additionally, we set $w(S)$ to be the uniform weight $1/(\ell+1)$. From Lemma 37, we have that

$$\frac{1}{\ell+1}\sum_{v \in S_{\ell+1}} C_1(\mathbf{x}, S_{\ell+1} \setminus \{v\}) = -\frac{\ell-1}{n-1}\left(\ell e(\mathbf{x}) + \frac{1}{n}\mathbf{x}^T\mathbf{L}\mathbf{x}\right) \leq -\frac{\ell-1}{n-1}\ell e(\mathbf{x}) .$$

Applying Proposition 23 with

$$\delta = \frac{\ell-1}{n-1}\ell e(\mathbf{x}) ,$$

we have that

$$1 - r_{\text{step}}(\mathbf{x}, S_\ell) \geq \begin{cases} -\left(\sqrt{2}-1\right)^2 \frac{s(s-1/n)}{1-1/n} , & s \geq s(n,\alpha) \\ -4\left(\sqrt{2}-1\right)^2 \left(\frac{s-1/n}{1-1/n}\right)^2 \frac{s^3}{(1-s)}\frac{1}{\alpha} , & s \leq s(n,\alpha) , \end{cases} \tag{7.16}$$

where $s = s(n, \alpha)$ solves the equation

$$\frac{4s^2(s-1/n)}{(1-s)(1-1/n)}\frac{1}{\alpha} = 1 \tag{7.17}$$

for $\alpha \in (0, 1]$ and $s(n, 0) = 1$.

Considering the limit of the bound in (7.16) as $n \to \infty$ leads to the bound in the following proposition. The proof is a simpler version of the proof of Proposition 32 in Chapter 6.

Proposition 43. *The per-step convergence factor of $\mathcal{A}(C_1, \ell)$ satisfies the following bound:*

$$1 - r_{\text{step}}(\mathcal{A}(C_1, \ell)) + \frac{c}{n} \geq \begin{cases} -\left(\sqrt{2}-1\right)^2 s^2 , & s \geq s(\infty,\alpha) \\ -4\left(\sqrt{2}-1\right)^2 \frac{s^5}{1-s}\frac{1}{\alpha} , & s \leq s(\infty,\alpha) , \end{cases} \tag{7.18}$$

where $c > 0$ is a constant and $s(\infty, \alpha)$ solves the equation

$$\frac{4s^3}{1-s} = \alpha$$

for $\alpha \in (0, 1]$ and $s(\infty, \alpha) = 0$ for $\alpha = 0$.

The following bound for $T(\epsilon, \mathcal{A}(C_1, \ell))$ is a result of Proposition 43.

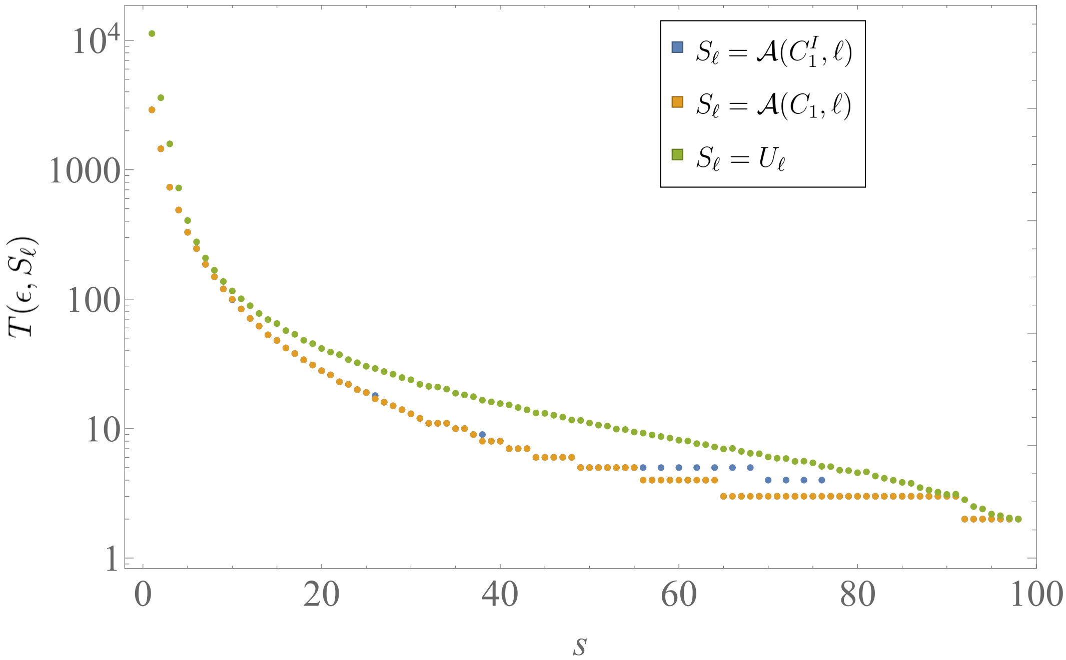
Theorem 44. *There exists a constant $c > 0$ such that for all $n \geq 3$ and all $\epsilon \in (0, 1)$, the ϵ-convergence time satisfies*

$$T(\epsilon, \mathcal{A}(C_1, \ell)) \leq \begin{cases} \frac{\log(\epsilon)}{\log\left(1+\frac{c}{n}-(\sqrt{2}-1)^2 s^2\right)}\,, & s \geq 1/2 \\ \frac{\log(\epsilon)}{\log\left(1+\frac{c}{n}-4(\sqrt{2}-1)^2 \frac{s^5}{1-s}\right)}\,, & s < 1/2\,. \end{cases}$$

Proof. The proof results from applying Propositions 15 and 43 and using the fact that $s(\infty, 1) = 1/2$. □

In Figure 7.1, we plot simulated ϵ-convergence times with meeting groups that we select from Algorithm 1 with $C = C_1^I$, from Algorithm 1 with $C = C_1$, and uniformly at random. We run simulations on the graph of two equally-sized cliques (see Example 9) and the $(n-1)$-to-1 graph (see Example 30). We set the initial vector in each simulation to -1 for each entry that corresponds to a node in one connected component and to 1 for each entry that corresponds to a node in the other connected component. The simulated ϵ-convergence times are the number t of meetings that are required for $e(x(t))/e(x(0)) < \epsilon$. For the $(n-1)$-to-1 graph, the simulations reveal a significant convergence-time improvement when meeting groups are generated by Algorithm 1 with $C = C_1^I$ and when meeting groups are generated by Algorithm 1 with $C = C_1$ compared to when meeting groups are generated uniformly at random (see the bottom panel of Figure 7.1). By contrast, our simulations using the graph of two equally-sized cliques indicate that selecting meeting groups using Algorithm 1 with either $C = C_1^I$ or $C = C_1$ results in a relatively small improvement over selecting meeting groups uniformly at random (see the top panel of Figure 7.1).

We now study the ϵ-convergence time when selecting meeting groups from Algorithm 1 with $C = C_1^O$. As we show in the next example, if we use the function C_1^O in Algorithm 1 on a graph that is not connected, C_1^O may not be able to provide any guidance to Algorithm 1 about which meeting group to select.



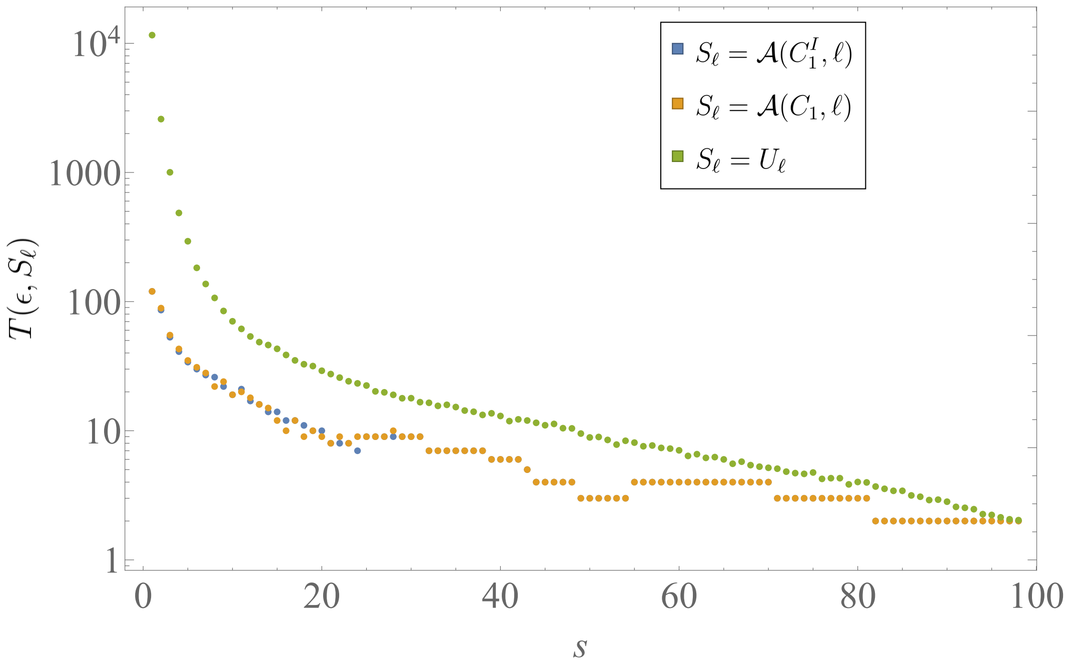


Figure 7.1: Simulated ϵ-convergence times with $n = 100$ and $\epsilon = .0001$ for (top) the graph of two equally-sized cliques and for (bottom) the $(n-1)$-to-1 graph.

Example 45. *Let G be a graph that consists of two complete graphs, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, that are disconnected from each other. Suppose that $x_v = x_w$ for all $v, w \in V_1$ and that $x_v = x_w$ for all $v, w \in V_2$. It then follows that $C_1^O(\mathbf{x}, S) = 0$ for all $S \subseteq V$. Therefore, the function C_1^O does not provide Algorithm 1 any guidance for meeting-group selection.*

Another limitation of using $C_1^O(\mathbf{x}, S)$ in Algorithm 1 is that even for a connected graph, the per-step convergence factor of $\mathcal{A}(C_1^O, \ell)$ can be $O(1/n)$ for all $s \in [2/n, (n-1)/n]$. We demonstrate this in the next example.

Example 46. *Let G be a graph that consists of m nodes $v_1, \dots, v_m$ with initial belief 0, m nodes $v_{m+1}, \dots, v_{2m}$ with initial belief $z > 0$, and 1 node v_{2m+1} with initial belief $y > z$. We denote the initial belief vector by $\mathbf{x}_0$. However, for notational convenience, we omit the subscript 0 when we indicate the components of $\mathbf{x}_0$ and the mean belief of nodes in a set S. The edge set of G consists of the edges $\{v_j, v_{m+j}\}$ and $\{v_{m+j}, v_{j+1}\}$ for $j \in \{1, \dots, m\}$, the edge $\{v_1, v_{2m}\}$, and the edges $\{v_j, v_{2m+1}\}$ for all $j \in \{1, \dots, m\}$.*

For all $\ell \in \{2, \dots, 2m-1\}$, one can make $1 - r_{\text{step}}(\mathbf{x}_0, S_\ell)$ arbitrarily small for an appropriate choice of z by choosing m to be large. To prove this fact, we first let $B := V \setminus \{v_{2m+1}\}$ and $D := V \setminus \{v_1\}$. We calculate the means

$$\bar{x}^B = \frac{z}{2}\,,$$
$$\bar{x}^D = \frac{y + zm}{2m}$$

and the out-group costs

$$-C_1^O(\mathbf{x}_0, B) = \bar{x}^B y m\,,$$
$$-C_1^O(\mathbf{x}_0, D) = 2\left(z - \bar{x}^D\right) z + \left(y - \bar{x}^D\right) y\,.$$

We let $z = 4y/m$ *and calculate*

$$-C_1^O(\mathbf{x}_0, B) = 2y^2\,,$$
$$-C_1^O(\mathbf{x}_0, D) = \frac{24 - 5m + 2m^2}{2m^2}y^2.$$

Because $-C_1^O(\mathbf{x}_0, D) \to y^2$ *as* $m \to \infty$*, it follows that*

$$-C_1^O(\mathbf{x}_0, B) > -C_1^O(\mathbf{x}_0, D)$$

for sufficiently large m*.*

Additionally, $-C_1^O(\mathbf{x}_0, V\backslash\{v_j\})$ *for* $j \in \{m+1, \ldots, 2m\}$ *is smaller than both* $-C_1^O(\mathbf{x}_0, B) = 2y^2$ *and* $-C_1^O(\mathbf{x}_0, D) = \frac{24-5m+2m^2}{2m^2}y^2$*. Therefore, for sufficiently large* m*, we see that* $S_{2m} = V \setminus \{v_j\}$ *for some* $j \in \{1, \ldots, m\}$*.*

We calculate that

$$\sum_{v\in B} \left(x_v - \bar{x}^B\right)^2 = m\left(\frac{2y}{m}\right)^2 + m\left(\frac{2y}{m}\right)^2 = \frac{8y^2}{m}\,. \tag{7.19}$$

For any set S' *of nodes and all* $v \in S'$*, we have the following inequality:*

$$\sum_{w\in S'\backslash\{v\}} \left(x_w - \bar{x}^{S'\backslash\{w\}}\right)^2 = \sum_{w\in S'} \left(x_w - \bar{x}^{S'}\right)^2 - \frac{|S'|}{|S'|-1}\left(x_v - \bar{x}^{S'}\right)^2 \leq \sum_{w\in S'} \left(x_w - \bar{x}^{S'}\right)^2.$$

It follows by induction using this inequality and (7.19) *that*

$$\sum_{v\in S} \left(x_v - \bar{x}^S\right)^2 \leq \frac{8y^2}{m} \tag{7.20}$$

for any set $S \subseteq B$*. Using the fact that* $m \geq 1$ *and* $z < y$*, we obtain*

$$e(\mathbf{x}_0) \geq (x_{2m+1} - \bar{x})^2 = \left(y - \frac{y + mz}{2m+1}\right)^2 = \left(\frac{2}{3}y - \frac{1}{3}z\right)^2 \geq \frac{1}{9}y^2\,. \tag{7.21}$$

We insert (7.20) *and* (7.21) *into* (7.19) *to obtain*

$$\frac{\sum_{v\in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x}_0)} \leq \frac{8y^2}{m}\frac{9}{y^2} = 72\frac{1}{m} \tag{7.22}$$

for all $S \subseteq S_{2m} = B$*. Using*

$$1 - r_{\text{step}}(\mathbf{x}_0, S) = (2 - \beta(\mathbf{x}, S))\beta(\mathbf{x}, S)\frac{\sum_{v\in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})} \leq \frac{\sum_{v\in S}(x_v - \bar{x}^S)^2}{e(\mathbf{x})}$$

and (7.22)*, we obtain the inequality*

$$1 - r_{\text{step}}(\mathbf{x}_0, S) \leq 72\frac{1}{m}\,, \tag{7.23}$$

which holds for sufficiently large m*. Using the fact that* $n = 2m + 1$ *and using the formula* (5.7) *from Definition 12, it then follows from* (7.23) *that*

$$1 - r_{\text{step}}(\mathcal{B}(C_1^o, \ell)) \leq 144\frac{1}{n-1} \tag{7.24}$$

for sufficiently large m *and for all* $\ell \in \{2, \ldots, 2m - 1\}$*. Therefore, for sufficiently large* n*, any upper bound that depends only on* n *and* ϵ *that we obtain from Proposition 15 is at least*

$$\frac{2\log(\epsilon)}{\log\left(1 - 144\frac{1}{n-1}\right)} = \frac{1}{72}\log(\epsilon)(n-1) + O(1),$$

where the asymptotic expression holds as $n \to \infty$ *and follows from* (5.14) *in Remark 17.*

From Examples 45 and 46, we conclude that using $C = C_1^O$ in Algorithm 1 with $C = C_1^O$ is ineffective for disconnected graphs and that $1 - r_{\text{step}}(\mathcal{A}(C_1^O, \ell))$ is no larger than $O(1/n)$ for all $\ell < n - 1$ for some connected graphs (such as the graph in Example 46). The question, then, is if Algorithm 1 with $C = C_1^O$ results in convergence to consensus for connected graphs. This question is answered in the affirmative in the next proposition.

Proposition 47. *There exists a constant* $c > 0$ *such that the following lower bound holds:*

$$1 - r_{\text{step}}(\mathcal{A}(C_1^O, \ell)) \geq \frac{1}{n^2}\frac{s^3}{(2-s)^3}\alpha\,. \tag{7.25}$$

Proof. Let $S_{\ell+1}, S_{\ell+2}, \ldots, S_{n-1}$ be generated by Algorithm 1 with inputs C_1^O and ℓ, let

$$v_\ell := \operatorname*{arg\,min}_{w \in S_{\ell+1}} C_1^O(x, S_{\ell+1} \setminus \{w\})\,,$$

and let $S'_\ell = S_{\ell+1} \setminus \{v_\ell\}$. From Lemma 25, it follows that

$$C_1^I(\mathbf{x}, S'_\ell) \leq -\frac{\ell}{n-\ell}\frac{C_1^O(\mathbf{x}, S'_\ell)^2}{\mathbf{x}^T\mathbf{L}\mathbf{x}}\,.$$

Because we chose S'_ℓ to minimize the out-group cost, we apply (3) from Lemma 37 to obtain

$$C_1^I(\mathbf{x}, S'_\ell) \leq -\frac{\ell}{n-\ell}\frac{(\ell-1)^2}{n^2(n-1)^2}\mathbf{x}^T\mathbf{L}\mathbf{x}\,.$$

Because $C_1^O(\mathbf{x}, S'_\ell) \leq 0$, it follows that

$$\mathcal{B}(\mathbf{x}, S'_\ell) \leq -\frac{\ell}{n-\ell}\frac{(\ell-1)^2}{n^2(n-1)^2}\mathbf{x}^T\mathbf{L}\mathbf{x}\,.$$

We use this this bound in (5.25) to obtain

$$1 - r_{\text{step}}(\mathbf{x}, S'_\ell) \geq \frac{1}{n^2}\frac{s}{1-s}\left(\frac{s-\frac{1}{n}}{1-\frac{1}{n}}\right)^2 \alpha\,. \tag{7.26}$$

Note that $s \geq 2/n$ implies that $n \geq 2/n$. Using this fact and the fact that

$$\left(\frac{s-\frac{1}{n}}{1-\frac{1}{n}}\right)^2 \tag{7.27}$$

is monotonically non-decreasing in n, we determine that

$$\left(\frac{s-\frac{1}{n}}{1-\frac{1}{n}}\right)^2 \geq \frac{s^2}{(s-2)^2}\,, \tag{7.28}$$

where we obtain the right-hand side of (7.28) by making the substitution $n \to s/2$ in (7.27). Using (7.28) in (7.26), it follows that

$$1 - r_{\text{step}}(\mathbf{x}, S'_\ell) \geq \frac{1}{n^2}\frac{s^3}{(1-s)(2-s)^2}\alpha\,. \tag{7.29}$$

We obtain (7.25) by using the fact that $1 - s \leq 2 - s$ and

$$1 - r_{\text{step}}(\mathbf{x}, S'_\ell) \geq 1 - r_{\text{step}}(\mathcal{A}(C_1^O, \ell))$$

in (7.29). □

Theorem 48. *For all $n \geq 3$ and all $\epsilon \in (0,1)$, the ϵ-convergence time satisfies the bound*

$$T(\epsilon, \mathcal{A}(C_1^O, \ell)) \leq \frac{\log(\epsilon)}{\log\left(1 - \frac{1}{n^2}\frac{s^3}{(2-s)^3}\lambda_2\right)}\,.$$

Proof. We use the lower bound $\alpha \geq \lambda_2$ from Proposition 47 to obtain

$$1 - r_{\text{step}}(\mathcal{A}(C_1^O, \ell)) + \frac{c}{n^3} \geq \frac{1}{n^2}\frac{s^3}{1-s}\lambda_2\,.$$

We finish the proof by inserting this bound into Proposition 15. □

In this chapter, we considered Algorithm 1 with different choices of $C(\mathbf{x}, S)$. In Theorems 42 and 44, we proved that using the algorithm with $C = C_1^I$ and $C = C_1$ leads to an ϵ-convergence time that depends only for the relative size s of the meeting group. The upper bound for the ϵ-convergence time that we determined for $\mathcal{A}(C_1^I, \ell)$ is smaller than the upper bound that we determined for $\mathcal{A}(C_1, \ell)$. We found similar convergence times in our simulations when we selected meeting groups using $\mathcal{A}(C_1^I, \ell)$ and $\mathcal{A}(C_1, \ell)$. For either choice, we observed faster convergence than we observed for uniform-at-random selection of meeting groups (see Figure 7.1). We also observed that the degree of improvement in convergence time over uniform-at-random meeting-group selection of $\mathcal{A}(C_1^I, \ell)$ and $\mathcal{A}(C_1, \ell)$ depends on graph structure (see Figure 7.1).

In Theorem 48, we presented an upper bound for the ϵ-convergence time of Algorithm 1 with $C = C_1^O$ that is approximately $O(n^2)$ for fixed s. In Example 45, we showed that for disconnected graphs the function $C = C_1^O$ may not be useful for choosing a meeting group when we it in Algorithm 1. In Example 46, we showed that $1 - r_{\text{step}}(\mathcal{A}(C_1, \ell))$ may be as poor as $O(1/n)$ for connected graphs. These results indicate that using $C = C_1^O$ in Algorithm 1 is likely to have poor performance compared to using either $C = C_1^I$ or $C = C_1$ in Algorithm 1.

Additional Proofs

Proof of Lemma 37. We prove only (1). The proofs of claims (2) and (3) are similar.

The proof is by finite induction starting at $\ell = n$ and decreasing to $\ell = 2$. Let $S_n = V$. Applying Lemma 24 with S_n, we obtain

$$\sum_{v\in V} C_1(\mathbf{x}, S_n \setminus \{v\}) = -(n-2)\sum_{v,w\in V}(x_v - x_w)^2 - \frac{n-2}{n-1}\sum_{v,w\in V} A_{vw}(x_v - x_w)^2 . \tag{7.30}$$

Following Algorithm 1, we choose v_{n-1} to be the node that minimizes $C_1(\mathbf{x}, S_n \setminus \{v_{n-1}\})$. By (7.30), the set $S_{n-1} = S_n \setminus \{v_{n-1}\}$ satisfies

$$C_1^O(\mathbf{x}, S_{n-1}) \leq -\frac{n-2}{n}\sum_{v,w\in V}(x_v - x_w)^2 - \frac{n-2}{n(n-1)}\sum_{v,w\in V} A_{vw}(x_v - x_w)^2 . \tag{7.31}$$

Inserting $\ell = n-1$ into (1) yields the same expression as (7.31).

Suppose that we generate S_k for $k \in \{\ell+1, \ldots, n-1\}$ by Algorithm 1, and suppose that S_k that for $k \in \{\ell+1, \ldots, n-1\}$ satisfy (1). We seek to show that (1) holds for $k = \ell$. Applying Lemma 24 with $S = S_{\ell+1}$ yields

$$\sum_{v\in S_{\ell+1}} C_1^I(\mathbf{x}, S_{\ell+1} \setminus \{v\}) = -(\ell-1)C_1^I(\mathbf{x}, S_{\ell+1}) \tag{7.32}$$

and

$$\sum_{v\in S_{\ell+1}} C_1^O(\mathbf{x}, S_{\ell+1} \setminus \{v\}) = -\frac{\ell-1}{\ell}\sum_{v,w\in S_{\ell+1}} A_{vw}(x_v - x_w)^2 + \frac{(\ell+1)(\ell-1)}{\ell} C_1^O(\mathbf{x}, S_{\ell+1}) . \tag{7.33}$$

Replacing $\ell+1$ with ℓ in (1) and applying (7.32) and (7.33), we obtain

$$\begin{aligned}\sum_{v\in S_{\ell+1}} C_1(\mathbf{x}, S_{\ell+1} \setminus \{v\}) = &-\frac{(\ell-1)\ell(\ell+1)}{(n-1)n}\sum_{v,w\in V}(x_v - x_w)^2 \\ &- \frac{\ell-1}{\ell}\sum_{v,w\in S_{\ell+1}} A_{vw}(x_v - x_w)^2 - \sum_{k=\ell+2}^{n}\frac{(\ell-1)(\ell+1)}{(k-1)k}\sum_{v,w\in S_k} A_{vw}(x_v - x_w)^2 .\end{aligned} \tag{7.34}$$

We obtain the inequality (1) by omitting all sums over sets S_k for $k < n$ in (7.34). □

Proof of Proposition 38. We first prove (7.1). To begin, we prove a formula for the in-group cost. Let S be a set of nodes of size $k > 1$. The in-group cost $C_1^I(\mathbf{x}, S \setminus \{v\})$ is

$$C_1^I(\mathbf{x}, S \setminus \{v\}) = -2(k-1) \sum_{w \in S \setminus \{v\}} \left(x_w - \bar{x}^{S \setminus \{v\}}\right)^2 ,$$

which we write as

$$C_1^I(\mathbf{x}, S \setminus \{v\}) = -2(k-1) \sum_{w \in S} \left(x_w - \bar{x}^{S \setminus \{v\}}\right)^2 + 2(k-1) \left(x_v - \bar{x}^{S \setminus \{v\}}\right)^2 . \tag{7.35}$$

Using

$$\bar{x}^{S \setminus \{v\}} = \frac{k}{k-1} \bar{x}^S - \frac{1}{k-1} x_v , \tag{7.36}$$

we determine that

$$x_w - \bar{x}^{S \setminus \{v\}} = x_w - \bar{x}^S + \frac{1}{k-1} \left(x_v - \bar{x}^S\right) . \tag{7.37}$$

Squaring both sides of (7.37) and summing over $w \in S$, we obtain

$$\sum_{w \in S} \left(x_w - \bar{x}^{S \setminus \{v\}}\right)^2 = \sum_{w \in S} \left(x_w - \bar{x}^S\right)^2 + \frac{k}{(k-1)^2} (x_v - \bar{x}^S)^2 . \tag{7.38}$$

We also obtain the equality

$$\left(x_v - \bar{x}^{S \setminus \{v\}}\right)^2 = \left(\frac{k}{k-1}\right)^2 \left(x_v - \bar{x}^S\right)^2 \tag{7.39}$$

from (7.36).

Inserting (7.38) and (7.40) into (7.35), we obtain

$$C_1^I(\mathbf{x}, S \setminus \{v\}) = -2(k-1) \sum_{w \in S} \left(x_w - \bar{x}^S\right)^2 + 2k \left(x_v - \bar{x}^S\right)^2 .$$

Therefore, (7.1) holds.

Let $v \in S_{k-1} \cap S_k$. By (7.37), it follows that

$$x_v - \bar{x}^{S_{k-1}} = x_v - \bar{x}^{S_k} + \frac{1}{k-1} \left(x_{v_{k-1}} - \bar{x}^{S_k}\right) . \tag{7.40}$$

By (7.1), we have the inequality $|x_v - \bar{x}^{S_k}| \geq |x_{v_{k-1}} - \bar{x}^{S_k}|$. These two facts together imply that

$$\left|\left(x_v - \bar{x}^{S_{k-1}}\right) - \left(x_v - \bar{x}^{S_k}\right)\right| \leq \frac{1}{k-1}|x_v - \bar{x}^{S_k}|\,. \tag{7.41}$$

This last inequality implies that

$$x_v = \bar{x}^{S_k} \iff x_v = \bar{x}^{S_{k-1}}\,.$$

It also follows from (7.41) that if $x_v \neq \bar{x}^{S_{k-1}}$, it must also be the case that $x_v - \bar{x}^{S_{k-1}}$ and $x_v - \bar{x}^{S_k}$ must have the same sign. This proves the proposition. □

Proof of Proposition 40. We prove this result by inducation.

The statement holds when $k = n$ because $S_n = V$. Suppose that conditions (1) and (2) hold for the set S_k, where $k \in \{\ell+2, \ldots, n\}$. Let $v \in S_{k-1}$. We suppose that $x_v < \bar{x}^{S_{k-1}}$; the proof for $x_v > \bar{x}^{S_{k-1}}$ is similar. By Proposition 38, we obtain $x_v < \bar{x}^{S_k}$. Using this inequality with condition (1), we see that $w \in S_k$ for all $w \in V$ that satisfy $x_w < x_v$. It then follows that $\left(x_w - \bar{x}^{S_k}\right)^2 > \left(x_v - \bar{x}^{S_k}\right)^2$ for all $w \in V$ such that $x_w < x_v$. Therefore, by (7.1), we obtain $x_w \in S_{k-1}$. This completes the proof. □

CHAPTER 8

Converging to Consensus with Uniform-at-Random Dyadic Meetings

In Chapters 6 and 7, we allowed meetings to occur between any nodes in a graph. In that scenario, edges in a graph indicate which pairs of nodes can influence one other, instead of indicating which pairs of nodes can interact with one other. In this chapter, however, we assume that edges indicate both which pairs of nodes can influence one another and which pairs of nodes can interact with one another. We derive an upper bound for the ϵ-convergence time of the group-soft-consensus model (3.5) with meeting groups selected uniformly at random from the dyads of G.

Let U_d denote the uniform random variable on the dyads of G. We denote a dyad by $S = \{p, q\}$, and we assume that p and q in $\{p, q\}$ are ordered so that $x_p \geq x_q$. Using

$$(x_p - \bar{x}^S)^2 = \frac{1}{4}(x_p - x_q)^2 = (x_q - \bar{x}^S)^2 \,,$$

we write the costs from Propositions 2 and 4 as

$$\begin{aligned}
C_1^I(\mathbf{x}, S) &= -(x_p - x_q)^2 \,, \\
C_1^O(\mathbf{x}, S) &= -\frac{1}{2}(x_p - x_q)\left(\sum_{v \neq q} A_{pv}(x_p - x_v) + \sum_{v \neq p} A_{qv}(x_v - x_q)\right), \\
C_1(\mathbf{x}, S) &= -\frac{1}{2}(x_p - x_q)\left(\sum_{v=1}^{n} A_{pv}(x_p - x_v) + \sum_{v=1}^{n} A_{qv}(x_v - x_q)\right), \\
C_2(\mathbf{x}, S) &= \frac{k_p + k_q + 2}{4}(x_p - x_q)^2 \,.
\end{aligned} \tag{8.1}$$

In the following example, we use the formulas in (8.1) to estimate the per-step convergence factor of U_d for the path graph.

Example 49. *Let G be the path graph with n nodes $1, \ldots, n$. The edges of the path graph consist of $\{i+1, i\}$ for $i \in \{1, \ldots, n-1\}$. Let the initial vector $\mathbf{x}_0 \in \mathbb{R}^n$ have components $x_i = i$ for $i \in \{1, \ldots, n\}$.*

We use the formulas from (8.1) to calculate that $C_1(\mathbf{x}, \{i+1, i\}) = 0$ for $i \in \{2, \ldots, n-2\}$ and

$$\begin{aligned} C_1(\mathbf{x}, \{2,1\}) &= C_1(\mathbf{x}, \{n, n-1\}) = -1\,, \\ C_1^I(\mathbf{x}, \{2,1\}) &= C_1^I(\mathbf{x}, \{n, n-1\}) = -1\,, \\ C_2(\mathbf{x}, \{2,1\}) &= C_2(\mathbf{x}, \{n, n-1\}) = 5/4\,. \end{aligned}$$

Using these calculations in the formula for the optimal degree of consensus (4.4), we obtain

$$\beta(\mathbf{x}, \{i+1, i\}) = \begin{cases} \frac{4}{5}\,, & i = 1 \text{ or } i = n-1 \\ 0\,, & \text{otherwise}\,. \end{cases}$$

We also calculate that $e(\mathbf{x}) = n(n^2-1)/12$. Because

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S) = (2 - \beta(\mathbf{x}, S))\beta(\mathbf{x}, S)\frac{-C_1^I(\mathbf{x}, S)}{e(\mathbf{x})}\,,$$

it follows that $\mathfrak{r}_{\text{step}}(\mathbf{x}, \{i+1, i\}) = 1$ for $i \in \{2, \ldots, n-2\}$ and

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S) \leq \left(2 - \frac{4}{5}\right)\frac{4}{5}\frac{6}{n(n^2-1)} = \frac{144}{25}\frac{1}{n(n^2-1)}$$

for $S = \{2,1\}$ and $S = \{n, n-1\}$. Therefore,

$$\mathrm{E}\left[1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, U_d)\right] \leq \frac{2}{n-1}\frac{144}{25}\frac{1}{n(n^2-1)} = \frac{288}{25}\frac{1}{n(n-1)(n^2-1)}\,.$$

We then use the facts that $r_{\text{step}}(U_d) \geq \mathfrak{r}_{\text{step}}(\mathbf{x}, U_d)$ for all $\mathbf{x} \in \mathbb{R}^n$ and $n^2 - 1 \geq (n-1)^2$ to conclude that

$$\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right] \leq \frac{288}{25}\frac{1}{n(n-1)^3}\,. \tag{8.2}$$

In the group-soft-consensus model (3.5), individuals choose a degree of consensus to minimize the cost that they incur from disagreeing with their social connections both in and outside meeting groups. In Example 21 from Chapter 5, we considered a "hard-consensus-only" model in which individual choose a degree of consensus of 1 without considering the cost of disagreeing with their social connections. For the hard-consensus-only model with uniform-at-random dyadic meeting groups, we noted in (5.19) that

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}_d)\right] \geq \frac{2}{n(n-1)^2}\,, \tag{8.3}$$

where $r_{\text{step}}(\mathcal{U}_d)$ denotes the per-step convergence factor of the hard-consensus-only model. Comparing (8.2) and (8.3), we see that $\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right]$ is at least a factor of $\frac{144}{25}\frac{1}{n-1}$ smaller than $\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}_d)\right]$. We conclude that the constraint on the degree of consensus in the soft-consensus model (i.e. the constraint that group members minimize the cost of disagreeing with their social connections) can substantially reduce the expected decrease of the distance from consensus due to a meeting.

In the following example, we show that $\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right]$ for the path graph is even smaller than we were able to show in Example 49.

Example 50. *Let G be the path graph with $2n$ nodes labeled $1, \ldots, 2n$. The edges of G consist of $\{i, i+1\}$ for $i \in \{1, \ldots, 2n-1\}$. Suppose that the initial vector $\mathbf{x}_0 \in \mathbb{R}^n$ has components x_i that are given by*

$$\begin{aligned} x_{(n+1)+i} - x_{(n+1)+i-1} &= 1 - \frac{i^2}{n^2}\,, \\ x_{n-(i-1)} - x_{n-i} &= 1 - \frac{i^2}{n^2} \end{aligned}$$

for $i \in \{0, \ldots, n-1\}$. We calculate that

$$\begin{aligned} C_1(\mathbf{x}_0, \{(n+1)+i, (n+1)+i-1\}) &= -2(1 - \frac{i^2}{n^2})\frac{1}{n^2} = -2\frac{n^2 - i^2}{n^4}\,, \\ C_1(\mathbf{x}_0, \{n-(i-1), n-i\}) &= -2(1 - \frac{i^2}{n^2})\frac{1}{n^2} = -2\frac{n^2 - i^2}{n^4} \end{aligned} \tag{8.4}$$

for $i \in \{1, \dots, n-2\}$ *and that*

$$C_1(\mathbf{x}_0, \{2,1\}) = -2(1 - \frac{(n-1)^2}{n^2}) = \frac{2-4n}{n^2}\,,$$
$$C_1(\mathbf{x}_0, \{2n, 2n-1\}) = -2(1 - \frac{(n-1)^2}{n^2}) = \frac{2-4n}{n^2}\,. \tag{8.5}$$

It follows from formula (4.4) *for the optimal degree of consensus that*

$$\beta(\mathbf{x}_0, S) \le \frac{-C_1(\mathbf{x}_0, S)}{C_2(\mathbf{x}_0, S)} \tag{8.6}$$

for $C_1(\mathbf{x}_0, S) < 0$*. Using* (8.6) *and the fact that* $(2-\beta)\beta \le 2\beta$*, we see that*

$$1 - r_{\text{step}}(\mathbf{x}_0, S) \le \frac{-C_1(\mathbf{x}_0, S)}{C_2(\mathbf{x}_0, S)} \frac{C_1^I(\mathbf{x}_0, S)}{e(\mathbf{x}_0)}\,. \tag{8.7}$$

From the formulas for $C_1^O(\mathbf{x}_0, S)$ *and* $C_2(\mathbf{x}_0, S)$ *from* (8.1) *and the fact that* $k_i \ge 1$ *for all nodes* i *in the path graph, we obtain*

$$\frac{C_1^I(\mathbf{x}_0, S)}{C_1^O(\mathbf{x}_0, S)} \le \frac{4}{k_p + k_q + 2} \le 1\,.$$

Using this inequality in (8.7) *yields*

$$1 - r_{\text{step}}(\mathbf{x}_0, S) \le \frac{C_1(\mathbf{x}_0, S)}{e(\mathbf{x}_0)}$$

and thus

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathbf{x}_0, U_d)\right] \le \frac{\mathrm{E}\left[-C_1(\mathbf{x}_0, U_d)\right]}{e(\mathbf{x}_0)}\,. \tag{8.8}$$

As a result of the calculations in (8.4) *and* (8.5)*, we find that*

$$\mathrm{E}\left[-C_1(\mathbf{x}_0, U_d)\right] = -\frac{1}{n-1}\left(2\frac{2-4n}{n^2} + 4\sum_{i=1}^{n-2}\frac{n^2-i^2}{n^4}\right) \le \frac{c_1}{n^2} \tag{8.9}$$

for some $c_1 > 0$*. We also find that*

$$e(\mathbf{x}_0) = 1/2 + 2\sum_{i=1}^{n-1}\left(\sum_{j=1}^{i}\left(1 - \frac{j^2}{n^2}\right)\right)^2 \ge c_2 n^3 \tag{8.10}$$

for some $c_2 > 0$*. Inserting the bounds* (8.9) *and* (8.10) *into* (8.8)*, we conclude that*

$$\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right] \le \frac{c_1}{c_2}\frac{1}{n^5}\,. \tag{8.11}$$

From the bounds (8.11) and (8.3), we conclude that there exists a constant $c > 0$ such that $\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right]$ is at least a factor of c/n^2 smaller than $\mathrm{E}\left[1 - r_{\text{step}}(\mathcal{U}_d)\right]$ for the path graph with n nodes. Consequently, the constraint in the soft-consensus model that group members minimize the cost of disagreeing with their social connections can reduce the expected decrease of the distance from consensus of a meeting by at least c/n^2 for some constant $c > 0$.

Given (8.11) from Example 50, the best bound for $T(\epsilon, U_d)$ that depends only on ϵ and n that one can obtain from Proposition 16 is

$$\frac{2\log(\epsilon)}{\log\left(1 - \frac{c}{n^5}\right)} = \frac{2}{c}\log(\epsilon^{-1})n^5 + O(1) \tag{8.12}$$

for some constant $c > 0$. (The asymptotic expression holds as $n \to \infty$ and is a result of using (5.15) from Remark 17.) Turning to the hard-consensus-only model of Example 21, we obtain the bound

$$\frac{2\log(\epsilon)}{\log\left(1 - \frac{2}{n(n-1)^2}\right)} \leq \log(\epsilon^{-1})n^3 + O(1) \tag{8.13}$$

for $T(\epsilon, \mathcal{U}_d)$ by using (5.19) from Proposition 16. (The asymptotic expression holds as $n \to \infty$ and is a result of using (5.15) from Remark 17.) The bounds (8.12) and (8.13) indicate that, for uniform-at-random dyadic updates, the group-soft-consensus model (3.5) may take approximately n^2 times as much time to converge as the hard-consensus-only model.

In Figure 8.1, we show simulated expected ϵ-convergence times of $T(\epsilon, U_d)$ and $T(\epsilon, \mathcal{U}_d)$ for the path graph and initial vector from Example 50. Our procedure for simulating the expected value of $T(\epsilon, U_d)$ on a graph G is to (1) run 100 trials of the group-soft-consensus model (3.5) with $S^t = U_d$ and record the first time t such that $e(\mathbf{x}(t)/\mathbf{x}(0)) < \epsilon$ for each trial and (2) take the mean of the recorded times from step (1). We follow the same procedure to simulate the expected value of $T(\epsilon, \mathcal{U}_d)$ except that we use the hard-consensus-only model of Example 21. We find that the simulated values $\mathrm{E}\left[T(\epsilon, U_d)\right]$ have a quintic dependence on the number n of nodes. By contrast, we find that the simulated values of $\mathrm{E}\left[T(\epsilon, \mathcal{U}_d)\right]$ exhibit a cubic dependence on n. These dependences on n align with the convergence-time behavior that is indicated by the bounds in (8.12) and (8.13).

We now prove an upper bound for $T(\epsilon, U_d)$. The key to our proof is a method to find a dyad $\{p, q\}$ such that we can lower-bound $1 - \imath_{\text{step}}(\mathbf{x}, \{p, q\})$; we state the method in Algorithm 2. The input of the method is a dyad $\{v, w\}$ and a vector $\mathbf{x} \in \mathbb{R}^n$. Let P denote the simple path (w, v).[1] The algorithm begins by searching through the neighbors of v for a node v' whose belief is maximal among those neighbors. If the belief of v' is larger than v, we place v' at the end of P. We then do the same for w, except that we search for a node w' whose belief is minimal among the neighbors of w. If the node w' has a belief smaller than w, we place w' at the start of P. At this point, we repeat the process with the starting and ending nodes of P in place of v and w, respectively. we repeat this process until there are no available nodes to add to the path. The dyad that is returned by the algorithm is the pair of adjacent nodes in P that maximizes

$$\beta(\mathbf{x}, \{p, q\})(x_p - x_q)^2 \,.$$

For the remainder of the chapter, we denote the dyad that is returned by Algorithm 2 for an input dyad $\{u, v\}$ by $S_{u,v}$, and we let the letters p and q denote the nodes in $S_{u,v}$ (i.e. $S_{u,v} = \{p, q\}$).

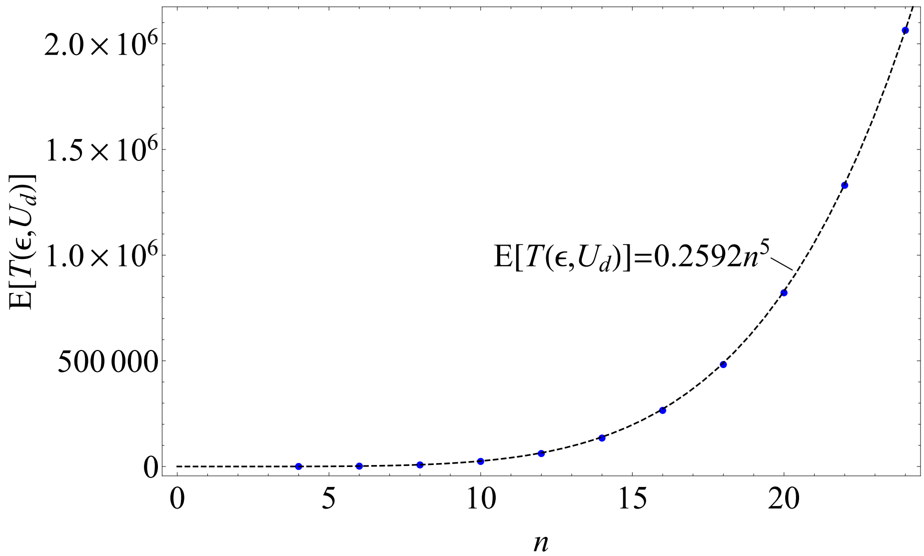
Using the formulas in (8.1), we write (5.22) from Chapter 5 for the special case of a dyad $S = \{u, v\}$ to obtain the lower bound

$$1 - \imath_{\text{step}}(\mathbf{x}, S) \geq \begin{cases} \frac{1}{2}\frac{(x_u - x_v)^2}{e(\mathbf{x})}\,, & -C_1(\mathbf{x}, S) \geq C_2(\mathbf{x}, S) \\ -\frac{2}{k_u + k_v + 2}\frac{C_1(\mathbf{x}, S)}{e(\mathbf{x})}\,, & 0 \leq -C_1(\mathbf{x}, S) \leq C_2(\mathbf{x}, S) \\ 0\,, & -C_1(\mathbf{x}, S) \leq 0\,. \end{cases} \tag{8.14}$$

In light of (8.14), we can determine a lower bound for $1 - \imath_{\text{step}}(\mathbf{x}, S_{u,v})$ by lower-bounding

$$\min\left(\frac{1}{2}(x_p - x_q)^2, -\frac{2}{k_p + k_q + 2}C_1(\mathbf{x}, \{p, q\})\right)\,. \tag{8.15}$$

[1]A *path* is a sequence of nodes with the property that each pair of consecutive nodes is adjacent in G. A *simple path* is a path that does not include repeated nodes.



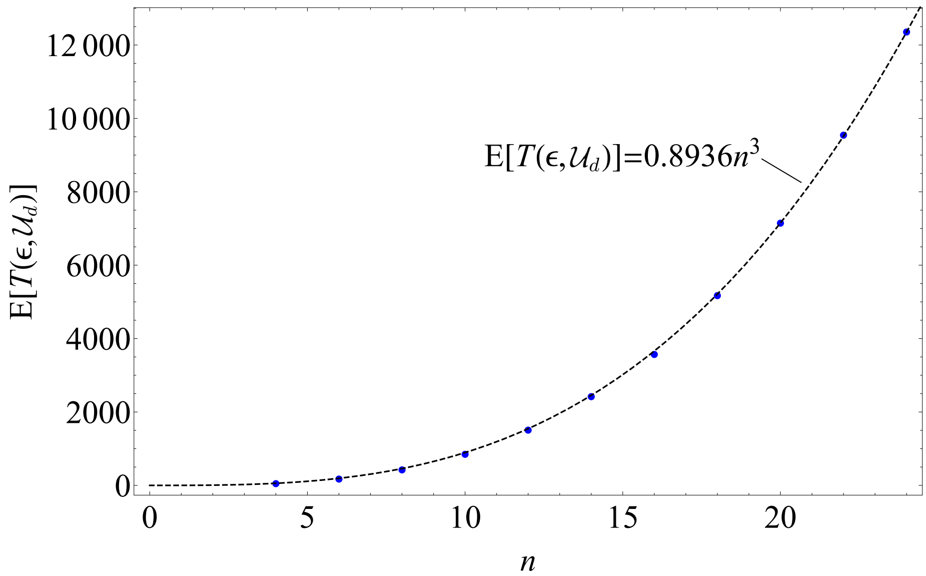


Figure 8.1: Simulated expected values of (top) $T(\epsilon, U_d)$ and (bottom) $T(\epsilon, \mathcal{U}_d)$ for the path graph with n nodes for $n \in \{4, 6, 8, \ldots, 24\}$ with the initial vector $\mathbf{x}_0$ from Example 50. We include the the dotted curves (top) $\mathrm{E}\left[T(\epsilon, U_d)\right] = 0.2591n^5$ and (bottom) $\mathrm{E}\left[T(\epsilon, \mathcal{U}_d)\right] = 0.8936n^3$ to exhibit the dependence of the simulation data on n.

Algorithm 2: FindDyad

Input: Dyad $\{v, w\}$ with $x_w > x_v$ and vector $\mathbf{x} \in \mathbb{R}^n$

Output: Dyad $\{p, q\}$

$MaxUpdate \longleftarrow \beta(\mathbf{x}, \{v, w\})(x_w - x_v)^2$
$MaxDyad \longleftarrow \{v, w\}$
$v' \longleftarrow \arg\min_{v \in \mathcal{N}(v)} x_v$
$w' \longleftarrow \arg\max_{v \in \mathcal{N}(w)} x_v$
while $x_{w'} > x_w$ or $x_{v'} < x_v$ **do**
    if $x_{v'} < x_v$ **then**
        if $\beta(\mathbf{x}, \{v', v\})(x_v - x_{v'})^2 > MaxUpdate$ **then**
            $MaxUpdate \longleftarrow \beta(\mathbf{x}, \{v', v\})(x_v - x_{v'})^2$
            $MaxDyad \longleftarrow \{v', v\}$
        end
        $v \longleftarrow v'$
    end
    if $x_{w'} > x_w$ **then**
        if $\beta(\mathbf{x}, \{w, w'\})(x_{w'} - x_w)^2 > MaxUpdate$ **then**
            $MaxUpdate \longleftarrow \beta(\mathbf{x}, \{w, w'\})(x_{w'} - x_w)^2$
            $MaxDyad \longleftarrow \{w, w'\}$
        end
        $w \longleftarrow w'$
    end
    $v' \longleftarrow \arg\min_{v \in \mathcal{N}(v)} x_v$
    $w' \longleftarrow \arg\max_{v \in \mathcal{N}(w)} x_v$
end
return $MaxDyad$

We state a lower bound of (8.15) in the following lemma.

Lemma 51. *Let $k_{\max}$ denote the maximum degree, and let k_{ave} denote the mean degree of the nodes in G. It follows that*

$$1 - r_{\text{step}}(\mathbf{x}, S_{u,v}) \geq \frac{1}{8}\frac{1}{n^3}\left(\frac{2}{k_{\max} + k_{\text{ave}}}\right)^{2n}\frac{(x_u - x_v)^2}{e(\mathbf{x})}. \tag{8.16}$$

From the result of Lemma 51, we obtain the following bound for $\mathrm{E}\left[1 - r_{\text{step}}(S_{u,v})\right]$, where we take the expectation over the dyads $\{u, v\}$.

Proposition 52. *Denote the algebraic connectivity of G by $\lambda_2(\mathbf{L})$. Using the notation of Lemma 51, the following bound holds:*

$$\mathrm{E}\left[1 - r_{\text{step}}(U_d)\right] \geq \frac{1}{8}\frac{\lambda_2(\mathbf{L})}{n^3 m^2}\left(\frac{2}{k_{\max} + k_{\text{ave}}}\right)^{2n}. \tag{8.17}$$

Proof. Denote the expectation with respect to the random variable U_d by $\mathrm{E}_d\left[\cdot\right]$. From Lemma 20, we obtain

$$\mathrm{E}_d\left[(x_u - x_v)^2\right] = \frac{1}{m}\sum_{\{u,v\}\in E}(x_u - x_v)^2 = \frac{1}{m}\mathbf{x}^T\mathbf{L}\mathbf{x} \geq \frac{1}{m}\lambda_2(\mathbf{L})e(\mathbf{x}). \tag{8.18}$$

Therefore, there exists a dyad $\{u^*, v^*\}$ such that

$$(u^* - v^*)^2 \geq \frac{1}{m}\lambda_2(\mathbf{L})e(\mathbf{x}).$$

By Lemma 51, it then follows that

$$1 - r_{\text{step}}(\mathbf{x}, S_{u^*,v^*}) \geq \frac{1}{8}\frac{\lambda_2(\mathbf{L})}{n^3 m}\left(\frac{2}{k_{\max} + k_{\text{ave}}}\right)^{2n}.$$

Therefore,

$$\mathrm{E}\left[1 - r_{\text{step}}(\mathbf{x}, U_d)\right] \geq \frac{1}{m}(1 - r_{\text{step}}(\mathbf{x}, S_{u^*,v^*})) \geq \frac{1}{8}\frac{\lambda_2(\mathbf{L})}{n^3 m^2}\left(\frac{2}{k_{\max} + k_{\text{ave}}}\right)^{2n}$$

for all $\mathbf{x} \in \mathbb{R}^n$. This proves Proposition 52. □

Using Proposition 52, we determine an upper bound for $T(\epsilon, U_d)$.

Theorem 53. *Let $k_{\max}$ denote the maximum degree of the nodes in G, and let k_{ave} denote the mean degree of the nodes in G. For all $n \geq 3$ and all $\epsilon \in (0,1)$, the ϵ-convergence time $T(\epsilon, U_d)$ satisfies the bound*

$$T(\epsilon, S_d) \leq \frac{2\log(\epsilon)}{\log\left(1 - \frac{1}{8}\frac{\lambda_2(\mathbf{L})}{n^3 m^2}\left(\frac{2}{k_{\max}+k_{\text{ave}}}\right)^{2n}\right)}\,.$$

Proof. We obtain this bound by using (8.17) from Proposition 52 in Proposition 16. □

By (5.15) from Remark 17, we write the ϵ-convergence time upper bound of Theorem 53 as

$$\frac{2\log(\epsilon)}{\log\left(1 - \frac{1}{8}\frac{\lambda_2(\mathbf{L})}{n^3 m^2}\left(\frac{2}{k_{\max}+k_{\text{ave}}}\right)^{2n}\right)} = 16\log(\epsilon^{-1})\frac{n^3 m^2}{\lambda_2(L)}\left(\frac{k_{\max}+k_{\text{ave}}}{2}\right)^{2n} + O(1)\,. \tag{8.19}$$

The bound in (8.19) depends exponentially on n. Consequently, the upper bound in Theorem 53 is extremely large for all but the smallest values of n. Determining whether there exist graphs for which the ϵ-convergence time depends exponentially on n is an important task for future work.

Additional Proofs

The primary goal of this section is to prove Lemma 51. Before we begin the proof, we introduce some definitions.

Definition 54. *We say that a node $v \in V$ is locally minimal (respectively, locally maximal) if $x_w \geq x_v$ (respectively, $x_w \leq x_v$) for all $w \in \mathcal{N}(v)$.*

Definition 55. *Let $P = (p_{-k}, \ldots, p_0, p_1, \ldots, p_{k+1})$ denote a sequence of nodes in a graph G. Denote*

$$p_k^{<} = p_{\max\{\ell<k\,:\,p_\ell \neq p_k\}}\,,$$

$$p_k^{>} = p_{\min\{\ell>k\,:\,p_\ell \neq p_k\}}$$

to be the previous and next nodes, respectively, that are distinct from p_k in the sequence. We call P a lazy simple path *if the sequence of nodes that we obtain from P by removing all consecutive, repeated nodes is a simple path. We denote the number of distinct nodes in P by $\mathcal{L}(P)$.*

Definition 56. *Let $G^+(\mathbf{x})$ denote the directed graph with node set V and edge set $\{(v,w) : \{v,w\} \in E,\ x_v > x_w\}$. Similarly, let $G^-(\mathbf{x})$ denote the directed graph with node set V and edge set $\{(v,w) : \{v,w\} \in E, x_v < x_w\}$.*

We say that a lazy path $P = (p_0, p_1, \ldots, p_k)$ is a monotone lazy path *if P is a lazy path in $G^+(\mathbf{x})$ or $G^-(\mathbf{x})$.*

We introduce a slight variation of Algorithm 2 in Algorithm 3. Algorithms 2 and 3 return the same values for the same inputs; however, Algorithm 3 generates a lazy path of the nodes that it finds when it computes this value. Using the notation of Algorithm 3, we denote this lazy path by $\mathcal{P} = (p_{-k}, \ldots, p_{k+1})$. In the following proposition, we state some properties of $\mathcal{P}$.

Proposition 57. *The following properties of $\mathcal{P}$ hold:*

(a) $\mathcal{P}$ is a lazy monotone path.

(b) Either (i) $p_{-\ell} \neq p_{-(\ell-1)}$ and $p_{\ell+1} = p_\ell$ or (ii) $p_{-\ell} = p_{-(\ell-1)}$ and $p_{\ell+1} \neq p_\ell$.

(c) $\mathcal{L}(\mathcal{P}) = k + 2$ (see Definition 55).

(d) The nodes p_{-k} and p_{k+1} are locally minimal and locally maximal nodes, respectively, in the sense of Definition 54.

(e) The lazy path $\mathcal{P}$ can include at most k copies of a node.

We use the following lemma in our proof of Lemma 59.

Algorithm 3: FindDyadPath

Input: Dyad $\{p_1, p_0\}$ and vector $\mathbf{x} \in \mathbb{R}^n$

Output: Dyad $\{p, q\}$

$k \longleftarrow 1$

$p'_{k+1} \longleftarrow \arg\max_{v \in \mathcal{N}(p_k)} x_v$

$p'_{-k} \longleftarrow \arg\min_{v \in \mathcal{N}(p_{-(k-1)})} x_v$

while $x_{p'_{k+1}} > x_{p_k}$ or $x_{p_{-(k-1)}} > x_{p'_{-k}}$ **do**

$$(p_{k+1}, p_{-k}) \longleftarrow \begin{cases} (p'_{k+1}, p_{-k})\,, & \text{if } x_{p'_{k+1}} - x_{p_k} \geq x_{p_{-(k-1)}} - x_{p'_{-k}} \\ (p_k, p'_{-k})\,, & \text{if } x_{p_{-(k-1)}} - x_{p'_{-k}} > x_{p'_{k+1}} - x_{p_k} \end{cases}$$

$k \longleftarrow k + 1$

$p'_{k+1} \longleftarrow \arg\max_{v \in \mathcal{N}(p_k)} x_v$

$p'_{-k} \longleftarrow \arg\min_{v \in \mathcal{N}(p_{-(k-1)})} x_v$

end

return $\arg\max_{\substack{\{p_\ell, p_\ell^<\} \\ -k<\ell\leq k}} \left(\beta(\mathbf{x}, \{p_\ell, p_\ell^<\})(x_{p_\ell} - x_{p_\ell^<})^2\right)$

Lemma 58. *Let $P = (p_{-k}, p_{-(k-1)}, \ldots, p_k, p_{k+1})$ be a lazy path, let γ be a positive real number, and assume that $\mathcal{L}(P) \geq 2$. If*

$$\sum_{j=1}^{n} A_{p_{\ell+1}j}(x_j - x_{p_{\ell+1}}) + \sum_{j=1}^{n} A_{p_{\ell+1}^<j}(x_{p_{\ell+1}^<} - x_j) > -\gamma \tag{8.20}$$

for all $\ell \in \{-(k-1), \ldots, k\}$, then

$$\sum_{j=1}^{n} A_{p_{k+1}j}(x_j - x_{p_{k+1}}) + \sum_{j=1}^{n} A_{p_{-k}j}(x_{p_{-k}} - x_j) > -(\mathcal{L}(P) - 1)\gamma\,. \tag{8.21}$$

Proof. Let $(r_1, \ldots, r_{\mathcal{L}(P)})$ denote the path that we obtain from P by removing consecutive, repeated nodes. Note that $r_1 = p_{-k}$ and $r_{\mathcal{L}(P)} = p_{k+1}$. From (8.20), for all $\ell \in \{1, \ldots, \mathcal{L}(P) - 1\}$, the following inequality holds:

$$\sum_{j=1}^{n} A_{r_{\ell+1}j}(x_j - x_{r_{\ell+1}}) + \sum_{j=1}^{n} A_{r_\ell j}(x_{r_\ell} - x_j) > -\gamma\,.$$

Summing this inequality over $\ell = 1, \ldots, \mathcal{L}(P) - 1$, we obtain

$$\sum_{j=1}^{n} A_{r_{\mathcal{L}(P)}j}(x_j - x_{r_{\mathcal{L}(P)}}) + \sum_{j=1}^{n} A_{r_1 j}(x_{r_1} - x_j) > -(\mathcal{L}(P) - 1)\gamma\,.$$

This last inequality is the same as (8.21). ☐

We now state and prove our main lemma for upper-bounding $T(\epsilon, U_d)$. Let A^+ denote the adjacency matrix of the directed graph $G^+(\mathbf{x})$, and let A^- denote the adjacency matrix of $G^-(x)$ (see Definition 56). Additionally, we denote the in-degree of a node $v \in V$ in the graph $G^+(\mathbf{x})$ by $k_v^+ = \sum_{j=1}^n A_{vj}^+$ and denote the in-degree in the graph $G^-(x)$ by $k_v^- = \sum_{j=1}^n A_{vj}^-$.

Lemma 59. *Let $\{p_1, p_0\}$ be a dyad from G, and let*

$$\delta(\mathcal{P}) = \frac{\frac{2^{\mathcal{L}(\mathcal{P})-1}}{\prod_{\ell=0}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{\ell+1}} + k^-_{p_{-\ell}} + 1)}}{\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{\ell=m}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{\ell+1}} + k^-_{p_{-\ell}} + 1)} + (\mathcal{L}(\mathcal{P}) - 1)}\,. \tag{8.22}$$

It then follows that there exists a dyad $\{p, q\}$ such that p and q are consecutive nodes in the path $\mathcal{P}$ and

$$\min\left(\frac{1}{2}(x_p - x_q)^2, -\frac{2}{k_p + k_q + 2}C_1(\mathbf{x}, \{p, q\})\right) \geq \frac{\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P}) - 1)}{k_p + k_q + 2}(x_{p_1} - x_{p_0})^2\,. \tag{8.23}$$

Proof. The proof is by finite induction. Let $\delta > 0$, and assume that

$$\sum_{j=1}^{n} A_{p_1 j}(x_w - x_{p_1}) + \sum_{j=1}^{n} A_{p_0 j}(x_{p_0} - x_w) > -\delta(x_{p_1} - x_{p_0})\,. \tag{8.24}$$

We rewrite (8.24) in the form

$$\begin{aligned}\sum_{j=1}^{n} A^+_{p_1 j}(x_w - x_{p_1}) + \sum_{j=1}^{n} A^-_{p_0 j}(x_{p_0} - x_w) \\ > -\delta(x_{p_1} - x_{p_0}) + 2(x_{p_1} - x_{p_0}) + \sum_{\substack{j=1 \\ j \neq p_0}}^{n} A^-_{p_1 j}(x_{p_1} - x_w) + \sum_{\substack{j=1 \\ j \neq p_1}}^{n} A^+_{p_0 j}(x_w - x_{p_0})\,.\end{aligned} \tag{8.25}$$

Using

$$\sum_{\substack{j=1\\ j\neq p_0}}^{n} A^-_{p_1 j}(x_{p_1}-x_w)\geq 0\,,$$
$$\sum_{\substack{j=1\\ j\neq p_1}}^{n} A^+_{p_0 j}(x_w-x_{p_0})\geq 0$$

in (8.24), we obtain

$$\sum_{j=1}^{n} A^+_{p_1 j}(x_w-x_{p_1})+\sum_{j=1}^{n} A^-_{p_0 j}(x_{p_0}-x_w)>(2-\delta)(x_{p_1}-x_{p_0})\,. \tag{8.26}$$

Let p_2 and p_{-1} be defined as in line 5 from Algorithm 3. By (8.26) and the definitions of p_2 and p_{-1}, exactly one of the following holds:

$$x_{p_2}-x_{p_1}>\frac{1}{k^+_{p_1}+k^-_{p_0}}(2-\delta)(x_{p_1}-x_{p_0})\,,$$
$$x_{p_0}-x_{p_{-1}}>\frac{1}{k^+_{p_1}+k^-_{p_0}}(2-\delta)(x_{p_1}-x_{p_0})\,. \tag{8.27}$$

Note that

$$\frac{2-\delta}{k^+_{p_1}+k^-_{p_0}+1}<1\,, \tag{8.28}$$

because $k^+_{p_1}+k^-_{p_0}\geq 1$ and $\delta>0$. It follows from (8.28) that

$$x_{p_1}-x_{p_0}>\frac{2-\delta}{k^+_{p_1}+k^-_{p_0}+1}(x_{p_1}-x_{p_0})\,. \tag{8.29}$$

Consequently, from (8.29) and (8.27), we see that

$$x_{p_2}-x_{p_2^<}>\frac{1}{k^+_{p_1}+k^-_{p_0}+1}(2-\delta)(x_{p_1}-x_{p_0})\,, \tag{8.30}$$
$$x_{p_{-1}^\geq}-x_{p_{-1}}>\frac{1}{k^+_{p_1}+k^-_{p_0}+1}(2-\delta)(x_{p_1}-x_{p_0})\,.$$

Let $P_k=(p_{-k},\ldots,p_{k+1})$ denote the lazy path of nodes that is generated by Algorithm 2 in the k^{th} step. Assume for induction that

$$\sum_{j=1}^{n} A_{p_{l+1}j}(x_j-x_{p_{l+1}})+\sum_{j=1}^{n} A_{p^<_{l+1}j}(x_{p^<_{l+1}}-x_j)>-\delta(x_{p_1}-x_{p_0}) \tag{8.31}$$

for all $l \in \{-k, \ldots, k\}$. Additionally, assume that

$$\begin{aligned} & x_{p_{k+1}} - x_{p^<_{k+1}} \\ & \quad > \left(\frac{2^k}{\prod_{l=0}^{k-1}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k-1} \frac{(m+1)2^{k-m-1}}{\prod_{l=m}^{k-1}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0}) \,, \\ & x_{p^>_{-k}} - x_{p_{-k}} \\ & \quad > \left(\frac{2^k}{\prod_{l=0}^{k-1}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k-1} \frac{(m+1)2^{k-m-1}}{\prod_{l=m}^{k-1}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0}) \,. \end{aligned} \tag{8.32}$$

Using (8.31) in Lemma 58 yields

$$\begin{aligned} \sum_{j=1}^{n} A_{p_{k+1}j}(x_j - x_{p_{k+1}}) + \sum_{j=1}^{n} A_{p_{-k}j}(x_{p_{-k}} - x_j) &> -(L(P_k) - 1)\delta(x_{p_1} - x_{p_0}) \\ &= -(k+1)\delta(x_{p_1} - x_{p_0}) \,, \end{aligned} \tag{8.33}$$

where the last equality is due to (c) from Proposition 57.

Using

$$\begin{aligned} \sum_{\substack{j=1 \\ j \neq p^<_{k+1}}}^{n} A^-_{p_{k+1}j}(x_{p_{k+1}} - x_j) &\geq 0 \,, \\ \sum_{\substack{j=1 \\ j \neq p^>_{-k}}}^{n} A^+_{p_{-k}j}(x_j - x_{p_{-k}}) &\geq 0 \,, \end{aligned}$$

and (8.33), we obtain

$$\begin{aligned} & \sum_{j=1}^{n} A^+_{p_{k+1}j}(x_j - x_{p_{k+1}}) + \sum_{j=1}^{n} A^-_{p_{-k}j}(x_{p_{-k}} - x_j) \\ & \qquad > -(k+1)\delta(x_{p_1} - x_{p_0}) + (x_{p_{k+1}} - x^<_{p_{k+1}}) + (x^>_{p_{-k}} - x_{p_{-k}}) \,. \end{aligned}$$

Combining this last inequality with (8.32), we find that

$$\begin{aligned}
&\sum_{j=1}^{n} A^{+}_{p_{k+1}j}(x_j - x_{p_{k+1}}) + \sum_{j=1}^{n} A^{-}_{p_{-k}j}(x_{p_{-k}} - x_j) \\
&\qquad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k-1} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} + (k+1) \right) \right) (x_{p_1} - x_{p_0})\,.
\end{aligned} \tag{8.34}$$

Let p_{k+2} and $p_{-(k+1)}$ be defined as in line 5 from Algorithm 3. By (8.34) and the definitions of p_{k+2} and $p_{-(k+1)}$, exactly one of the following holds:

$$\begin{aligned}
&x_{p_{k+2}} - x_{p_{k+1}} \\
&\qquad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0})\,, \\
&x_{p_{-k}} - x_{p_{-(k+1)}} \\
&\qquad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0})\,.
\end{aligned} \tag{8.35}$$

For notational convenience, we let

$$\prod_{l=k}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1) = 1\,. \tag{8.36}$$

Using the convention (8.36), we have

$$\sum_{m=0}^{k-1} \frac{(m+1)2^{k-m-1}}{\prod_{l=m}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} \le \sum_{m=0}^{k} \frac{(m+1)2^{k-m-1}}{\prod_{l=m}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)}\,. \tag{8.37}$$

From (8.37) and

$$\frac{2}{k^{+}_{p_{k+1}} + k^{-}_{q_{-k}} + 1} \le 1\,,$$

we obtain

$$\begin{aligned}
&\left(\frac{2^{k}}{\prod_{l=0}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k-1} \frac{(m+1)2^{k-m-1}}{\prod_{l=m}^{k-1}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0}) \\
&\qquad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k}(k^{+}_{p_{l+1}} + k^{-}_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0})\,.
\end{aligned}$$

Therefore, from the definitions of p_{k+2} and $p_{-(k+1)}$ and by (8.35), we conclude that

$$\begin{aligned} & x_{p_{k+2}} - x_{p^<_{k+2}} \\ & \quad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0}), \\ & x_{p^>_{-(k+1)}} - x_{p_{-(k+1)}} \\ & \quad > \left(\frac{2^{k+1}}{\prod_{l=0}^{k}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} - \delta \left(\sum_{m=0}^{k} \frac{(m+1)2^{k-m}}{\prod_{l=m}^{k}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} \right) \right) (x_{p_1} - x_{p_0}) \,. \end{aligned} \tag{8.38}$$

By (d) from Proposition 57, Algorithm 3 must terminate and it must generate a lazy monotone path $\mathcal{P}$ between a locally minimal node $p_{-(\mathcal{L}(\mathcal{P})-2)}$ and a locally maximal node $p_{\mathcal{L}(\mathcal{P})-1}$. Applying (8.34) with $k = \mathcal{L}(\mathcal{P}) - 2$, we obtain

$$\begin{aligned} 0 > & \left(\frac{2^{\mathcal{L}(\mathcal{P})-1}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} \right. \\ & \left. -\delta \left(\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} + (\mathcal{L}(\mathcal{P}) - 1) \right) \right) (x_{p_1} - x_{p_0}) \end{aligned} \tag{8.39}$$

because $p_{-(\mathcal{L}(\mathcal{P})-2)}$ and $p_{\mathcal{L}(\mathcal{P})-1}$ are locally minimal and locally maximal nodes, respectively. However, taking

$$\delta(\mathcal{P}) = \frac{\frac{2^{\mathcal{L}(\mathcal{P})-1}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)}}{\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}} + k^-_{p_{-l}} + 1)} + (\mathcal{L}(\mathcal{P}) - 1)} \tag{8.40}$$

in (8.39) leads to a contradiction. Therefore, the assumption that

$$\sum_{j=1}^{n} A_{p_{l+1}j}(x_j - x_{p_{l+1}}) + \sum_{j=1}^{n} A_{p^<_{l+1}j}(x_{p^<_{l+1}} - x_j) > -\delta(\mathcal{P})(x_{p_1} - x_{p_0}) \tag{8.41}$$

for all $l \in \{-(\mathcal{L}(\mathcal{P})-3), \dots, \mathcal{L}(\mathcal{P})-1\}$ leads to a contradiction, so there exists a dyad $\{p, q\}$ of distinct consecutive nodes in $\mathcal{P}$ such that

$$\sum_{j=1}^{n} A_{pj}(x_j - x_p) + \sum_{j=1}^{n} A_{qj}(x_q - x_j) \leq -\delta(\mathcal{P})(x_{p_1} - x_{p_0}) \,. \tag{8.42}$$

For some step k', we have either (i) $p_{k'} = p$ and $p_{k'}^{<} = q$ or (ii) $p_{-k'} = q$ and $p_{-k'}^{>} = p$. For this value of k', we use the fact that the right-hand side of (8.38) is strictly decreasing in k to replace k' with $\mathcal{L}(\mathcal{P}) - 3$ in the right-hand side and obtain

$$x_p - x_q > \tag{8.43}$$
$$\left(\frac{2^{\mathcal{L}(\mathcal{P})-2}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+} + k_{p_{-l}}^{-} + 1)} - \delta(\mathcal{P}) \left(\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-3-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+} + k_{p_{-l}}^{-} + 1)} \right) \right) (x_{p_1} - x_{p_0}) \,.$$

We insert the definition of $\delta(\mathcal{P})$ into the right-hand side of (8.43) to obtain

$$\frac{2^{\mathcal{L}(\mathcal{P})-2}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+} + k_{p_{-l}}^{-} + 1)} \left(1 - 2 \frac{\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-3-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+}+k_{p_{-l}}^{-}+1)}}{\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+}+k_{p_{-l}}^{-}+1)} + (\mathcal{L}(\mathcal{P}) - 1)} \right) (x_{p_1} - x_{p_0}) \,, \tag{8.44}$$

and we simplify (8.44) to conclude that

$$x_p - x_q > \frac{2^{\mathcal{L}(\mathcal{P})-2}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+} + k_{p_{-l}}^{-} + 1)} \frac{\mathcal{L}(\mathcal{P}) - 1}{\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k_{p_{l+1}}^{+}+k_{p_{-l}}^{-}+1)} + (\mathcal{L}(\mathcal{P}) - 1)} (x_{p_1} - x_{p_0}) \,. \tag{8.45}$$

Using the definition of $\delta(\mathcal{P})$ in (8.45), we obtain

$$x_p - x_q > \frac{\delta(\mathcal{P})(\mathcal{L}(\mathcal{P}) - 1)}{2} (x_{p_1} - x_{p_0}) \,. \tag{8.46}$$

Therefore,

$$(x_p - x_q)^2 > \frac{\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P}) - 1)^2}{4} (x_{p_1} - x_{p_0})^2 \,. \tag{8.47}$$

Using $\mathcal{L}(\mathcal{P}) - 1 \geq 1$ and $k_p + k_q + 2 \geq 4$ in (8.48), we obtain

$$(x_p - x_q)^2 > \frac{\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P}) - 1)}{k_p + k_q + 2} (x_{p_1} - x_{p_0})^2 \,. \tag{8.48}$$

Using the formula for $C_1(\mathbf{x}, \{p, q\})$ from (8.1) along with (8.42), we find that

$$-\frac{2}{k_p + k_q + 2} C_1(\mathbf{x}, \{p, q\}) > \frac{\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P}) - 1)}{k_p + k_q + 2} (x_{p_1} - x_{p_0})^2 \,. \tag{8.49}$$

The inequalities (8.49) and (8.48) imply the inequality (8.23). □

Proof of Lemma 51. Note that

$$\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)} = \sum_{m=0}^{\mathcal{L}(\mathcal{P})-3}(m+1)\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}\frac{2}{(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)}\,. \tag{8.50}$$

Using the inequality $k_{p_l}+k_{q_{-l}}+1\geq 2$ in the right-hand side of (8.50), we obtain

$$\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3}(m+1)\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}\frac{2}{(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)} \leq \sum_{m=0}^{\mathcal{L}(\mathcal{P})-3}(m+1) = \frac{(\mathcal{L}(\mathcal{P})-2)(\mathcal{L}(\mathcal{P})-1)}{2}\,. \tag{8.51}$$

We combine (8.50) and (8.51) to deduce that

$$\sum_{m=0}^{\mathcal{L}(\mathcal{P})-3} \frac{(m+1)2^{\mathcal{L}(\mathcal{P})-2-m}}{\prod_{l=m}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)} \leq \frac{(\mathcal{L}(\mathcal{P})-2)(\mathcal{L}(\mathcal{P})-1)}{2}\,. \tag{8.52}$$

Let $k_{\max}(\mathcal{P})$ denote the maximum degree of the nodes in the path $\mathcal{P}$, and let $k_{\text{tot}}(\mathcal{P})$ denote the sum of the degrees of the nodes in $\mathcal{P}$. As a result of (c) and (e) from Proposition 57, the lazy path $\mathcal{P}$ can have at most $\mathcal{L}(\mathcal{P})-1$ copies of a node and it must have at least one copy of every node. It then follows that

$$\sum_{l=-k}^{k+1} k_{p_l} \leq (\mathcal{L}(\mathcal{P})-1)k_{\max}(\mathcal{P}) + (k_{\text{tot}}(\mathcal{P})-k_{\max}(\mathcal{P}))\,. \tag{8.53}$$

We also use $k^+_{p_{l+1}}+k^-_{p_{-l}}+1 < k_{p_{l+1}}+k_{p_{-l}}$ to obtain

$$\prod_{l=0}^{\mathcal{L}(\mathcal{P})-2}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1) < \prod_{l=0}^{\mathcal{L}(\mathcal{P})-2}(k_{p_{l+1}}+k_{p_{-l}})\,. \tag{8.54}$$

Consequently, using (8.53), (8.54), and the inequality of arithmetic and geometric means, it follows that

$$\begin{aligned}\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1) &< \prod_{l=0}^{\mathcal{L}(\mathcal{P})-2}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)(2k_{\max}(\mathcal{P})) \\ &\leq \left(\frac{\mathcal{L}(\mathcal{P})k_{\max}(\mathcal{P})+k_{\text{tot}}(\mathcal{P})}{\mathcal{L}(\mathcal{P})}\right)^{\mathcal{L}(\mathcal{P})}\,.\end{aligned} \tag{8.55}$$

As a result of (8.55), we obtain

$$\frac{2^{\mathcal{L}(\mathcal{P})-2}}{\prod_{l=0}^{\mathcal{L}(\mathcal{P})-3}(k^+_{p_{l+1}}+k^-_{p_{-l}}+1)} > \frac{1}{4}\left(\frac{2}{k_{\max}(\mathcal{P})+k_{\text{ave}}(\mathcal{P})}\right)^{\mathcal{L}(\mathcal{P})}\,. \tag{8.56}$$

We use (8.52) and (8.56) in the equation for $\delta(\mathcal{P})$ in (8.22) to deduce the following path-dependent lower bound:

$$\delta(\mathcal{P}) \geq \frac{1}{2} \frac{\left(\frac{2}{k_{\max}(\mathcal{P})+k_{\text{ave}}(\mathcal{P})}\right)^{\mathcal{L}(\mathcal{P})}}{\mathcal{L}(\mathcal{P})(\mathcal{L}(\mathcal{P})-1)}.$$

It then follows that

$$\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P})-1) \geq \frac{1}{4} \frac{\left(\frac{2}{k_{\max}(\mathcal{P})+k_{\text{ave}}(\mathcal{P})}\right)^{2\mathcal{L}(\mathcal{P})}}{\mathcal{L}(\mathcal{P})^2(\mathcal{L}(\mathcal{P})-1)}. \tag{8.57}$$

The left-hand side of (8.57) is monotonically non-increasing in $\mathcal{L}(\mathcal{P})$. From this fact and $\mathcal{L}(\mathcal{P}) \leq n$, we obtain

$$\delta(\mathcal{P})^2(\mathcal{L}(\mathcal{P})-1) \geq \frac{1}{4} \frac{1}{n^3} \left(\frac{2}{k_{\max}+k_{\text{ave}}}\right)^{2n}. \tag{8.58}$$

Because $C_1(\mathbf{x}, S_{u,v}) < 0$, we have that

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S_{u,v}) \geq \min\left(\frac{1}{2}(x_p - x_q)^2, -\frac{2}{k_p + k_q + 2} C_1(\mathbf{x}, \{p, q\})\right) \frac{1}{e(\mathbf{x})} \tag{8.59}$$

as a consequence of (8.15). Applying both (8.23) from Lemma 59 and (8.58) to equation (8.59), we conclude that

$$1 - \mathfrak{r}_{\text{step}}(\mathbf{x}, S_{u,v}) \geq \frac{1}{8} \frac{1}{n^3} \left(\frac{2}{k_{\max}+k_{\text{ave}}}\right)^{2n} \frac{(x_u - x_v)^2}{e(\mathbf{x})}.$$

This completes the proof. □

CHAPTER 9

Conclusions

In this thesis, we introduced a new model of social influence and studied the dynamics of this model. Our model, which we developed in Chapter 3, is meant to describe the belief changes of individuals during group interactions. The belief changes in our model are a result of both explicit social influence from other group members and implicit social influence from the social connections of group members.

In Chapters 4–8, we studied the behavior of our model. The only possible steady state of our model is a consensus in which all individuals have a belief that equals the mean belief over all individuals. However, the time that is required for beliefs to reach consensus depends crucially on which individuals interact. In Chapter 6, we assumed that a group of interacting individuals consists of individuals chosen uniformly at random from the set of all individuals. We fixed the size of a group, and we considered how the convergence time to consensus depends on the group size. In Chapter 7, we introduced strategies for identifying a group of individuals that is able to cause a significant reduction in the total amount of disagreement in a network by coming to a agreement. We studied the convergence time to consensus of our model when we use our strategies to select groups of individuals to interact. In Chapter 8, we required that the groups of individuals that interact be dyads. Under this constraint, we examined the convergence time to consensus when we select groups of individuals to interact uniformly-at-random from the dyads of a graph. In each of Chapters 6, 7, and 8, we derived upper bounds for the convergence time to consensus of our model. We also performed simulations and calculations to examine the convergence time of our model for specific graphs.

We found evidence in Chapters 6, 7, and 8 that the convergence time of our model can be at least quadratic in the number of individuals. Because a quadratic convergence time can be extremely large, this result emphasizes the importance of checking convergence times when considering a steady state as indicative of finite-time behavior. Additionally, if large convergence times are possible, then it is important in situations in which one desires fast convergence to consensus to have strategies to accelerate convergence. Our investigation of strategies in Chapter 7 provides evidence that a strategy can meaningfully decrease the convergence time to consensus. In Chapter 8, we observed that by adding a mechanism in our model to incoporate indirect social influence we caused the convergence time to consensus to increase by a factor that is at least quadratic in the number of individuals. This large increase in convergence time suggests that it may be important to incorporate the effects of both indirect and direct social influence into social-influence models.

In our research, we were guided by two goals: (1) to obtain a simple model of social influence that is inspired from psychological considerations and that is amenable to mathematical analysis and (2) to analyze convergence times to examine sociologically-relevant properties of our model. With our introduction of our model in Chapter 3 and our convergence-time analyses in Chapters 4–8, we hope to have contributed to these two goals. We believe that further research in social-influence modeling that aligns with these two goals will uncover fascinating mathematical properties and improve scientific understanding of social influence.

[AKP18] Rediet Abebe, Jon Kleinberg, David Parkes, and Charalampos E Tsourakakis. "Opinion dynamics with varying susceptibility to persuasion." In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1089–1098, New York, NY, USA, 2018. Association for Computing Machinery.

[AO11] Daron Acemoglu and Asuman Ozdaglar. "Opinion Dynamics and Learning in Social Networks." *Dynamic Games and Applications*, **1**:3–49, 2011.

[AS08] Alan I Abramowitz and Kyle L Saunders. "Is polarization a myth?" *The Journal of Politics*, **70**(2):542–555, 2008.

[Asc51] Solomon E Asch. "Effects of group pressure upon the modification and distortion of judgments." In H Guetzknow, editor, *Groups, Leadership and Men; Research in Human Relations*, pp. 177–190, Pittsburgh, PA, USA, 1951. Carnegie Mellon University Press.

[Asc55] Solomon E Asch. "Opinions and social pressure." *Scientific American*, **193**(5):31–35, 1955.

[Asc56] Solomon E Asch. "Studies of independence and conformity: I. A minority of one against a unanimous majority." *Psychological Monographs: General and Applied*, **70**(9):1–70, 1956.

[Axe97] Robert Axelrod. "The dissemination of culture: A model with local convergence and global polarization." *Journal of Conflict Resolution*, **41**(2):203–226, 1997.

[Bar59] Dean C Barnlund. "A comparative study of individual, majority, and group judgment." *The Journal of Abnormal and Social Psychology*, **58**(1):55–60, 1959.

[BBC92] Ann Bettencourt, Marilynn B Brewer, Marian Rogers Croak, and Norman Miller. "Cooperation and the reduction of intergroup bias: The role of reward structure and social orientation." *Journal of Experimental Social Psychology*, **28**(4):301–319, 1992.

[BBC14] Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. "Maximizing social influence in nearly optimal time." In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 946–957, Philadelphia, PA, USA, 2014. Society for Industrial and Applied Mathematics.

[BCB94] Fletcher A Blanchard, Christian S Crandall, John C Brigham, and Leigh Ann Vaughn. "Condemning and condoning racism: A social context approach to interracial settings." *Journal of Applied Psychology*, **79**(6):993–997, 1994.

[BCI20] Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. "Networks beyond pairwise interactions: Structure and dynamics." *arXiv:2006.01764*, 2020.

[BCN20] Luca Becchetti, Andrea Clementi, and Emanuele Natale. "Consensus dynamics: An overview." *ACM SIGACT News*, **51**(1):58–104, 2020.

[BGP06] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. "Randomized gossip algorithms." *IEEE Transactions on Information Theory*, **52**(6):2508–2530, 2006.

[Bou84] Pierre Bourdieu. *Distinction: A Social Critique of the Judgement of Taste*. Harvard University Press, Cambridge, MA, USA, 1984.

[Bov48] Everett W Bovard Jr. "Social norms and the individual." *The Journal of Abnormal and Social Psychology*, **43**(1):62–69, 1948.

[Bre84] Marilynn B Brewer. "Beyond the contact hypothesis: Theoretical perspectives on desegregation." *Groups in Contact: The Psychology of Desegregation*, pp. 281–302, 1984.

[BS96] Rod Bond and Peter B Smith. "Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task." *Psychological Bulletin*, **119**(1):111–137, 1996.

[Bur90] John W Burton. *Conflict: Resolution and Prevention*. St. Martin's Press, New York, NY, USA, 1990.

[Bur01] J M Burger. "Social Influence, Psychology of." In J Wright, editor, *International Encyclopedia of the Social & Behavioral Sciences*, pp. 14320–14325. Pergamon Press, Oxford, UK, 2001.

[BW17] Andrei Boutyline and Robb Willer. "The social structure of political echo chambers: Variation in ideological homophily in online networks." *Political Psychology*, **38**(3):551–569, 2017.

[CBB10] Ciro Cattuto, Wouter Van den Broeck, Alain Barrat, Vittoria Colizza, Jean-François Pinton, and Alessandro Vespignani. "Dynamics of person-to-person interactions from distributed RFID sensor networks." *PLOS ONE*, **5**(7):e11596, 2010.

[CEE92] Christer Carlsson, Dieter Ehrenberg, Patrik Eklund, Mario Fedrizzi, Patrik Gustafsson, Paul Lindholm, Galina Merkuryeva, Tony Riissanen, and Aldo GS Ventre. "Consensus in distributed soft environments." *European Journal of Operational Research*, **61**(1–2):165–185, 1992.

[CM07] Damon Centola and Michael Macy. "Complex contagions and the weakness of long ties." *American Journal of Sociology*, **113**(3):702–734, 2007.

[CS73] Peter Clifford and Aidan Sudbury. "A model for spatial conflict." *Biometrika*, **60**(3):581–588, 1973.

[CWH06] Richard J Crisp, Judi Walsh, and Miles Hewstone. "Crossed categorization in common ingroup contexts." *Personality and Social Psychology Bulletin*, **32**(9):1204–1218, 2006.

[DDM17] Yucheng Dong, Zhaogang Ding, Luis Martínez, and Francisco Herrera. "Managing consensus based on leadership in opinion dynamics." *Information Sciences*, **397**:187–205, 2017.

[De 07] Nair Maria Maia De Abreu. "Old and new results on algebraic connectivity of graphs." *Linear Algebra and Its Applications*, **423**(1):53–73, 2007.

[DeG74] Morris H DeGroot. "Reaching a consensus." *Journal of the American Statistical Association*, **69**(345):118–121, 1974.

[DG55] Morton Deutsch and Harold B Gerard. "A study of normative and informational social influences upon individual judgment." *The Journal of Abnormal and Social Psychology*, **51**(3):629, 1955.

[DG87] Gerardine Desanctis and R Brent Gallupe. "A foundation for the study of group decision support systems." *Management science*, **33**(5):589–609, 1987.

[DLR72] Norman C Dalkey, Ralph Lewis, and Daniel L Rourke. *Studies in the Quality of Life: Delphi and Decision-making.* Lexington Books, Landham, MD, USA, 1972.

[DM94] Jeannine Drew and the 3M Meeting Management Team. *Mastering Meetings: Discovering the Hidden Potential of Effective Business Meetings.* McGraw-Hill, New York, NY, USA, 1994.

[DNA00] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. "Mixing beliefs among interacting agents." *Advances in Complex Systems*, **03**(01n04):87–98, 2000.

[DPW20] Ru-Xi Ding, Iván Palomares, Xueqing Wang, Guo-Rui Yang, Bingsheng Liu, Yucheng Dong, Enrique Herrera-Viedma, and Francisco Herrera. "Large-Scale decision-making: Characterization, taxonomy, challenges and future directions from an Artificial Intelligence and applications perspective." *Information Fusion*, **59**:84–102, 2020.

[DS82] Michael Doyle and David Straus. *How to Make Meetings Work: The New Interaction Method.* Jove Books, New York, NY, USA, 1982.

[Dur82] Emile Durkheim. *Rules of Sociological Method.* Simon and Schuster, New York, NY, USA, 1982.

[DZK18] Yucheng Dong, Min Zhan, Gang Kou, Zhaogang Ding, and Haiming Liang. "A survey on the fusion process in opinion dynamics." *Information Fusion*, **43**:57–65, 2018.

[DZZ18] Yucheng Dong, Quanbo Zha, Hengjie Zhang, Gang Kou, Hamido Fujita, Francisco Chiclana, and Enrique Herrera-Viedma. "Consensus reaching in social network group decision making: Research paradigms and challenges." *Knowledge-Based Systems*, **162**:3–13, 2018.

[Einnd] Gil Einstein. "Sherif's Study of Suggestibility.", n.d. Retrieved 06-01-2020 from http://eweb.furman.edu/~einstein/general/social/suggest.htm.

[Eps99] Joshua M Epstein. "Agent-based computational models and generative social science." *Complexity*, **4**(5):41–60, 1999.

[FA08] Morris P Fiorina and Samuel J Abrams. "Political polarization in the American public." *Annual Review of Political Science*, **11**:563–588, 2008.

[FAP08] Morris P Fiorina, Samuel A Abrams, and Jeremy C Pope. "Polarization in the American public: Misconceptions and misreadings." *The Journal of Politics*, **70**(2):556–560, 2008.

[Fes50] Leon Festinger. "Informal social communication." *Psychological Review*, **57**(5):271, 1950.

[FGR16] Seth Flaxman, Sharad Goel, and Justin M Rao. "Filter bubbles, echo chambers, and online news consumption." *Public Opinion Quarterly*, **80**(S1):298–320, 2016.

[FH16] Santo Fortunato and Darko Hric. "Community detection in networks: A user guide." *Physics Reports*, **659**:1–44, 2016.

[FI99] Juan Manuel Falomir and Federica Invernizzi. "The role of social influence and smoker identity in resistance to smoking cessation." *Swiss Journal of Psychology*, **58**(2):73–84, 1999.

[FJ90] Noah E Friedkin and Eugene C Johnsen. "Social influence and opinions." *Journal of Mathematical Sociology*, **15**(3–4):193–206, 1990.

[FJ99] Noah E Friedkin and Eugene C Johnsen. "Social influence networks and opinion change." *Advances in Group Processes*, **16**(1):1–29, 1999.

[FJ11] Noah E Friedkin and Eugene C Johnsen. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics.* Cambridge University Press, Cambridge, UK, 2011.

[FMF17] Andreas Flache, Michael Mäs, Thomas Feliciani, Edmund Chattoe-Brown, Guillaume Deffuant, Sylvie Huet, and Jan Lorenz. "Models of social influence: Towards the next frontiers." *Journal of Artificial Societies and Social Simulation*, **20**(4):2, 2017.

[Fre56] John R P French Jr. "A formal theory of social power." *Psychological Review*, **63**(3):181–194, 1956.

[Fri15] Noah E Friedkin. "The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem." *IEEE Control Systems Magazine*, **35**(3):40–51, 2015.

[FW16] Joseph P Forgas and Kipling D Williams. *Social Influence: Direct and Indirect Processes*. Routledge, New York, NY, USA, 2016.

[Gar09] R Kelly Garrett. "Echo chambers online?: Politically motivated selective exposure among Internet news users." *Journal of Computer-Mediated Communication*, **14**(2):265–285, 2009.

[GBC18] Douglas Guilbeault, Joshua Becker, and Damon Centola. "Complex contagions: A decade in review." In S Lehmann and YY Ahn, editors, *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, pp. 3–25. Springer International Publishing, Cham, Switzerland, 2018.

[GBK09] Eric Gilbert, Tony Bergstrom, and Karrie Karahalios. "Blogs are echo chambers: Blogs are echo chambers." In *2009 42nd Hawaii International Conference on System Sciences*, pp. 1–10. IEEE, 2009.

[GDG17] 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*, pp. 81–90. Association for Computing Machinery, 2017.

[Gra73] Mark S Granovetter. "The Strength of Weak Ties." *American Journal of Sociology*, **78**(6):1360–1380, 1973.

[GS14] Javad Ghaderi and Rayadurgam Srikant. "Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate." *Automatica*, **50**(12):3209–3215, 2014.

[Har59] Frank Harary. "A criterion for unanimity in French's theory of social power." In D Cartwright, editor, *Studies in Social Power*, pp. 168–182. University of Michigan, Ann Arbor, MI, USA, 1959.

[HCC17] Enrique Herrera-Viedma, Francisco Javier Cabrerizo, Francisco Chiclana, Jian Wu, Manuel Jesus Cobo, and Samuylov Konstantin. "Consensus in group decision making and social networks." *Studies in Informatics and Control*, **26**(3):259–268, 2017.

[HCK14] Enrique Herrera-Viedma, Francisco Javier Cabrerizo, Janusz Kacprzyk, and Witold Pedrycz. "A review of soft consensus models in a fuzzy environment." *Information Fusion*, **17**:4–13, 2014.

[HKK14] Rainer Hegselmann, Stefan König, Sascha Kurz, Christoph Niemann, and Jörg Rambau. "Optimal opinion control: The campaign problem." *Journal of Artificial Societies and Social Simulation*, **18**(3):18, 2014.

[HL75] Richard A Holley and Thomas M Liggett. "Ergodic theorems for weakly interacting infinite systems and the voter model." *The Annals of Probability*, **3**(4):643–663, 1975.

[HL78] Ted L Huston and George Levinger. "Interpersonal attraction and relationships." *Annual Review of Psychology*, **29**(1):115–156, 1978.

[HL06] Dorit S Hochbaum and Asaf Levin. "Methodologies and algorithms for group-rankings decision." *Management Science*, **52**(9):1394–1408, 2006.

[HOM96] S Alexander Haslam, Penelope J Oakes, Craig McGarty, John C Turner, Katherine J Reynolds, and Rachael A Eggins. "Stereotyping and social influence: The mediation of stereotype applicability and sharedness by the views of in-group and out-group members." *British Journal of Social Psychology*, **35**(3):369–397, 1996.

[KDG03] David Kempe, Alin Dobra, and Johannes Gehrke. "Gossip-based computation of aggregate information." In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pp. 482–491. Institute of Electrical and Electronics Engineers, 2003.

[Kel10] Herbert C Kelman. "Conflict resolution and reconciliation: A social-psychological perspective on ending violent conflict between identity groups." *Landscapes of Violence*, **1**(1):5, 2010.

[KF86] J Kacprzyk and M Fedrizzi. "'Soft' consensus measures for monitoring real consensus reaching processes under fuzzy preferences." *Control and Cybernetics*, **15**(3–4):309–323, 1986.

[KF88] Janusz Kacprzyk and Mario Fedrizzi. "A 'soft' measure of consensus in the setting of partial (fuzzy) preferences." *European Journal of Operational Research*, **34**(3):316–325, 1988.

[KF89] Janusz Kacprzyk and Mario Fedrizzi. “A ‘human-consistent’ degree of consensus based on fuzzy logic with linguistic quantifiers.” *Mathematical Social Sciences*, **18**(3):275–290, 1989.

[KKT03] David Kempe, Jon Kleinberg, and Éva Tardos. “Maximizing the spread of influence through a social network.” In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, New York, NY, USA, 2003. Association for Computing Machinery.

[KML16] Athanasios Kolios, Varvara Mytilinou, Estivaliz Lozano-Minguez, and Konstantinos Salonitis. “A comparative study of multiple-criteria decision-making methods under stochastic inputs.” *Energies*, **9**(7):566, 2016.

[Knu97] Donald Ervin Knuth. *The Art of Computer Programming*, volume 3. Pearson Education, Santa Monica, CA, USA, 1997.

[Kra00] Ulrich Krause. “A discrete nonlinear and non-autonomous model of consensus formation.” In *Communications in Difference Equations: Proceedings of the Fourth International Conference on Difference Conference on Difference Equations*, pp. 227–236, Boca Raton, FL, USA, 2000. CRC Press.

[LEV81] Nan Lin, Walter M Ensel, and John C Vaughn. “Social resources and strength of ties: Structural factors in occupational status attainment.” *American Sociological Review*, pp. 393–405, 1981.

[LFW18] Yuchen Li, Ju Fan, Yanhao Wang, and Kian-Lee Tan. “Influence maximization on social graphs: A survey.” *IEEE Transactions on Knowledge and Data Engineering*, **30**(10):1852–1872, 2018.

[LL85] Barry Loewer and Robert Laddaga. “Destroying the consensus.” *Synthese*, **62**:79–95, 1985.

[LRS11] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. “How social influence can undermine the wisdom of crowd effect.” *Proceedings of the National Academy of Sciences*, **108**(22):9020–9025, 2011.

[LT75] Harold A Linstone and Murray Turoff. *The Delphi Method: Techniques and Applications*. Addison-Wesley, Reading, MA, USA, 1975.

[Lyn96] Nancy A Lynch. *Distributed Algorithms*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 1996.

[MHB81] John B McConahay, Betty B Hardee, and Valerie Batts. “Has racism declined in America? It depends on who is asking and what is asked.” *Journal of Conflict Resolution*, **25**(4):563–579, 1981.

[Mil63] Stanley Milgram. "Behavioral study of obedience." *The Journal of Abnormal and Social Psychology*, **67**(4):371–378, 1963.

[MLN69] Serge Moscovici, Elisabeth Lage, and Martine Naffrechoux. "Influence of a consistent minority on the responses of a majority in a color perception task." *Sociometry*, pp. 365–380, 1969.

[MMG07] Alan Mislove, Massimiliano Marcon, Krishna P Gummadi, Peter Druschel, and Bobby Bhattacharjee. "Measurement and analysis of online social networks." In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pp. 29–42, New York, NY, USA, 2007. Association for Computing Machinery.

[MSC01] 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.

[MTS13] Pavlin Mavrodiev, Claudio J Tessone, and Frank Schweitzer. "Quantifying the effects of social influence." *Scientific Reports*, **3**:1360, 2013.

[MVP18] X Flora Meng, Robert A Van Gorder, and Mason A Porter. "Opinion formation and distribution in a bounded-confidence model on various networks." *Physical Review E*, **97**(2):022312, 2018.

[Ng16] Sik Hung Ng. "Influencing through the power of language." In J P Forgas and K D Williams, editors, *Social Influence: Direct and Indirect Processes*, pp. 185–197, New York, NY, USA, 2016. Routledge.

[OT11] Alex Olshevsky and John N Tsitsiklis. "Cenvergence speed in distributed consensus and averaging." *SIAM Review*, **53**(4):747–772, 2011.

[Oys] Fred the Oyster. "Asch experiment." Wikipedia.org. Retrieved 05-28-2020 from https://commons.wikimedia.org/wiki/File:Asch_experiment.svg.

[Pet98] Thomas F Pettigrew. "Intergroup contact theory." *Annual Review of Psychology*, **49**(1):65–85, 1998.

[PFT19a] Rohit Parasnis, Massimo Franceschetti, and Behrouz Touri. "On graphs with bounded and unbounded convergence times in social Hegselmann-Krause dynamics." In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 6431–6436, New York, NY, USA, 2019. Institute of Electrical and Electronics Engineers.

[PFT19b] Rohit Parasnis, Massimo Franceschetti, and Behrouz Touri. "On the convergence properties of social Hegselmann–Krause dynamics." *arXiv:1909.03485*, 2019.

[PG16] Mason A Porter and James P Gleeson. *Dynamical Systems on Networks: A Tutorial*. Number 4 in Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Springer International Publishing, Cham, Switzerland, 2016.

[PMC16] Luis G Pérez, Francisco Mata, Francisco Chiclana, Gang Kou, and Enrique Herrera-Viedma. "Modelling influence in group decision making." *Soft Computing*, **20**(4):1653–1665, 2016.

[POM09] Mason A Porter, Jukka-Pekka Onnela, and Peter J Mucha. "Communities in networks." *Notices of the AMS*, **56**(9):1082–1097, 1164–1166, 2009.

[PT17] Anton V Proskurnikov and Roberto Tempo. "A tutorial on modeling and analysis of dynamic social networks. Part I." *Annual Reviews in Control*, **43**:65–79, 2017.

[PT18] Anton V Proskurnikov and Roberto Tempo. "A tutorial on modeling and analysis of dynamic social networks. Part II." *Annual Reviews in Control*, **45**:166–190, 2018.

[PTC17] Anton V Proskurnikov, Roberto Tempo, Ming Cao, and Noah E Friedkin. "Opinion evolution in time-varying social influence networks with prejudiced agents." *IFAC-PapersOnLine*, **50**(1):11896–11901, 2017.

[QSS16] Walter Quattrociocchi, Antonio Scala, and Cass R Sunstein. "Echo chambers on Facebook." *SSRN:2795110*, 2016. Available at http://dx.doi.org/10.2139/ssrn.2795110.

[RBH54] John H Rohrer, Seymour H Baron, EL Hoffman, and DV Swander. "The stability of autokinetic judgments." *The Journal of Abnormal and Social Psychology*, **49**(4, Pt.1):595–597, 1954.

[Red19] Sidney Redner. "Reality-inspired voter models: A mini-review." *Comptes Rendus Physique*, **20**(4):275–292, 2019.

[RN01] Nicholas C Romano and Jay F Nunamaker. "Meeting analysis: Findings from research and practice." In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, p. 13, New York, NY, USA, 2001. Institute of Electrical and Electronics Engineers.

[Sco88] John Scott. "Social network analysis." *Sociology*, **22**(1):109–127, 1988.

[SGA13] Hammad Sheikh, Jeremy Ginges, and Scott Atran. "Sacred values in the Israeli-Palestinian conflict: Resistance to social influence, temporal discounting, and exit strategies." *Annals of the New York Academy of Sciences*, **1299**:11–24, 2013.

[She36] Muzafer Sherif. *The Psychology of Social Norms*. HarperCollins Publishers, New York, NY, USA, 1936.

[SSJ16] Charles Stangor, Gretchen B Sechrist, and John T Jost. "Social influence and intergroup beliefs: The role of perceived social consensus." In J P Forgas and K D Williams, editors, *Social Influence: Direct and Indirect Processes*, pp. 235–252, New York, NY, USA, 2016. Routledge.

[TFM16] Károly Takács, Andreas Flache, and Michael Mäs. "Discrepancy and disliking do not induce negative opinion shifts." *PLOS ONE*, **11**(6):e0157948, 2016.

[Tho38] Robert L Thorndike. "The effect of discussion upon the correctness of group decisions, when the factor of majority influence is allowed for." *The Journal of Social Psychology*, **9**(3):343–362, 1938.

[Wat02] Duncan J Watts. "A simple model of global cascades on random networks." *Proceedings of the National Academy of Sciences of the United States of America*, **99**(9):5766–5771, 2002.

[WCF17] Jian Wu, Francisco Chiclana, Hamido Fujita, and Enrique Herrera-Viedma. "A visual interaction consensus model for social network group decision making with trust propagation." *Knowledge-Based Systems*, **122**:39–50, 2017.

[WCG91] Stephen Worchel, Joel Cooper, and George R Goethals. *Understanding Social Psychology*. Thomson Brooks/Cole Publishing Co, Independence, KY, USA, 5th edition, 1991.

[WDL18] Xinjue Wang, Ke Deng, Jianxin Li, Jeffery Xu Yu, Christian S Jensen, and Xiaochun Yang. "Targeted influence minimization in social networks." In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 689–700, Cham, Switzerland, 2018. Springer International Publishing.

[WLO94] Wendy Wood, Sharon Lundgren, Judith A Ouellette, Shelly Busceme, and Tamela Blackstone. "Minority influence: A meta-analytic review of social influence processes." *Psychological Bulletin*, **115**(3):323–245, 1994.

[XB04] Lin Xiao and Stephen Boyd. "Fast linear iterations for distributed averaging." *Systems & Control Letters*, **53**(1):65–78, 2004.