Network Based Models of Opinion Formation: Consensus and Beyond

by

Dylan Weber

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2021 by the
Graduate Supervisory Committee:

Sebastien Motsch, Chair
Dieter Armbruster
John Fricks
Nicolas Lanchier
Rodrigo Platte

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

Understanding the evolution of opinions is a delicate task as the dynamics of how one changes their opinion based on their interactions with others are unclear. The rise of social media has drastically changed the conception of opinion formation as there are now a myriad of smaller actors that steer opinions one way or the other. Therefore, opinion formation can be argued to be an example of self-organized dynamics where large scale behaviors emerge without a central authority similar to a flock of birds or a shoal of fish. In this dissertation several contributions to this body of work are presented.

First, two linear models of opinion formation are studied - one stochastic in nature and one deterministic. Both models are defined in terms of an underlying graph; the effects of the structure of the graph on the long time behavior of the models are studied in all cases of graph topology. Special attention is given to the emergence of a *consensus* among the agents and a condition on the graph that is necessary and sufficient for convergence to a consensus is provided.

Next, a variant of the well known bounded confidence dynamics is proposed with the goal of inducing unconditional convergence to a consensus. The defining feature of these dynamics which are dubbed No one left behind dynamics, is the introduction of a local control on the agents which preserves the connectivity of the interaction network. A rigorous demonstration that these dynamics result in unconditional convergence to a consensus is provided and a relaxed version of the control is presented that maintains many qualitative features of the bounded confidence dynamics.

Finally a model aimed at a better understanding of the polarization of opinions in addition to consensus is introduced. Most models of opinion formation only include attractive forces among the agents and only capture polarization through an echo-chamber effect. The addition of repulsive forces is mathematically challenging as repulsion can cause divergence

of the model. A model which dubbed the *Döppelganger model* is introduced which includes repulsive forces, show that this model remains bounded and characterize its convergence.

DEDICATION

For Da.

ACKNOWLEDGMENTS

Thank you to my family and my (soon to be) wife for their unwavering love and support. This work would not exist without the support and guidance of my adviser, Dr. Sebastien Motsch. I'd also like to thank Dr. Simone Göttlich and her lab at the University of Mannheim for the hospitality and the many hours of useful and interesting discussion during my time as a visiting researcher in the Spring of 2019. Finally, to my committee and all others who are too numerous to name (you know who you are) who shared their thoughts, ideas, and had some fun with me - thank you.

TABLE OF CONTENTS

Page

CHAPTER Page

Chapter 1

INTRODUCTION

How are opinions formed? What mechanisms cause individuals to think similarly or in disparate ways? Understanding the evolution of opinions is a delicate task as the dynamics of how one changes their opinion based on their interactions with others are unclear. Moreover, the rise of social media has drastically changed our conception of opinion formation. Rather than a few central outlets (e.g. TV channels, newspapers) acting as a common and external frame of reference, there are now a myriad of smaller actors that steer opinions one way or the other. Therefore, opinion formation can be argued to be an example of self-organized dynamics where large scale behaviors emerge without a central authority similar to a flock of birds or a shoal of fish [22, 79, 29, 9, 35, 118].

The idea that one's opinion is the result of a summation of influences from other individuals dates back to the beginnings of inquiry into social dynamics. However, it wasn't until the 1950s that a formulation of this idea precise enough to be amenable to mathematical analysis appeared in [48]. Here, an individual's opinion is modeled as a continuous value on a one dimensional spectrum that changes based on that individual's connections with other individuals in a social network modeled as a directed graph. The importance that an individual affords another's opinion is modeled by positive weights on the edges of the directed graph. A given individual feels a "pull" exerted on their opinion value by those to whom it is connected which is proportional to the weight of the connection. Ultimately, the assumption that an individual changes their opinion due to a summation of influences is captured quite literally in these early models as an individual's new opinion value was modeled as a sum of the opinion value's to those that individual is connected in the directed

graph weighted by values defined on those connections. Several different models of this class were studied in this period with minor differences and *consensus*, when all individual opinion values converge on a single value, emerged as a ubiquitous feature [39, 56].

Despite the success of this framework at codifying assumptions about the dynamics of social interaction, work on models of this type was sparse in the following 40 years. From a mathematical point of view, these early models are similar in the sense that they are all linear and all define only positive weights between individuals. From a modeling point of view the tendency of these models to evolve to consensus is not ideal and it was recognized [2] that assumptions that produce more varied outcomes in the same framework would require non linear models. However, the analysis of non linear models is notoriously complex and it was not until the late 1990s, after the advent of computing, that a detailed study of a nonlinear model for modeling opinion dynamics was introduced by Hegselmann and Krause in [57]. Here, the non linearity introduced is meant to formalize the assumption that two individuals will only interact with each other in the first place if they already agree in some sense, formally if their opinion values are within a prescribed range; the confidence that an individual has in another is bounded. However if they do interact they do in a manner exactly analogous to the earlier class of models. This assumption causes the formation of clusters of individuals to be a standard phenomena, consensus is rare and much work following was done analyzing and characterizing the observed clustering behavior. Similar to the earlier models the structure of the interaction network is central, however in this case that structure depends on the distribution of opinions in the network. In particular, the observed clustering of opinions is equivalent to a fragmenting of the social network.

Following the introduction of Hegselmann-Krause model, a larger body of work analyzing models in this framework began to emerge [95, 121, 80, 29, 57, 109, 14, 13]. The rise of social networks as the ubiquitous forum for the exchange of information in our society

and the meta data associated with these networks further motivated the continued study of opinion formation from an agent-based modeling point of view. The defining feature of the agent-based approach is the study of how locally defined interaction rules affect globally observed behavior among the agents. Models of opinion formation often have the feature that agents can only interact if they are connected in an underlying network structure. Therefore, a hallmark of the study of these models is examining how the interplay between the topology of the underlying network and the interaction rules affect the distribution of opinions among the agents [85, 106, 119]. Of particular interest is how these factors can lead to the emergence of a consensus among the agents. Models in this context have the generic assumption that the opinion of a given agent is continuously influenced by those to whom it is connected in the network according to relatively simple interaction rules which are globally defined. Often these rules carry an assumption of *local consensus*; if agents interact only with each other then they should agree in some sense. This assumption can also be interpreted as saying that there are only attractive forces present among the agents. One might assume that the attractive nature of the interactions causes the emergence of consensus to be a ubiquitous feature of this class of models, however this is not the case. The manner in which agents are connected in the underlying network has a large effect on the distribution of opinions observed among the agents [85, 119]. The interplay between the network structure and the interaction rules can often cause the analysis of these models to be very involved; a popular strategy is to use simplifying assumptions on the network structure such as symmetry of connections or static connections that do not change throughout the evolution of the model [106, 109, 123, 95]. Given an interaction rule, the second strategy could be viewed as studying a "linearization" of a model with the same interaction rule but dynamic connections. A main takeaway from the study of these models is that a necessary condition for the emergence of a consensus is the persistence of

a suitable degree of connectivity in the network throughout the evolution of the dynamics. This allows for "heterophilic" interactions; agents with disparate opinions interact and due to the attractive nature of the interaction rules eventually agree [85].

In this dissertation several contributions to this body of work are presented and plans for future research and preliminary results are presented. This work is unified in the viewpoint that interplay between the underlying interaction network and the observed opinions is the critical consideration when analyzing models of this type.

We first discuss two linear models of opinion formation; one deterministic in nature and one stochastic. In both cases we are interested in how the structure of the graph affects the long-time behavior of the model. We are particularly interested in conditions on the graph that cause the opinion of every agent to converge to a *consensus* - i.e. every agent converges on the same opinion in the long-time limit. Moreover we analyze if the addition of stochasticity helps or hinders the emergence of a consensus; if consensus occurs, we are interested in how the connectivity of the graph and stochasticity affect the speed of convergence. We examine and characterize both models' behavior in all possible graph topologies: i) undirected and connected (referred to as a *simple* graph), ii) directed and strongly connected, and iii) directed and weakly connected or disconnected.

In the deterministic case we find that the model converges regardless of graph structure. In the case that the underlying graph is simple, we find that a consensus is reached that is the *initial average* of opinions. Similarly, if the graph is strongly connected, but not necessarily undirected, we again find that a consensus emerges, however it is not in general the initial average. The most challenging case occurs when the underlying network is only weakly connected. In general, a consensus does not form. However, we are able to find necessary and sufficient conditions on the structure of the graph for convergence to

a consensus regardless of the initial condition: the graph needs to have only one *isolated strongly connected component*.

In the stochastic case we similarly find consensus in the case of a strongly connected network using markedly different techniques in the arguments. In contrast to the deterministic model, the stochastic model does not converge in general in the case of a weakly connected graph. Despite this, we find that the condition that the graph have exactly one isolated strongly connected component prevails and is equivalent to convergence to a consensus in the stochastic case as well. We explore this link between the models further as we find that the deterministic dynamics can be recovered in expectation from the stochastic dynamics.

We discuss speed of convergence for both the deterministic and stochastic models and quantify it in terms of the *algebraic connectivity* of the graph, also known as the Fielder number[45, 69, 70], which also happens to be the spectral gap of the operator associated with consensus dynamics. The algebraic connectivity measures, in a sense, how *well connected* the graph is and we find generally that better connection in the graph results in faster convergence of the models. However, we find a contrast between the models as the convergence rate depends on the number of agents in the stochastic case whereas the rate in the deterministic case depends only on the connectivity of the graph. In short, the spectral gap reduces drastically in the stochastic model.

We next study a class of models inspired by the Hegselmann-Krause bounded confidence model [57] in which the connections between agents are dynamic; a connection forms between agents when their opinions are within an interaction range. This dynamic (combined with an attractive interaction rule) causes the formation of "clusters" of opinions in the long time limit to be a generic behavior; consensus is rare. For this reason much of the study of this class of models has focused on characterizing the clustering behavior [14, 36, 63, 81, 15, 71]. We take a different viewpoint and instead investigate controls on collections of

agents [24, 99] otherwise evolving according to bounded confidence dynamics that result in consensus. The interaction range in bounded confidence dynamics causes interactions between agents to be "homophilic"; agents only interact with agents who are sufficiently "similar". This tendency causes the interaction network of a collection of agents to quickly become disconnected and prevents a consensus from occurring despite the fact that agents who do interact attract each other. Therefore, the controls that we impose on agents are generally motivated by maintaining connectivity among the agents.

We investigate two different ways of augmenting the bounded confidence dynamics with the goal of achieving consensus. The first strategy which we dub the *no one left behind dynamics* imposes the rule that once agents become connected they remain connected. We prove rigorously that under this augmentation of the bounded confidence dynamics, connectivity of the initial interaction network is sufficient for a consensus to emerge for agents who's opinion can be of arbitrary dimension. If we restrict agent opinions to being one dimensional we can quantify how fast a consensus is reached as we derive explicit convergence rates. Here, an interesting phenomenon is observed as we find that the convergence occurs in two stages. Before all agents are within the interaction range provided by the bounded confidence dynamics the convergence is linear, afterwards the convergence to consensus spontaneously becomes exponential. The preservation of connections among pairs of agents is sufficient to preserve the connectivity of the network however it isn't necessary. If the existence of *paths* between agents is maintained then the connectivity of the interaction network is maintained as well. The second strategy which we dub the *relaxed no one left behind dynamics* takes advantage of this observation and demands that agents who are connected by a *path* in the interaction network remain connected by a path. We find numerical evidence that this less restrictive control is sufficient for consensus as well. We also demonstrate numerically that this strategy is, in a sense, an interpolation between the

bounded confidence dynamics and the no one left behind dynamics - most agents evolve according to the bounded confidence dynamics and a high degree of clustering initially occurs. However several "bridging" agents alter their trajectories in order to maintain connectivity of the interaction network and ensure convergence to a consensus.

Finally, in a slight departure from the previous two models, we present an investigation of a model hierarchy intended to model the food seeking behavior of an organism known as a *slimemold*. While different in inspiration we believe that the microscopic model in particular serves as an interesting example of how agent based models of opinion formation could be extended (for example to include the possibility of repulsive forces between the agents). Additionally, the interest in modeling slimemold behavior stems mainly from its noted ability to act itself as a model for a host of collective phenomena; it would be interesting to examine opinion formation in this context. Finally, the framework of deriving a density based model from the agent based model is an approach that would be interesting to apply more widely to agent based models of opinion formation. *P. polycephalum* , or a slime mold, is an amoeboid organism that is notable for its ability to perform complex tasks despite its relatively simple biological structure and lack of a brain. Surprisingly, slime molds are known to be able to solve mazes, construct robust networks, and solve shortest path and spanning tree problems [18, 92, 3, 5, 72, 11, 4, 89, 94]. The simplicity of the slime mold organism and the complexity of its emergent behaviors have inspired research into *P. polycephalum* as a model for collective behavior [64, 66, 113, 65]. A hallmark of research in collective dynamics is complex global structure emerging as a result of relatively simple local interactions between agents. Indeed, cellular automata models inspired by *P. polycephalum* food seeking behavior have been found to be good models of other collective phenomena such as the formation of transportation networks in a country or the "cosmic web" of stellar material between galaxies [21, 6].

When seeking food, *P. polycephalum* moves to concentrate most of its mass on food sources - however it maintains a network of mass between food sources. In the absence of food it begins to retract to a central mass. Past research into models of *P. polycephalum* has mainly attempted to model the general food seeking behavior of physarum or its food seeking behavior within a maze. In the first camp most of the models take a microscopic, cellular automata approach - the slime mold is represented by a collection of particles that evolve discretely in time, space, or both. Different interaction rules and sometimes even different "types" of agents are employed to model different slime mold behaviors, however almost all models include a rule that causes slime mold agents to move towards food sources [78, 120, 77, 116, 55]. Most attempts to model the maze solving behavior of the slime mold take a macroscopic approach. Some studies note that the *P. polycephalum* moves and forms networks by moving fluid through tubules within its mass - much like an electrical network [113, 114, 112]. These authors focus on modeling *P. polycephalum* as a network of ordinary differential equations that describe the flow of *P. polycephalum* through network components. Others try to be more specific and attempt to directly model the density of *P. polycephalum* on each maze component. Here the maze is described as a network topology and *P. polycephalum* is characterized with one dimensional partial differential equations for chemotaxis (in particular the Keller-Segel model) defined on the edges of the network with coupling conditions at the junctions of the maze [19, 20]. Both of these macroscopic descriptions attempt to model the path that *P. polycephalum* builds between food sources once they are located. In particular they do not attempt to model the general food seeking behavior of the organism.

In this work we propose a model hierarchy aimed at modeling the food seeking behavior of *P. polycephalum*. The main behaviors we attempt to model are aggregation of mass on food sources while maintaining connected paths between food sources and retraction in

the absence of food sources. We first define an agent based model that is continuous in time and space through a system of stochastic differential equations. The evolution of each agent is governed by three terms - an interaction term with other agents that causes agents to aggregate, a drift term that causes agents to move towards food sources and a noise term intended to model slime mold foraging behavior in the absence of food sources. However any simulation of the microscopic model that includes a suitable number of agents to accurately portray *P. polycephalum* behavior would be computationally intensive. Therefore, we demonstrate under the assumption of propagation of chaos that as the number of agents approaches infinity, the marginal distribution of each agent converges to a deterministic density based model in the form of an aggregation-diffusion equation. This macroscopic model includes terms that correspond to the drift, interaction and noise terms present in the microscopic model. We then embark on an analysis of the macroscopic equation from the point of view of modeling *P. polycephalum* food seeking behavior. We first discuss a framework for choosing an appropriate interaction kernel through an analysis of the steady states of the macroscopic model in the case that there are no food sources. Then, we rigorously show that there is a parameter regime in which the only possible steady state of the macroscopic model is zero and thus this regime is not suitable for modeling *P. polycephalum* food seeking behavior. Finally, we simulate the macroscopic model using the *blob method for aggregation-diffusion equations* presented in [28] and find that it is necessary to scale the three main terms of the model in order to reproduce aggregation on food sources. We then simulate the evolution of the model in different scaling regimes and with different choices of interaction kernel and qualitatively examine which regimes result in the main behaviors we attempt to model. We find that there are several regimes that result in the qualitative desired behaviors.

From a modeling point of view our approach distinguishes itself from the previously described literature in several ways. As previously discussed, past models of *P. polycephalum* behavior mainly took either a discrete agent based or macroscopic viewpoint - our approach unifies these strategies as our macroscopic description is formally derived from the agent based model. Additionally, most models of general *P. polycephalum* food seeking behavior were from the agent based view point; our model hierarchy allows us to present a macroscopic description of this behavior. Finally, the macroscopic description that results is an aggregation diffusion equation with a drift term; while known, to our knowledge such models have not been widely studied. Aggregation diffusion equations without a drift term are well known to exhibit "blow up" below a critical mass - we do not observe such a phenomena in our simulations (although we do not vary the mass) - it could be an interesting area of future research to investigate the effect of the drift term on the occurrence of blow up.

There are many ways in which this work can be extended. In many models proposed to describe the formation of consensus [95, 81, 80, 119, 85, 58, 39, 73], the interaction network has a regularizing effect; it helps neighbors share a similar opinion. The underlying assumption is that individuals can only exert *attractive* forces on each other. In other words there is an assumption of local consensus, if two individuals interact they should agree in some sense. Therefore, many existing models do not include a dynamic of *disagreement*, individuals either interact and agree or do not interact at all. A natural way to model this dynamic is by including negative weights in the interaction graph. However, naively adding negative weights in previously proposed models usually results in a divergence of the dynamics. A careful investigation of how to introduce repulsive forces among the individuals in a way that is natural to the problem is required. We propose to study a model we dub the *Doppelgänger model*. Intuitively, this model introduces a dynamic of disagreement

by stating that when two agents, i and j, disagree the opinion of agent i is attracted to the *opposite* opinion of that of agent j, in lieu of being repulsed by the opinion of agent j. We study these dynamics in two main regimes. In both regimes the dynamics considered are nonlinear, however in the *pseudo-linear* regime the interaction network among agents remains fixed whereas in the fully nonlinear regime the interaction network varies according to an interaction function similar to the formulation found in [58], however in this case the interaction function is allowed to be negative. We demonstrate that in both regimes that the dynamics remain bounded. Additionally, in the pseudo-linear regime we are able to characterize the convergence of the model in terms of the static interaction network.

Another main area of future research is to combine our modeling framework with experimental data. This is critical to the motivation of the modeling pursuits as it might be the case that opinion dynamics as measured by social media data is too noisy of a process to truly be modeled by dynamical systems. Further it remains to be seen if opinion dynamics are truly an example of self organized behavior - it could be that several influencers dictate the majority of the behavior. Experimental data are derived from social media activity such as Twitter for which we have at our disposal a large data-set. Thus, a major challenge is to extract opinions of users along with their interaction networks. The advent of computing represents a major change in that landscape. There are now multiple methods to measure the graph structures and the opinion of users using *sentiment analysis* [51, 75, 82, 87] and *graph clustering* [108, 115, 61, 50, 74, 43, 93, 96, 68, 107]. Sentiment analysis and network scaling have been used in concert in several studies [59, 32, 40]. Usually the aim is to use graph clustering to identify communities within the networks and then leverage sentiment analysis to characterize the opinions of those communities in a static way. However, less work has been done on the dynamics of how those communities were formed and how they may coalesce/fragment in the future. In this area, using sentiment analysis and graph

clustering in concert to examine the change in the structure of the network and opinions of the users over time could provide valuable insights to the interplay between those dynamics and ultimately inform mathematical modeling efforts to further understand the more intricate dynamics at play.

The final proposed research aim is to combine our modeling and measurement frameworks in order to develop a strategy to mitigate polarization if it occurs. This approach is two fold. First, we will recast our theoretical models in the language of an optimization problem in order to develop an approach to identify the nodes in a given network who are most important in the sense of preventing polarization. Specifically we will introduce "moderating nodes" to the models and attempt to minimize the observed variance of opinions by controlling how the moderating node interacts with the network. Once this theoretical framework is developed it can be tested in the data by identifying two polarized communities and using them as the input to the previously discussed control problem. Once susceptible nodes are identified via the models, efforts could be made (for example by a public moderating account) to prevent the fragmentation of the interaction network; for instance by linking those users with content they would not otherwise be served by the social network algorithms.

Chapter 2

DETERMINISTIC VERSUS STOCHASTIC CONSENSUS DYNAMICS ON GRAPHS

The content in this chapter closely follows the content in [119]. 1In this chapter we study two agent based models of opinion formation - one stochastic in nature and one deterministic. Both models are defined in terms of an underlying graph; we study how the structure of the graph affects the long time behavior of the models in all possible cases of graph topology. We are especially interested in the emergence of a *consensus* among the agents and provide a condition on the graph that is necessary and sufficient for convergence to a consensus in both models. This investigation reveals several contrasts between the models - notably the convergence rates - which are explored through analytical arguments and several numerical experiments.

2.1 Consensus dynamics: a deterministic model

2.1.1 Model definition and preliminary properties

We first study the following deterministic consensus model. Many of the results in this section concerning strongly connected networks are well known and are included for context - see for example [33].

Definition 1 *Given a collection of N agents, let $s_i \in \mathbb{R}$ represent the opinion of the ith agent. The **consensus model** (CM) is defined by the dynamics:*

$$s_i' = \sum_{j=1}^{N} a_{ij}(s_j - s_i) \quad \textit{where} \quad a_{ij} \geq 0. \tag{2.1}$$

In vector notation, introducing $\mathbf{s} = (s_1, ..., s_N)^T$, *the dynamics can be written as:*

$$\mathbf{s}' = -L\mathbf{s}, \tag{2.2}$$

where

$$L = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & -a_{ji} & \\ & -a_{ij} & \ddots & \\ & & & \sigma_N \end{bmatrix} \tag{2.3}$$

and

$$\sigma_i = \sum_{j=1}^{N} a_{ij}, \tag{2.4}$$

where we assume (without loss of generality) that $a_{ii} = 0$ *for all i.*

For the sake of clarity of presentation we restrict opinions to being one dimensional. We note that the consensus model can be easily extended for opinions s_i in $\mathbb{R}^n$ and that all results discussed will still hold with analogous proofs.

If we interpret a_{ij} as measuring the amount of influence agent j exerts on agent i, (in particular, in this interpretation if $a_{ij} = 0$ then agent j does not influence agent i) then we can interpret the matrix L as encoding the structure of a network on which the agents are interacting. In this interpretation, each vertex of the graph G represents an agent and an edge from vertex i to vertex j represents that agent i is exerting influence on agent j directly, i.e. that $a_{ji} > 0$. When one agent influences another directly, the influenced agent adjusts its opinion towards the influencer's at a rate proportional to how much influence is being exerted.

Remark 1 We note that L is the transpose of the *Laplacian matrix* of the directed graph G (with the convention that the diagonal entries of L measure the *indegree* of the corresponding vertex in G.)

In this work we will study how different conditions on the matrix L affect the behavior of the model. We will mostly be concerned with whether a *consensus* is formed among the agents, i.e. if the opinions of every agent converge to the same value.

Definition 2 *We say that the consensus model converges to a* ***consensus*** *if there exists* $\alpha \in \mathbb{R}$ *such that:*

$$\lim_{t \to +\infty} \mathbf{s}(t) = \alpha \mathbf{1}, \tag{2.5}$$

where $\mathbf{1} = (1, 1, .., 1)^T$.

We will also be interested in convergence to a consensus that is *unconditional*.

Definition 3 *The consensus model* ***converges to a consensus unconditionally*** *if a consensus is reached for any choice of the initial state* $\mathbf{s}(0)$.

Since L encodes the structure of a network, the conditions we place on L can be viewed as conditions on the structure of G, the network on which the agents are interacting. We will show that in the case where no extra assumptions are made about L that the model always converges, although not necessarily to a consensus. We find that in the case where L is irreducible that the model converges unconditionally to a consensus; if we add the assumption that L is symmetric we find that the opinion the agents convene on is the mean of their *initial* opinions. Assuming that L is irreducible is equivalent to assuming that G is a directed graph which is *strongly connected*. Adding the assumption that L is symmetric is equivalent to assuming that G is a connected undirected graph (often called a *simple graph*) where every interaction between agents is mutual. However, we find that these assumptions - while sufficient for unconditional convergence to a consensus - aren't necessary and can be weakened. We finally present conditions on L that are equivalent to unconditional convergence to a consensus and are weaker then irreducibility. In terms of G these conditions are equivalent to demanding that G be *weakly connected* with only one

isolated strongly connected component. In Fig. 1, we give an illustration of these cases of study for L and two examples of the time evolution of opinions are plotted in Fig. 2. Similar work including nonlinearities in the case of undirected connected graphs can be found in [106, 109].

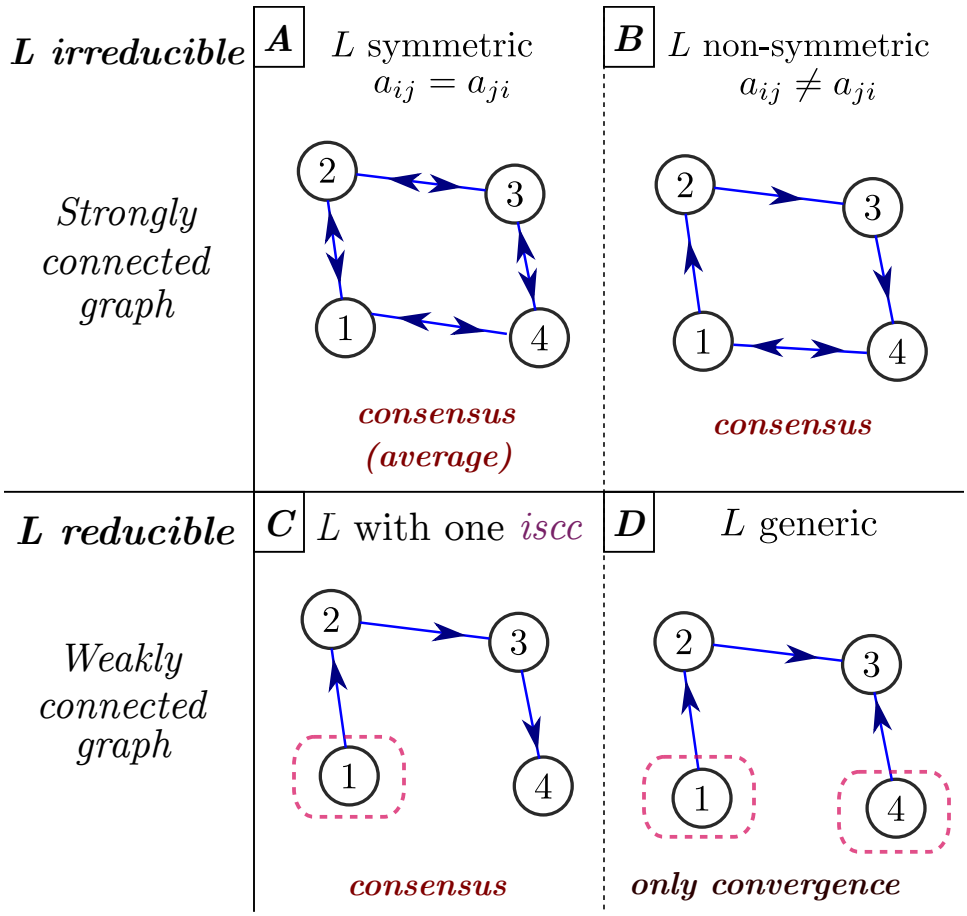

Figure 1: The four cases (A, B, C, D) for the structure of the network: from the strongest assumption on L (symmetric irreducible) to a generic matrix (*weakly* connected or disconnected). A sufficient condition to reach consensus is to have L with only one *isolated strongly connected component*.

Before investigating the cases discussed above we present a well known property of L.

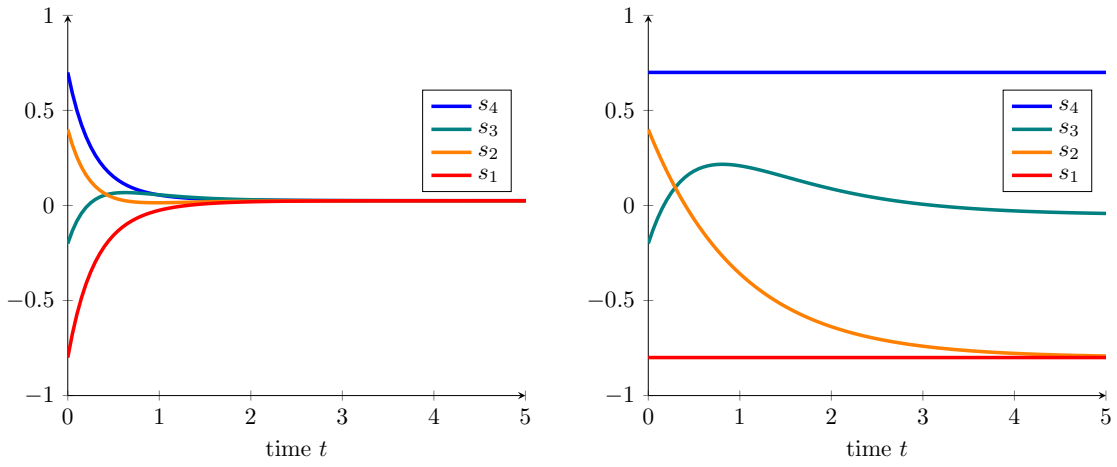

Figure 2: Evolution of the opinions $s_i(t)$ using an irreducible symmetric L (case A Figure 1) and a reducible L (case D Figure 1) from Fig. 1. Notice that agents $\{1\}$ and $\{4\}$ do not directly or indirectly influence each other.

Regardless of the coefficients a_{ij}, L has a special structure which provides information on its spectrum.

Proposition 1 *The matrix L is a diagonal dominant matrix, its eigenvalues satisfy $Re(\lambda_i) \geq 0$ and no eigenvalue λ_i is purely imaginary except zero.*

Proof

It is direct consequence of the Gershgorin theorem (see Appendix A.1). ■

Note also that 0 is an eigenvalue of L (i.e. $\lambda_1 = 0$) since the one-vector $\mathbf{1}$ is always an eigenvector of L since L is a row stochastic matrix. The main challenge is now to find conditions on L such that the 0 eigenvalue is either simple or has as many eigenvectors as its algebraic multiplicity.

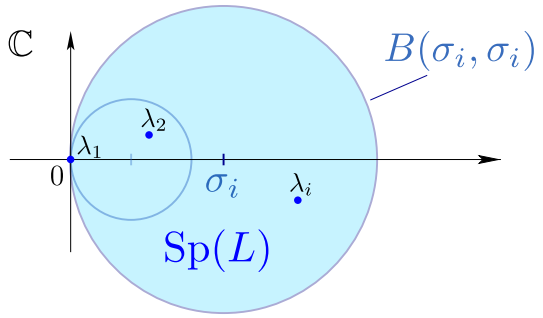

Figure 3: The spectrum of L, denoted $\mathrm{Sp}(L)$, is contained in the Gershgorin discs $B(\sigma_i, \sigma_i)$. Thus, the eigenvalues of L are either zero (λ_1 in this representation) or have a real part strictly positive.

2.1.2 Strongly connected networks (L irreducible)

We first study the case where L is assumed to be irreducible (see Fig. 1 **A** and **B**). In terms of the network G on which the agents interact, this means that G is *strongly connected*. Intuitively we are assuming that given any two agents i and j, that agent i exerts influence over agent j (either directly or indirectly) and vice versa. Formally, in terms of L, we introduce the following definitions.

Definition 4 *Consider N agents and $1 \leq i,\ j \leq N$. We say that agent i is influencing agent j if there exists a path $q_1, \ldots, , q_k$ such that:*

$$q_1 = i, \quad q_k = j, \quad \text{and } a_{q_m, q_{m+1}} > 0 \text{ for all } 1 \leq m \leq k-1. \tag{2.6}$$

and we write $i \sim j$.

In other words, L is irreducible if and only if for any $i \neq j$ we have both $i \sim j$ and $j \sim i$.

Definition 5 *A graph G is said to be strongly connected if for any vertices i and j:*

$$i \sim j \quad \textbf{\textit{and}} \quad j \sim i. \tag{2.7}$$

A graph G is said to be weakly connected if for any vertices i and j:

$$i \sim j \quad \textbf{\textit{or}} \quad j \sim i. \tag{2.8}$$

The information about the eigenvalues of L given in Proposition 1 is not sufficient to characterize the long-term behavior of the model. We need to further investigate the algebraic multiplicity of the zero eigenvalue of L. Since we are assuming that L is irreducible, we can deduce that the eigenvalue $\lambda_1 = 0$ is in fact simple.

Lemma 1 *If L is irreducible then the eigenvalue* 0 *is simple.*

Proof

Apply Perron-Frobenus theorem (see Appendix A.1). ■

Using the properties of L stated above we can now exploit the following general fact from the theory of ordinary differential equations to deduce that the model converges to a consensus unconditionally in the case where L is irreducible. This result will prove useful for showing the convergence of the model and studying consensus in the case of an arbitrary choice for L as well.

Lemma 2 *Given a linear system defined by*

$$\mathbf{x}' = A\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0.$$

Assume A has a zero eigenvalue of multiplicity m with m linearly independent associated eigenvectors and every non-zero eigenvalue λ of A satisfies $Re(\lambda) < 0$. Then

$$\lim_{t \to +\infty} \mathbf{x}(t) = \mathbf{u}$$

where $\mathbf{u}$ *is in the center subspace of A.*

Proof

See Appendix A.2. ■

Since in the case where L is irreducible its zero eigenvalue is simple, we obtain immediately the convergence of the consensus model to a consensus.

Theorem 1 *If L is irreducible then the consensus model converges to a consensus unconditionally.*

Proof

By Proposition 1 we know that the eigenvalues of of $-L$ satisfy $\mathrm{Re}(\lambda_i) \leq 0$. By Lemma 1 we know that 0 is a simple eigenvalue of $-L$ associated to $\mathbf{1}$ and therefore by Lemma 2 we can deduce that the model converges to a consensus as the center subspace of L is spanned by $\mathbf{1}$. ■

2.1.2.1 Connected undirected networks (L symmetric)

We now study a well known case of when L is irreducible by adding the assumption that L is also symmetric. Here we are implicitly assuming that every direct interaction between individuals is mutual. We know from the last section that a consensus will be reached unconditionally, however in this case we will show that the opinion the agents convene on is the *initial average of opinions*. We note that these results have been well studied in the past; we include them to highlight the special role that symmetry plays in the consensus model and to contextualize results that arise in the subsequent study of the stochastic analogue to this model - especially convergence rates. The average opinion is defined as:

$$\bar{s}(t) := \frac{1}{N} \sum_{i=1}^{N} s_i(t). \tag{2.9}$$

We first show that the average opinion is conserved by the consensus model in this case.

Lemma 3 *If L is symmetric then the average opinion $\bar{s}(t)$ (2.9) is conserved by the consensus model.*

Proof

Using the symmetry $a_{ij} = a_{ji}$, we find:

$$\bar{s}\,'(t) \;= \frac{1}{N}\sum_{i=1}^{N} s_i'(t) = \frac{1}{N}\sum_{i,j} a_{ij}(s_j(t) - s_i(t)) = 0. \qquad \blacksquare$$

We can now show that the opinion the agents convene on is the average initial opinion and that the convergence is exponential with rate at least λ_2, the second smallest eigenvalue of L, also known as the *Fieldler number* or *algebraic connectivity* of the graph G.

Remark 2 We will point out that we can impose a weaker condition than symmetry on L and still maintain the conservation of the average opinion (and therefore convergence to a consensus that is the initial average opinion). Recall that if L is not symmetric then the graph associated to L is a directed graph. If we assume that the graph is *balanced*, i.e. that L satisfies:

$$\sum_{j=1}^{N} a_{ij} = \sum_{j=1}^{N} a_{ji} \qquad \text{for all } 1 \le i \le N, \tag{2.10}$$

then $\mathbf{1}$ must be a left eigenvector of L associated to zero as well (equivalently a right zero eigenvector of L^T). In this case the average opinion is conserved as:

$$\bar{s}\,'(t) = \langle -L\mathbf{s}, \mathbf{1}\rangle = \langle \mathbf{s}, -L^T\mathbf{1}\rangle = 0 \tag{2.11}$$

using (2.10).

However, in the following corollary we will maintain the assumption that L is symmetric in order to guarantee that it can be diagonalized.

Corollary 1 *Suppose L is irreducible and symmetric. Then the solution of the consensus model, $\mathbf{s}(t)$, satisfies:*

$$\mathbf{s}(t) \to \bar{s}(0)\,\mathbf{1} \quad \textit{as} \quad t \to +\infty$$

with $\bar{s}(0)$ *the initial average opinion* (2.9)*. Moreover,* $|s_i(t) - \bar{s}(0)| \leq Ce^{-\lambda_2 t}$ *where C depends only on the initial condition and* λ_2 *is the second largest eigenvalue of L.*

Proof

See Appendix A.1 ■

2.1.3 Weakly connected and disconnected networks (arbitrary L)

In this section we remove the assumption that L is irreducible. In terms of the network associated to L, this translates to assuming that the network on which the agents interact is merely *weakly connected* or disconnected. The case of weakly connected or disconnected graphs is more delicate. In this case we are implicitly assuming that individuals, or communities of individuals, may be isolated from influence from the network. Under these assumptions on L a consensus might not always be reached. See for example the right plot in Figure 2. However we will show that the dynamics always converge and provide a condition on L weaker than irreducibility that is equivalent to unconditional convergence to a consensus. Our main tool for studying the dynamics in this case is Lemma 2 and the decomposition of L according to its strongly connected components (similar to the so-called *Frobenius normal form*).

Given a weakly connected or disconnected graph we may partition its set of vertices into *strongly connected components*. Two vertices i and j are in the same strongly connected component if there exists a directed path from i to j and vice versa (i.e. $i \sim j$ and $j \sim i$). If we treat each strongly connected component as a vertex of a new graph G', and notice that this graph is necessarily directed acyclic, then by relabeling the vertices of G in a way that agrees with the topological ordering of the vertices in G' we can represent L in the following

form (see Appendix A.3 and Figure 4).

$$L = \begin{bmatrix} B_1 & & & \\ & \ddots & 0 & \\ & & B_k & \\ & * & & \ddots \\ & & & & B_d \end{bmatrix}. \tag{2.12}$$

Each block, B_k, on the diagonal is irreducible as they correspond to strongly connected components of G. Notice that d denotes the number of strongly connected components, thus when L is irreducible we have $d = 1$. We give an example of such a relabeling in Figure 4.

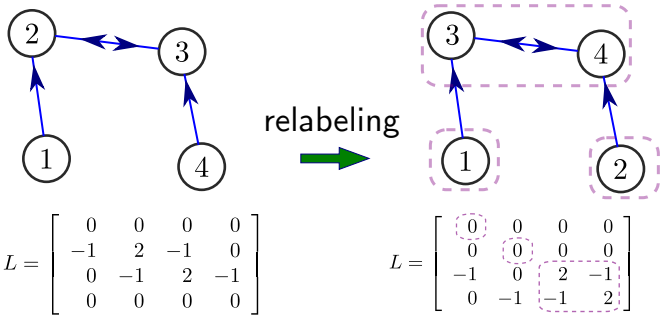

Figure 4: An example of relabeling to obtain the block form.

Notice that by the definition of L each block B_k is diagonal dominant:

$$b_{ii} \geq \sum_{j=1, j\neq i}^{N_k} -b_{ij}, \tag{2.13}$$

where b_{ij} are the coefficients of the matrix B_k and N_k is the total number of vertices in B_k. We can further distinguish between blocks.

Definition 6 *Assume L is given in the block form* (2.12)*. A block $B_k = (b_{ij})$ is called isolated if it satisfies:*

$$b_{ii} = \sum_{j=1, j\neq i}^{N_k} -b_{ij} \qquad \text{for all } 1 \leq i \leq N_k. \tag{2.14}$$

In other words, vertices from the block B_k are only influenced by vertices in block B_k:

$$\textbf{\textit{isolated}}: \qquad \textit{if } i \in B_k \textit{ and } j \notin B_k, \textit{ then } j \nsim i. \tag{2.15}$$

Note however that i might influence j (i.e. $i \sim j$). A non-isolated block B_k is called absorbing if the vertices from block B_k can only influence vertices in B_k:

$$\textbf{\textit{absorbing}}: \qquad \textit{if } i \in B_k \textit{ and } j \notin B_k, \textit{ then } i \nsim j. \tag{2.16}$$

Note again that j might influence i (i.e. $j \sim i$).

Notice that a block can be neither absorbing nor isolated. In the example from Figure 4, the matrix L is composed of two isolated blocks and one absorbing block and has no blocks that are neither. Isolated blocks of L correspond to *isolated strongly connected components* of the graph G on which the agents interact. We now establish a correspondence between isolated blocks and zero eigenvectors of L.

Proposition 2 *If L has m isolated blocks there exist at least m linearly independent eigenvectors of L associated to the zero eigenvalue.*

Proof

Assume that B_k is an isolated block of L and that B_k is $N_k \times N_k$. Then, the vector $\mathbf{1}_{N_k} = (1, 1, ..., 1)^T$ of length N_k is a right eigenvector of B_k associated to the zero eigenvalue. Since B_k is irreducible and is isolated we have by the Perron-Frobenius theorem that $\mathbf{1}_{N_k}$ is the only zero right eigenvector of B_k (see the proof of Lemma 1). Since for any matrix the collection of left and right eigenvalues as well as the dimensions of the corresponding eigenspaces are equal, there must exist a unique *left* eigenvector of B_k associated to zero denoted $\bar{\mathbf{u}}_k$. We then construct the following vector

$$\mathbf{w}_k = (0, \ldots, 0, \bar{\mathbf{u}}_k, 0, \ldots, 0)^T$$

where the non-zeros entries correspond the indices of the block B_k in the decomposition (2.12). Using that B_k is isolated, we deduce that $\mathbf{w}_k$ is a *left* eigenvector of L, i.e. $\mathbf{w}_k^T L = \mathbf{0}^T$.

We can now repeat this argument for all isolated blocks B_k to produce a collection of left eigenvectors $\mathbf{w}_k^T$ of L associated to the zero eigenvalue $\lambda_1 = 0$. These eigenvectors are necessarily linearly independent as they are all nonzero in disjoint coordinates. Therefore, since the dimensions of the left and right eigenspaces of L are equal, if there are m isolated blocks of L there must be m linearly independent right eigenvectors of L associated to zero.■

So, if L has m isolated blocks then it has at least m linearly independent zero eigenvectors. We will eventually show that these are the only zero eigenvectors of L and deduce convergence by Lemma 2; as a first step we show that non isolated blocks of L do not have a zero eigenvalue.

Proposition 3 *Suppose that B_k is a non-isolated block. Then B_k does not have a zero eigenvalue.*

Proof

Denote by b_{ij} the components of B_k. Since B_k is a non-isolated block, there exists i_0 such that:

$$b_{i_0 i_0} > \sum_{j=1, j\neq i_0}^{N_k} -b_{i_0 j}.$$

In particular, the constant vector $\mathbf{1}_{N_k} = (1, \ldots, 1)^T$ cannot be a zero-eigenvector of B_k. Suppose for the sake of contradiction that B_k had a zero eigenvector, $\mathbf{u}$, and let u_i be the largest entry of $\mathbf{u}$ i.e. $u_i = |\mathbf{u}|_\infty$ (we assume without loss of generality that $u_i > 0$). Then we have:

$$\begin{aligned} b_{ii}u_i + \sum_{j=1, j\neq i}^{N_k} b_{ij}u_j \;=\; 0 \qquad &\Rightarrow \qquad (-\sum_{j=1, j\neq i}^{N_k} b_{ij})u_i + \sum_{j=1, j\neq i}^{N_k} b_{ij}u_j \;\leq\; 0 \\ &\Rightarrow \qquad \sum_{j=1, j\neq i}^{N_k} b_{ij}(u_j - u_i) \;\leq\; 0. \end{aligned}$$

Since $b_{ij} \leq 0$ (for $j \neq i$) and $u_j - u_i \leq 0$, we must have $u_j = u_i$ if $b_{ij} < 0$. In other words, the coefficients of the eigenvector $\mathbf{u}$ are constant on the indices connected to i:

$$u_j = |\mathbf{u}|_\infty \quad \text{if} \quad i \sim j.$$

Iterating the argument, we deduce that $u_j = |\mathbf{u}|_\infty$ if there exists a path joining i to j. Since B_k is irreducible we deduce that $u_j = |\mathbf{u}|_\infty$ for all j. Therefore, $\mathbf{u} = (|\mathbf{u}|_\infty, \ldots, |\mathbf{u}|_\infty)$. This is a contradiction as $\mathbf{1}_{N_k}$ is not a zero eigenvector of B_k. ■

Combining propositions 2 and 3 leads to the following theorem.

Theorem 2 *The consensus model* (2.1) *converges (but not necessarily to a consensus).*

Proof

We want to show that the multiplicity of the zero eigenvalue $\lambda_1 = 0$ is equal to the number of (independent) zero eigenvectors $\mathbf{w}$ of L. From proposition 2, we already know that there exist m zero eigenvectors with m the number of isolated blocks. Thus, it remains to show that the multiplicity of the zero eigenvalue is m.

Using the decomposition (2.12), the spectrum of L can be written as a union

$$\mathrm{Sp}(L) = \bigcup_{k=1}^{d} \mathrm{Sp}(B_k) \tag{2.17}$$

where the matrices B_k correspond to the diagonal blocks in the decomposition (2.12) (i.e. the strongly connected components). Since only isolated blocks B_k have zero eigenvalues (Proposition 3), the spectrum of L contains exactly m zero eigenvalues. Therefore we obtain the convergence of the dynamics using Lemma 2. ■

Remark 3 In the proof of Lemma 2 (see Appendix A.2) we find that we can bound the convergence of $\mathbf{s}(t)$ in terms of the spectral gap of L, denoted λ_2, some $0 < \epsilon < \mathrm{Re}(\lambda_2)$ and some $C > 0$:

$$\left\|\mathbf{s}(t)\right\| \leq Ce^{-(\mathrm{Re}(\lambda_2)+\epsilon)t}. \tag{2.18}$$

If L is symmetric, λ_2 is known as the *algebraic connectivity* of the graph G and is a measure of how well connected G is. In this context the bound (2.18) can be interpreted as stating that the consensus model converges faster if the network on which agents interact is better connected.

We can observe from Theorem 2 that the number of zero eigenvectors of L is in correspondence with the number of isolated blocks of L. This observation supplies us with a condition on L that is equivalent to unconditional convergence to a consensus.

Corollary 2 *The consensus model converges to a consensus unconditionally if and only if L has exactly one isolated block.*

Proof

The vector $\mathbf{1}$ is always a zero eigenvector of L. Since L has only one isolated block, the eigenvalue 0 must be simple by Propositions 2 and 3. Therefore, $\mathbf{1}$ is the only zero eigenvector of L and therefore Lemma 2 implies that the dynamics converge to a consensus as the center subspace of L is spanned by $\mathbf{1}$. ■

Remark 4 Recall that isolated blocks of L correspond to isolated blocks of the graph G on which the agents interact. We can interpret an isolated block as "leading" the rest of the network. Given a network with only one isolated block the consensus model implies that nodes in an isolated block receive no influence from nodes outside of the isolated block and since we know by Corollary 2 that the entire network reaches a consensus, we deduce that the consensus converged to by the whole network is the consensus converged to by agents in the isolated block. The network can fail to converge to a consensus if there are multiple isolated blocks that converge on different opinions (See the right plot in Figure 2) - this clearly depends on the initial opinions of the agents in the network. Thus, if a network has multiple isolated blocks the consensus model does *not* converge to a consensus

unconditionally (although a consensus might still be reached if all isolated blocks happen to converge on the same opinion).

2.1.4 Consensus as a control problem

We now slightly shift our perspective to consider a question inspired by control theory. Given N agents with initial opinions $\{ s_1(0), ..., s_N(0) \}$ and a desired opinion $s^* \in \mathbb{R}^d$ that one would like every agent to converge on, we investigate if there is a way to connect the agents in a network so that under the consensus model we have:

$$\lim_{t \to +\infty} \mathbf{s}(t) = s^* \mathbf{1}.$$

We can use our observations about how isolated blocks affect the behavior of the consensus model to answer this question. For ease of notation we denote:

$$s_{max} := s_i \quad \text{where} \quad |s_i| = \max \{ \ | s_1(0)|, ..., |s_N(0)| \} ,$$

$$s_{min} := s_j \quad \text{where} \quad |s_j| = \min \{ \ | s_1(0)|, ..., |s_N(0)| \} .$$

We also note that there is no way to connect the agents to achieve the desired result if $s^* \notin \text{Conv}(\{ s_1(0), ..., s_N(0) \})$ (the convex hull of initial opinions) as the consensus model forces the convex hull of $\{ s_1(t), ..., s_N(t) \}$ to decrease in diameter as time evolves.

Proposition 4 *Given* $(s_1(0), ..., s_N(0))$ *and* $s^* \in$ *Conv*$(\{ s_1(0), ..., s_N(0) \})$ *there exists L such that:*

$$\lim_{t \to +\infty} \mathbf{s}(t) = s^* \mathbf{1}.$$

Proof

For ease of notation, assume

$$s_{max} = s_1(0) \quad \text{and} \quad s_{min} = s_2(0).$$

We first consider the two-agent network who's evolution is given by

$$\begin{pmatrix} s_1'(t) \\ s_2'(t) \end{pmatrix} = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix} \begin{pmatrix} s_{max} \\ s_{min} \end{pmatrix}. \tag{2.19}$$

We find that if we choose α and β such that

$$s^* = \frac{1}{\alpha + \beta}(\beta s_{max} + \alpha s_{min}),$$

we have

$$\lim_{t \to \infty} \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} s^* \\ s^* \end{pmatrix}.$$

So, given N agents, if we connect agents 1 and 2 with the weights found above and the rest of the agents in a "chain", i.e. we choose an L of the form:

$$L = \begin{pmatrix} \beta & -\beta & 0 & 0 & 0 & \dots & 0 \\ -\alpha & \alpha & 0 & 0 & 0 & \dots & 0 \\ 0 & -1 & & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & -1 & 1 & \dots & 0 \\ & & & \vdots & & & \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

then the network on which the agents interact will only have one isolated block consisting of agent 1 and agent 2. Our choice of weights guarantees that the opinions of agents 1 and 2 converge to s^* and since they comprise the only isolated block, Corollary 2 guarantees that we must have that:

$$\lim_{t \to +\infty} \mathbf{s}(t) = s^* \mathbf{1}$$

as desired. ■

2.2 Consensus dynamics: stochastic approach

2.2.1 Introduction

We now adopt another point of view to model opinion dynamics. Rather than changing their opinions gradually according to their neighbors, agents now "jump" their opinion to one of their neighbors. The weights a_{ij} now encode the *probability* of agent i switching its opinion to agent j. This stochastic analogue to the consensus model has received less attention than its deterministic counterpart studied above. Many of the same phenomena found in the deterministic case have analogues in this case, however a contrast in convergence rates can be observed. In the context of opinion formation, jumping models the assumption that people rarely take into account the opinions of all of those they interact with when deciding how to change their opinion and instead appeal to someone they view as an authority. If we interpret the weights as encoding the influence one agent has over another, we can interpret the model as saying that one usually adopts the opinion of the individual who exerts the most influence over them. However this shouldn't always be the case as individuals will occasionally "rebel" against their prevailing local influence. Mathematically, we will use a Poisson process of rate a_{ij} to model this event; opinion S_i "jumps" to opinion S_j at rate a_{ij}.

$$(S_i, S_j) \overset{\text{rate } a_{ij}}{\rightsquigarrow} (S_j, S_j). \tag{2.20}$$

The dynamics are now a continuous time Markov chain instead of a system of differential equations. We begin by formally defining the dynamics in terms of the stochastic process S_i that describes the evolution of the opinion of agent i.

Definition 7 *Consider N agents with opinion $S_i(t) \in \mathbb{R}$, $t \geq 0$. Let $(a_{ij})_{ij}$ be the entries of the adjacency matrix of the graph G on which the agents interact. For each tuple (i, j)*

with $i \neq j$, we associate an independent Poisson process $N^{ij}(t)$ with rate a_{ij}. The ***stochastic consensus model*** *(SCM) is defined as:*

$$\mathrm{d}S_i = \sum_{j=1, j\neq i}^{N} (S_j - S_i)\,\mathrm{d}N^{ij}. \tag{2.21}$$

Again, for the sake of clarity of presentation we restrict opinions to being one dimensional and we note that the stochastic consensus model can be easily extended for opinions in $\mathbb{R}^n$ and that all results discussed will still hold with analogous proofs. Notice that one can write the Poisson process as $N^{ij}(t) = N(a_{ij} \cdot t)$ where $N(t)$ is a Poisson process of rate 1. Each time a Poisson process N^{ij} increases by one unit, the opinion of agent i, S_i, "jumps" to S_j. We illustrate these dynamics in Figure 5 on a graph with only three nodes and two links.

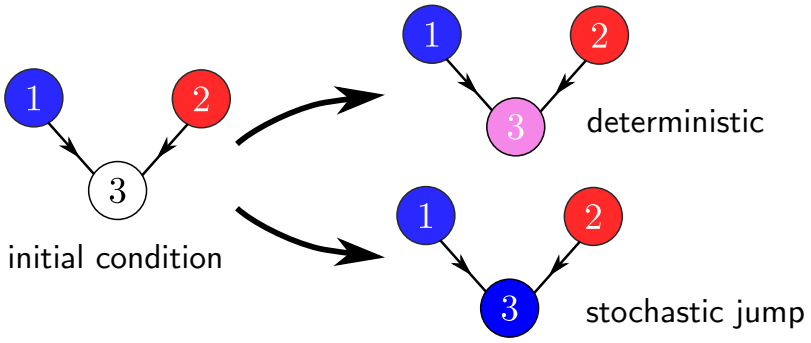

Figure 5: Illustration of the deterministic and stochastic dynamics. The node 3 receives influence from both node 1 and node 2. In the deterministic dynamics (2.1), the opinion of s_3 will become a weighted average of opinion s_1 and s_2. However, in the stochastic dynamics (2.21), the opinion S_3 will continue switching between opinions S_1 and S_2.

The stochastic consensus model can also be written in terms of its generator. For any pair (i, j), denote by Φ_{ij} the function:

$$\Phi_{ij}(s_1, \ldots, s_i, \ldots, s_j, \ldots, s_N) \quad = \quad (s_1, \ldots, s_j, \ldots, s_j, \ldots, s_N). \tag{2.22}$$

Let $\mathbf{S} = (S_1, \ldots, S_N)$. For any smooth test function $\varphi : \mathbb{R}^N \to \mathbb{R}$, we have:

$$\mathrm{d}\mathbb{E}[\varphi(\mathbf{S})] = \sum_{1\leq i,j\leq N,\, j\neq i} a_{ij}\mathbb{E}[\varphi(\Phi_{ij}(\mathbf{S})) - \varphi(\mathbf{S})]\,\mathrm{d}t. \tag{2.23}$$

The stochastic model can also be more simply described through its associated embedded Markov chain. Instead of describing the continuous evolution of each opinion $S_i(t)$, we only record the evolution of each jump:

$$S_i^n = S_i(T_n)$$

where T_n is the time of the n^{th} jump. Notice that $T_{n+1} - T_n$ is an exponential random variable with frequency σ - the rate of any jump occurring:

$$\sigma = \sum_{i=1}^{N} \sigma_i = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} a_{ij}, \tag{2.24}$$

by the superposition property of independent Poisson processes. The evolution of $(S_i^n)_{1\leq i\leq N}$ is then given by a discrete Markov process described by the transition probability matrix:

$$p_{ij} = \mathbb{P}\Big((S_i^n, S_j^n) \rightsquigarrow (S_j^{n+1}, S_j^{n+1})\Big) = \frac{a_{ij}}{\sigma}, \tag{2.25}$$

where a_{ij} is the rate of agent i switching to the opinion of agent j and σ given by (2.24). The generator of this Markov chain can be described as:

$$\mathbb{E}[\varphi(\mathbf{S}^{n+1})] = \mathbb{E}[\varphi(\mathbf{S}^n)] + \sum_{1\leq i,j\leq N,\, j\neq i} p_{ij}\mathbb{E}[\varphi(\Phi_{ij}(\mathbf{S}^n)) - \varphi(\mathbf{S}^n)]. \tag{2.26}$$

Remark 5 The stochastic consensus model can be approximated and simulated by Algorithm 1 with discrete timestep Δt. Let T be the duration of the simulation and t be the current time.

We will now investigate the long-time behavior of these dynamics and analyze how the structure of the graph G affects the convergence.

Remark 6 Taking the expectation of the stochastic dynamics (2.21), we recover the deterministic model (2.1). Indeed, denote

$$\bar{s}_i = \mathbb{E}[S_i(t)], \tag{2.27}$$

Algorithm 1 Consensus model (stochastic)

procedure STOCHASTIC CONSENSUS
  while $t \leq T$ **do**
    for $i \in \{1, \ldots, N\}$ **do**
      if $\text{Unif}(0,1) < 1 - e^{-\sigma_i \Delta t}$ **then**
        Pick an agent j according to Prob{ Picking agent j } $= \frac{a_{ij}}{\sigma_i}$
        Set $S_i(t) = S_j(t-1)$
      end if
    end for
    $t = t + \Delta t$
  end while
end procedure

using that $\mathbb{E}[N^{ij}(t)] = a_{ij} \cdot t$ [101] and that $\mathrm{d}N^{ij}(t)$ is independent of $S_i(t)$ and $S_j(t)$, we deduce:

$$\mathrm{d}\bar{s}_i = \sum_{j=1, j\neq i}^{N} (\bar{s}_j - \bar{s}_i) a_{ij} \, \mathrm{d}t \tag{2.28}$$

which corresponds to the deterministic dynamics (2.1). Thus, we can deduce the behavior of the deterministic model from the stochastic model using a simple Monte-Carlo method. Conversely, the deterministic model gives information about the average behavior of the stochastic model. We also notice that in the case that all agents are connected ($a_{ij} = 1$ for all i, j), we recover a special case of the so-called "choose-the-leader" dynamics [26, 25].

2.2.2 Convergence (in probability)

Analogous to our discussion of the deterministic consensus model, we are going to prove that the stochastic model converges to a consensus unconditionally if the graph, G, has only one isolated strongly connected component. However, the convergence will be much slower than the deterministic model. The key tool is to use the notions of *absorbing states* and *absorbing Markov chains*.

Definition 8 (Absorbing states and absorbing Markov chain) *Let* $\mathbf{S}^n$ *be a Markov chain on a state space* C*. A state* $\mathbf{S} \in C$ *is called absorbing if* $p_{\mathbf{S},\mathbf{S}} = 1$ *– that is, the probability of staying in* $\mathbf{S}$ *given that the chain has arrived in* $\mathbf{S}$ *is 1. The Markov chain* $\mathbf{S}^n$ *is called an absorbing Markov chain if for any starting state* $\mathbf{S}^0 \in C$*, the chain can reach an absorbing state in finitely many transitions with positive probability. A Markov chain is absorbed if it reaches an absorbing state.*

For instance, in our case, $\mathbf{1} = (1, 1, .., 1)^T$ is an absorbing state. We can find sufficient conditions on the network G for the embedded Markov chain $\mathbf{S}^n$ of the stochastic consensus model to be absorbing.

Theorem 3 *Assume that the adjacency matrix* $L = (a_{ij})_{ij}$ *has only one isolated strongly connected component. Then the stochastic consensus model (SCM) converges with probability* 1 *to a consensus.*

Proof

We are going to show the convergence of the embedded Markov chain $\{\mathbf{S}^n\}_n$ from which we can easily deduce the convergence of the continuous process $\{\mathbf{S}(t)\}_{t\geq 0}$.

Notice first that for a given initial condition $\mathbf{S}^0$, the embedded Markov chain $\mathbf{S}^n$ evolves in a finite space C of size N^N. Indeed, each component $\mathbf{S}_i^n$ has to be one of the initial opinion components $\left(\mathbf{S}_i^0\right)_{1\leq i\leq N}$. Thus, if we can show that $\{\mathbf{S}^n\}_n$ is an absorbing Markov chain, then $\mathbf{S}^n$ will be absorbed with probability 1 [Theorem 11.3, [54]].

Consider now any opinion S_i^0 in the isolated strongly connected component of the graph. By assumption on the adjacency matrix L, there exists for any j a path joining i to j (2.6). Thus, there is a strictly positive probability (i.e. $p_{i,q_2} \dots p_{q_{k-1},j}$ where $q_2, \dots, q_{k-1}$ path joining i to j) such that $S_j^k = S_i^k = S_i^0$ after k steps. Iterating the argument for all j, we find a non-zero probability that $S_j^n = S_i^0$ for all j with n finite (with a rough upper-bound of N^N). Thus, we conclude that $\{\mathbf{S}^n\}_n$ is an absorbing Markov chain. ■

Remark 7 From Theorem 3, we can deduce a simple proof for corollary 2 concerning the convergence of the deterministic model. Indeed, for any $\varepsilon > 0$ and initial condition $\mathbf{S}^0$, there exists a time T such that for any $t \geq T$ we have $\mathbb{P}\Big(\mathbf{S}(t) \text{ is a consensus}\Big) \geq 1 - \varepsilon$. Taking the expectation of the stochastic process, $\mathbf{S}(t) = \mathbb{E}[\mathbf{S}(t)]$, we deduce that $\mathbf{S}(t)$ converges to the expected consensus. However notice that in using this argument we lose the rate of convergence as we did not analyze the eigenvalues of the dynamics.

2.2.3 Decay rate for undirected graphs

Theorem 3 gives a necessary and sufficient condition for the unconditional convergence to consensus of the stochastic model. However, we do not have any information about the rate of convergence of the process. In the deterministic case, we have seen that the convergence is exponential with a rate given in terms of the second largest eigenvalue of the matrix L, λ_2. We would like to investigate if one could obtain a similar explicit convergence rate for the stochastic model. With this aim, we are going to investigate the empirical variance of the process:

$$V(S(t)) = \frac{1}{N}\sum_{i=1}^{N} |S_i(t) - \bar{S}(t)|^2 = \frac{1}{2N^2}\sum_{1\leq i,j\leq N} |S_i(t) - S_j(t)|^2, \tag{2.29}$$

where $\bar{S}(t)$ is the empirical mean:

$$\overline{S}(t) = \frac{1}{N}\sum_{i=1}^{N} S_i(t). \tag{2.30}$$

We will need to assume that the interactions are symmetric, i.e. that L is symmetric.

Lemma 4 *Suppose that the adjacency matrix L is symmetric, then the empirical variance* $V(\mathbf{S})$ (2.29) *decays according to*

$$\frac{d}{dt}\mathbb{E}[V(\mathbf{S})] = -\frac{1}{N^2}\mathbb{E}[\sum_{1\leq i,j\leq N} a_{ij}|S_j - S_i|^2]. \tag{2.31}$$

Proof

Using the generator of the stochastic process (2.23), we find:

$$\frac{d}{dt}\mathbb{E}[V(\mathbf{S})] = \sum_{1\leq i,j\leq N} a_{ij}\mathbb{E}[V(\Phi_{ij}(\mathbf{S})) - V(\mathbf{S})].$$

Moreover, notice the following equality:

$$V(\Phi_{ij}(s)) = V(s) + \frac{2}{2N^2}\sum_{k=1}^{N}\Big(|S_k - S_j|^2 - |S_k - S_i|^2\Big) - \frac{2}{2N^2}|S_j - S_i|^2.$$

Indeed, when i jumps to opinion j (i.e. $\Phi_{ij}(\mathbf{S})$), all the '(i,k)' index become '(j,k)' but the difference between opinion i and j is zero. We then deduce:

$$V(\Phi_{ij}(\mathbf{S})) + V(\Phi_{ji}(\mathbf{S})) = 2V(\mathbf{S}) - \frac{4}{2N^2}|S_j - S_i|^2.$$

Furthermore, exploiting the symmetry $a_{ij} = a_{ji}$ we get that

$$a_{ij}V(\Phi_{ij}(\mathbf{S})) + a_{ji}V(\Phi_{ji}(\mathbf{S})) \quad = \quad (a_{ij} + a_{ji})\Big[V(\mathbf{S}) - \frac{2}{2N^2}|S_j - S_i|^2\Big].$$

Thus

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[V(\mathbf{S})] &= \sum_{1\leq i,j\leq N}\mathbb{E}[a_{ij}(V(\Phi_{ij}(\mathbf{S})) - V(\mathbf{S}))] \\ &= \mathbb{E}\Big[\sum_{i<j}\Big(a_{ij}V(\Phi_{ij}(\mathbf{S})) + a_{ji}V(\Phi_{ji}(\mathbf{S})) - (a_{ij}+a_{ji})V(\mathbf{S})\Big)\Big] \\ &= -\frac{1}{N^2}\mathbb{E}\Big[\sum_{i<j}(a_{ij}+a_{ji})|S_j - S_i|^2\Big] = -\frac{1}{N^2}\mathbb{E}\Big[\sum_{1\leq i,j\leq N} a_{ij}|S_j - S_i|^2\Big]\end{aligned}$$ ■

We can now deduce a lower bound for the previous equality (2.31) in terms of λ_2, the second smallest eigenvalue of the Laplacian matrix and the algebraic connectivity of the graph G.

Lemma 5 *Suppose that the Laplacian matrix L of the network G is symmetric and let λ_2 be its second smallest eigenvalue. Then:*

$$\sum_{1\leq i,j\leq N} a_{ij}|s_j - s_i|^2 \geq 2\lambda_2 N \cdot V(\mathbf{S}). \tag{2.32}$$

with $V(\mathbf{S})$ the variance (2.29)*.*

Proof

From the above, we note that

$$\langle L\mathbf{S}, \mathbf{S}\rangle = \sum_{i=1}^{N}\Big(\sum_{j=1}^{N} a_{ij}(S_i - S_j)\Big)S_i = \sum_{i,j} a_{ij}(S_i - S_j)S_i$$
$$= \sum_{j,i} a_{ji}(S_j - S_i)S_j = \sum_{j,i} a_{ij}(S_j - S_i)S_j$$

where we use the symmetry of L. Therefore we can write

$$\langle L\mathbf{S}, \mathbf{S}\rangle = \frac{1}{2}\Big(\sum_{i,j} a_{ij}(S_j - S_i)S_j + \sum_{i,j} a_{i,j}(S_j - S_i)S_i\Big)$$
$$= \frac{1}{2}\sum_{i,j} a_{ij}(S_i - S_j)(S_i - S_j) = \frac{1}{2}\sum_{i,j} a_{ij}|S_i - S_j|^2.$$

Notice also that the empirical variance can be written in two ways:

$$\frac{1}{2N^2}\sum_{i,j}|S_i - S_j|^2 = \frac{1}{N}\sum_{i}|S_i - \overline{S}|^2 = \frac{1}{N}\|\mathbf{S} - \overline{S}\mathbf{1}\|^2,$$

where $\overline{S}$ is the average opinion (2.30). Since $L\mathbf{1} = 0$, we also deduce:

$$\langle L(\mathbf{S} - \overline{S}\mathbf{1}), \mathbf{S} - \overline{S}\mathbf{1}\rangle = \langle L\mathbf{S}, \mathbf{S}\rangle.$$

Decomposing the symmetric matrix L as $L = PDP^T$ with $D = \mathrm{diag}(0, \lambda_2, \ldots, \lambda_N)$, we deduce:

$$\langle L\mathbf{S}, \mathbf{S}\rangle = \langle L(\mathbf{S} - \overline{\mathbf{S}}), \mathbf{S} - \overline{\mathbf{S}}\rangle$$
$$= \langle DP^T(\mathbf{S} - \overline{\mathbf{S}}), P^T(\mathbf{S} - \overline{\mathbf{S}})\rangle = \langle D\mathbf{w}, \mathbf{w}\rangle$$

with $\mathbf{w} = P^T(\mathbf{s} - \bar{\mathbf{s}})$. Thus,

$$\langle L\mathbf{S}, \mathbf{S}\rangle = \sum_{i=1}^{N}\lambda_i|w_i|^2 = \sum_{i=2}^{N}\lambda_i|w_i|^2 \geq \lambda_2\sum_{i=2}^{N}|w_i|^2.$$

Notice that the first component $\mathbf{w}_1$ has to be zero as the first row of P^T is $\alpha\mathbf{1}$ for some $\alpha \in \mathbb{R}$.

$$\lambda_2\sum_{i=2}^{N}|w_i|^2 = \lambda_2\sum_{i=1}^{N}|w_i|^2 = \lambda_2\|\mathbf{w}\|^2.$$

Since the matrix P preserves the norm, we have:

$$\langle L\mathbf{S}, \mathbf{S}\rangle \geq \lambda_2 \|\mathbf{w}\|^2 = \lambda_2 \|S - \overline{S}\mathbf{1}\|^2 = \lambda_2 N \cdot V(\mathbf{S}).$$

Thus

$$\sum_{i,j} a_{ij}|S_j - S_i|^2 = 2\langle L\mathbf{S}, \mathbf{S}\rangle \geq 2\lambda_2 N V(\mathbf{S}). \qquad \blacksquare$$

Using these tools we can now characterize the convergence rate of the stochastic consensus model on a symmetric network.

Theorem 4 *Assume L is symmetric and connected. Denote λ_2 its second eigenvalue. Then*

$$\mathbb{E}[V(\mathbf{S}(t))] \leq e^{\frac{-\lambda_2 t}{N}} \mathbb{E}[V(\mathbf{S}(0))].$$

Proof

From the previous lemmas we observe that

$$\frac{d}{dt}\mathbb{E}[V(\mathbf{S}(t))] = -\frac{1}{2N^2}\mathbb{E}\left[\sum_{1\leq i,j\leq N} a_{ij}|S_j - S_i|^2\right] \leq -\frac{\lambda_2}{N}\mathbb{E}[V(\mathbf{S}(0))].$$

Applying Gronwall's lemma gives the result. $\blacksquare$

Remark 8 Thus we see that the expected value of the variance of the opinions decays at rate $\frac{\lambda_2}{N}$; the dependence on the number of agents N is naturally contrasted with the deterministic case where the decay rate is only λ_2, in other words, the deterministic dynamics converge much faster than the stochastic dynamics.

Recall that in the deterministic case studied earlier we found that in the case of a balanced graph that the consensus converged to is the initial average of opinions. We can make a related observation concerning the stochastic model by exploiting the fact that the average, $\bar{S}$, is a martingale. The following corollary also gives a succinct proof of the convergence of the stochastic dynamics in this case.

Corollary 3 *Suppose G is a balanced directed graph* (2.10) *and strongly connected. Then with probability 1, the dynamics* $S(t)$ *converge to a consensus opinion* S^* *at time* $T < \infty$ *and the consensus satisfies* $\mathbb{E}[S^*] = \frac{1}{N}\sum_1^N S(0)$.

Proof

Let $\bar{S}(t) = \frac{1}{N}\sum_{i=1}^N S_i(t)$ the average. Using eq. (2.23), we deduce:

$$\mathrm{d}\mathbb{E}[\bar{S}(t)] = \sum_{1\le i,j\le N} a_{ij}\mathbb{E}[\Phi_{ij}(\bar{S}(t)) - \bar{S}(t)] = \sum_{1\le i,j\le N} \frac{a_{ij}}{N}\mathbb{E}[S_j(t) - S_i(t)] = 0$$

since the graph is balanced (see remark 2). Therefore $\bar{S}(t)$ is a martingale. Moreover, $\bar{S}(t)$ is bounded with at most N^2 states (finite) and therefore there exists a stopping time T that is finite with probability 1 such that $\bar{S}(t) \to \bar{S}^*$ almost surely. Then, by the Optional Stopping Theorem,

$$\mathbb{E}[S^*] = \mathbb{E}[\bar{S}(T)] = \mathbb{E}[\bar{S}(0)] = \bar{S}(0). \qquad \blacksquare$$

Since the graph G is also strongly connected, we know by Theorem 3 that S^* must be a consensus.

2.3 Consensus dynamics: Numerical simulations

We will now present some numerical experiments to further analyze more subtle features of both of the consensus models studied. We first examine explicitly how the *algebraic connectivity* of the interaction network affects the speed of convergence in the case of a symmetric L - recall that in both the deterministic and stochastic case that we were able to bound the speed of convergence in terms of the algebraic connectivity. Then, we focus on the case where there are multiple isolated blocks as we have shown in this case that a consensus does *not* form unconditionally. We are interested in how the opinions of isolated

blocks affect the distribution of opinions in absorbing components. To this end we will examine the case of a "battle" between two isolated blocks with opinions at the extreme ends of the possible spectrum of opinions. Each of these isolated blocks will have the same number of connections to a central absorbing component whose internal connections are all symmetric.

2.3.1 Algebraic connectivity and speed of convergence

Earlier in the discussion of the deterministic model we showed that we could characterize the convergence rate of the model in terms of λ_2, the eigenvalue of L with the second smallest real part. In the case that L was symmetric (an undirected network) we could say even more - that the rate of convergence of the dynamics was at least of of order $O(e^{-\lambda_2 t})$. The second eigenvalue λ_2 is known as the **algebraic connectivity** of the graph G, it measures in a sense *how well the graph is connected*. In this interpretation, since λ_2 bounds the speed of convergence, the dynamics converge faster if the graph is *better connected*. We illustrate with several examples. We give two examples of graphs with the same number of vertices ($|V| = 20$) but different adjacency matrices. In Fig. 6–left, we present a graph where each vertex is connected with its 4 neighbors. In Fig. 6-right, the graph is composed of two sub-graphs connected by only one link resulting in a *weak* connectivity. Indeed, computing the algebraic connectivity λ_2, we observe that the graph on the left has a higher value even though it has fewer edges.

To illustrate how this affects the speed of convergence, we perform two simulations of the dynamics starting from the same initial configuration but using the two graphs presented in figure 6. The results are given in Fig. 7. We observe that the right figure converges quickly to two "clusters" but then these clusters are very slow to move together. This is due

to the *weak* connectivity of the graph. In Fig. 8, we estimate the variance (denoted $\mathbf{V}(t)$ in the stochastic dynamics) of the opinions over time :

$$\mathbf{V}(t) = \frac{1}{2N^2} \sum_{1 \le i,j \le N} |s_j(t) - s_i(t)|^2. \tag{2.33}$$

The variances are decaying to zero (indicating that a consensus is forming). As expected, the variance is decaying faster for the graph with the larger λ_2 (left-figure). We also include the evolution of the stochastic dynamics on the graph in Fig. 6–right, from the same initial condition, in Fig. 8–left. Notice that in Fig. 8–right that the variance of the stochastic dynamics is the slowest to decay - this is an illustration of how the stochastic dynamics slow down the speed of convergence.

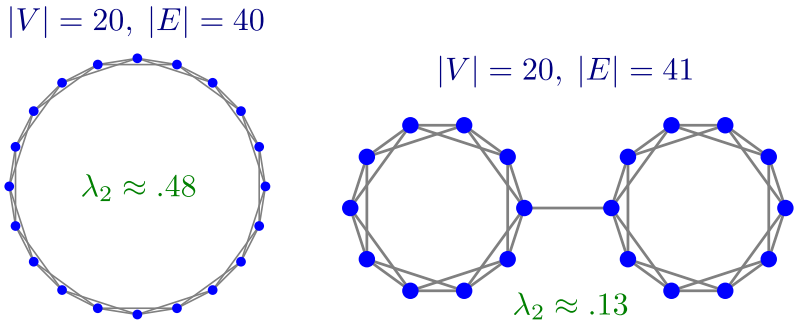

Figure 6: Examples of values of λ_2 for two graphs. Notice that even though the right graph has more edges, its algebraic connectivity λ_2 is lower.

2.3.2 Outside of consensus - multiple isolated blocks

Here we examine numerically the effects of more than one isolated block on the long-term behavior of both models. We set up the initial opinions so that each isolated block is at each extreme of the possible spectrum of opinions and that opinions in the central component are randomly distributed uniformly among the possible opinions.

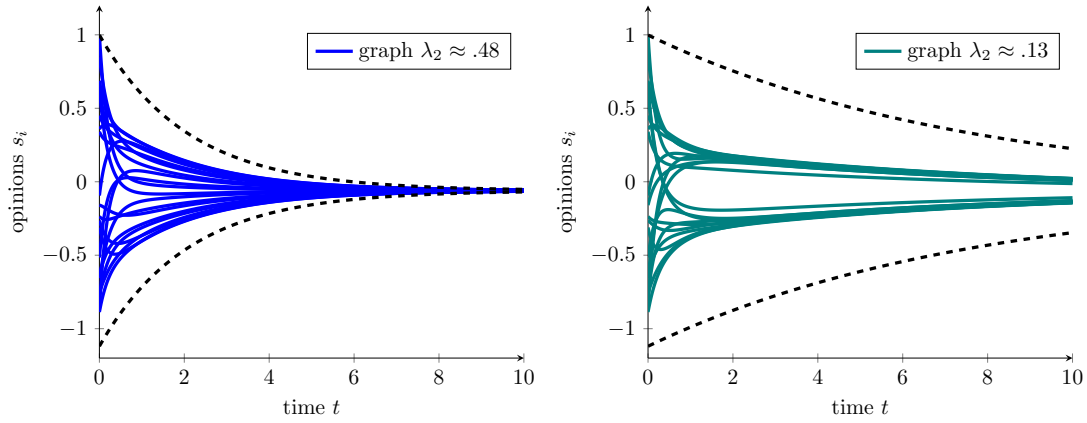

Figure 7: Evolution of opinions $s_i(t)$ over time for the two graphs of Fig. 6. The dash line indicates a decay with rate λ_2 (worst-case scenario).

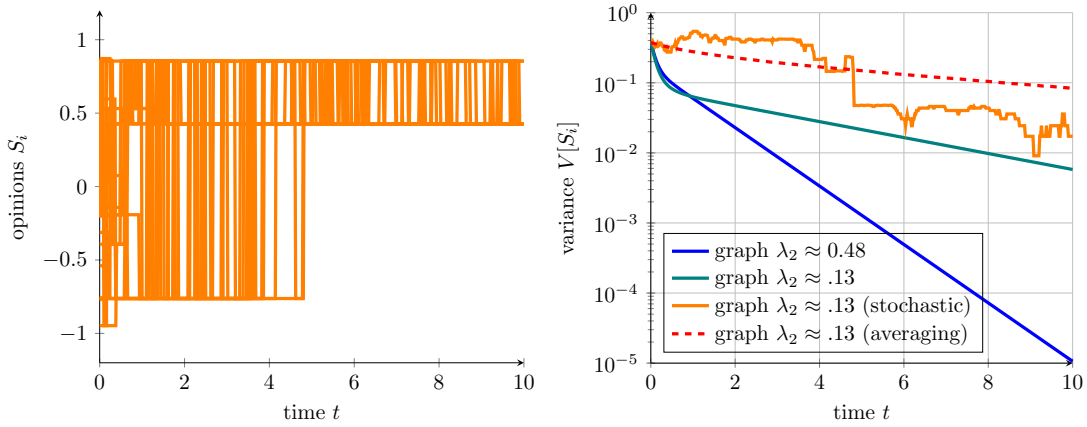

Figure 8: Left: Evolution of opinions $S_i(t)$ for the stochastic dynamics on the graph with $\lambda_2 \approx .13$. After $t = 5$, the opinions $S_i(t)$ are jumping back and force between two values. They will eventually converge to a single value due to Theorem 3. **Right:** Evolution of the variance of opinions $V[S_i]$ for the three illustrations (i.e. Fig. 7 and Fig. 8-left). The decay is faster when λ_2 is larger. Moreover, we estimate the average decay of the variance for the stochastic dynamics $\mathbb{E}[V]$ (using 10,000 realization) plotted in red.

After running a simulation of the dynamics we find that the middle component now exhibits a "gradient" of the possible opinions. Nodes close to the blue isolated block in the sense of graph distance exhibit an opinion close to that of the isolated block, likewise for nodes near the red isolated block. Nodes in the middle of the component exhibit an opinion close to neutral. This shows that in the case of multiple isolated blocks that the

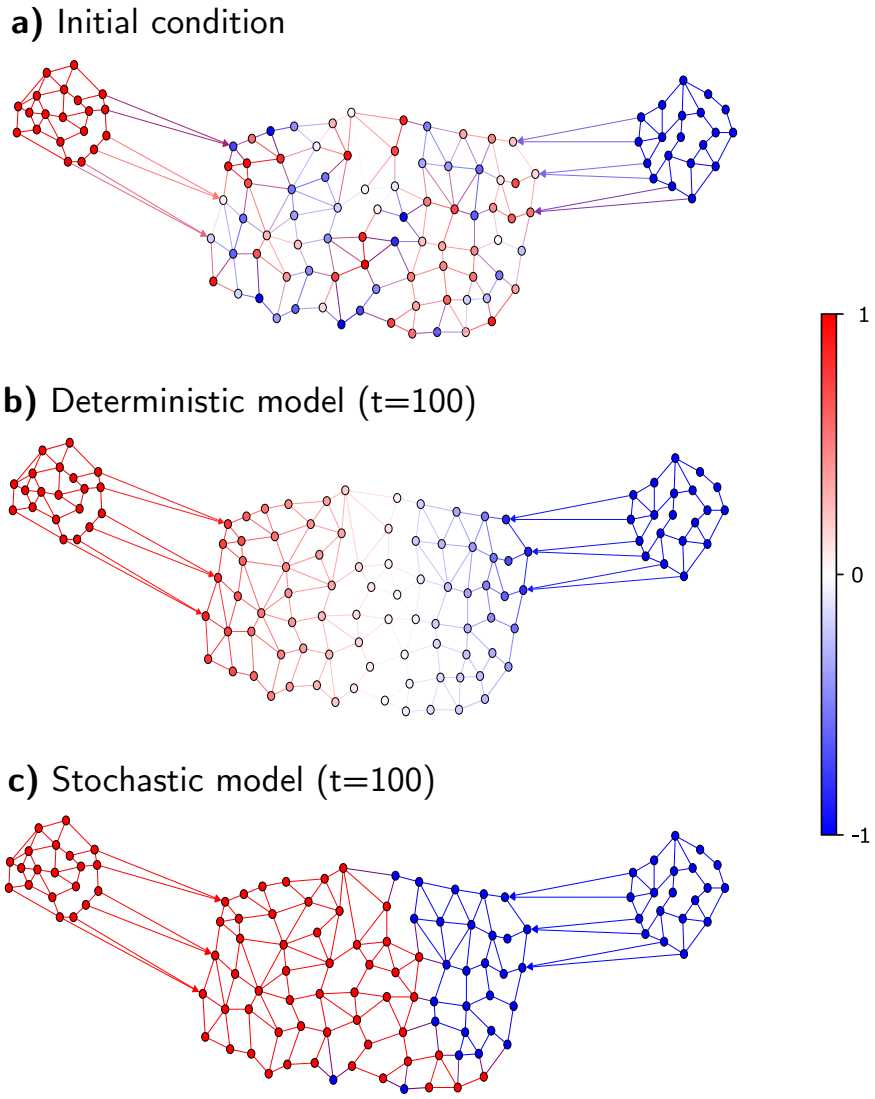

Figure 9: a) Initial state of the network in the battle scenario. b) The final state of the network in the battle scenario for the deterministic model. Nodes close to isolated blocks exhibit an opinion close to that of the isolated block. c) The state of the network at $t = 100$ for the stochastic model.

graph distance of a node in an absorbing component from isolated blocks plays a role in determining the final opinion of that node.

We now examine the analogous situation for the stochastic model using the same initial condition. We have already seen in the theory of the model that the agents in the central component cannot converge to a limiting opinion distribution as there are multiple isolated

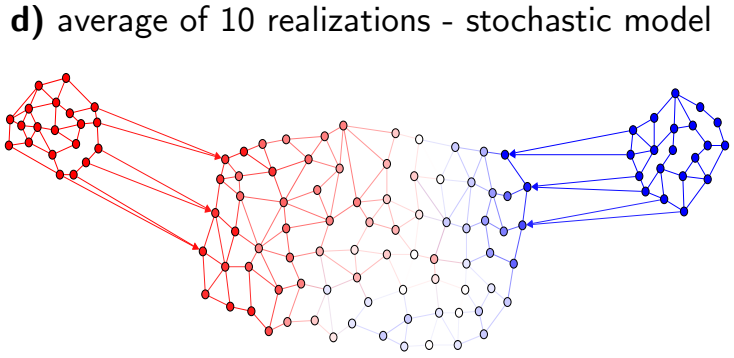

Figure 10: d) The average of 10 realizations of the stochastic model at $t = 100$. Notice that analogous to the deterministic case, that nodes close to an isolated block have an average opinion close to the opinion of the isolated block.

blocks in the network. However, to get an idea of how the network evolves in the stochastic case we show the state at $t = 100$.

Notice that similar to the deterministic case that it appears that nodes closer to an isolated block in the graph sense are likely to have the same opinion as that isolated block. However, this could very well be an artifact of the realization shown. To see that this phenomenon does indeed occur in the stochastic case we recall that we have shown that while the dynamics do not converge in this case, they are the same as the deterministic dynamics *in expected value*.

If we take the mean of 10 realizations at $t = 100$ we recover a gradient of opinions in the central component similar to what we found in the deterministic case. This indicates that indeed if a node is closer to an isolated block in the graph sense then it is more likely to have the opinion of that isolated block and serves to illustrate that the dynamics of the stochastic model coincide with the deterministic model in expected value.

Chapter 3

BOUNDED CONFIDENCE DYNAMICS AND GRAPH CONTROL: ENFORCING CONSENSUS

This chapter closely follows [73]. A generic feature of bounded confidence type models is the formation of clusters of agents. In this chapter we propose and study a variant of bounded confidence dynamics with the goal of inducing unconditional convergence to a consensus. The defining feature of these dynamics which we name the *No one left behind dynamics* is the introduction of a local control on the agents which preserves the connectivity of the interaction network. We rigorously demonstrate that these dynamics result in unconditional convergence to a consensus. The qualitative nature of our argument prevents us quantifying how fast a consensus emerges, however we present numerical evidence that sharp convergence rates would be challenging to obtain for such dynamics. Finally, we propose a relaxed version of the control. The dynamics that result maintain many of the qualitative features of the bounded confidence dynamics yet ultimately still converges to a consensus as the control still maintains connectivity of the interaction network.

3.1 Bounded confidence opinion dynamics

We will consider a collection of N agents where the opinion of the ith agent is denoted by $\mathbf{x}_i \in \mathbb{R}^d$. We will be concerned with a class of opinion dynamics that take the form:

$$\dot{\mathbf{x}}_i = \sum_{j=1}^{N} a_{ij}(\mathbf{x}_j - \mathbf{x}_i), \quad a_{ij} = \frac{\Phi_{ij}}{\sum_{k=1}^{N} \Phi_{ik}} \quad \text{with } \Phi_{ij} = \Phi(\left|\mathbf{x}_j - \mathbf{x}_i\right|). \tag{3.1}$$

Here, Φ represents the so-called *interaction function* and can be thought of as encoding how much influence one agent exerts over another, i.e. if $a_{ij} \neq 0$ then agent j is influencing

agent i with strength a_{ij}. The coefficients a_{ij} can be thought of as encoding the structure of a directed network on which the agents interact. As the a_{ij}'s are time dependent, the structure of the network changes in time as well. We will refer to this network as the *interaction network* and denote it by $G = (V, E)$ where V is the set of agents. Throughout the following we will assume that the interaction function has compact support on the interval $[0, 1]$, that it is positive on its support and has a minimum and maximum on this interval:

$$m = \min_{r\in[0,1]} \Phi(r) \quad , \quad M = \max_{r\in[0,1]} \Phi(r). \tag{3.2}$$

This assumption encodes that individuals will only interact and share ideas if their opinions are close enough to begin with. Notice that these conditions on the interaction function allow for a discontinuity at $x = 1$ from above in general; a prototypical example would be the indicator function on the interval $[0, 1]$, i.e. $\Phi(r) = \mathbb{1}_{[0,1]}(r)$.

This model represents a continuous version of the *bounded confidence* opinion dynamics introduced in [57]. Notice that in general $a_{ij} \neq a_{ji}$ and thus the dynamics are not symmetric (the center of mass is not preserved). However, if $a_{ij} > 0$ then we must also have $a_{ji} > 0$, in other words if j influences i then i must influence j.

Remark 9 Notice that since the interaction coefficients a_{ij} satisfy $\sum_{j=1}^{N} a_{ij} = 1$ we can rewrite the dynamics (3.1) as:

$$\dot{\mathbf{x}}_i = \overline{\mathbf{x}}_i - \mathbf{x}_i, \quad \overline{\mathbf{x}}_i = \sum_{j=1}^{N} a_{ij}\mathbf{x}_j, \tag{3.3}$$

so $\mathbf{x}_i$ moves towards $\overline{\mathbf{x}}_i$; the average opinion of all agents within the interaction radius of agent i weighted by their influence on agent i (see figure 11).

In this work we will be concerned with conditions that cause the opinions of all agents to converge to a consensus. We will see that due to the local consensus assumption, the notion of *connectivity* is crucial to the formation of a consensus.

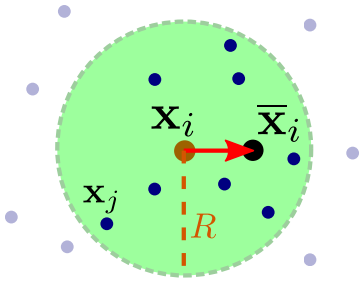

Figure 11: The movement of an agent according to the bounded confidence dynamics (3.3).

Definition 9 *We say that the configuration of agents* $\{\mathbf{x}_1, ..., \mathbf{x}_N\}$ *is* ***connected*** *if for any two agents* i *and* j *there exists a path between* i *and* j*; that is a subset* $\{i_1, ..., i_m\} \subseteq \{1, ..., N\}$ *such that* $i_1 = i$, $i_m = j$ *and:*

$$a_{i_k, i_{k+1}} \neq 0 \quad \textit{for all } k.$$

Clearly, connectedness of the initial configuration is a necessary condition for the emergence of a consensus. Due to the local consensus assumption, two agents who initially do not have a path between them will never become connected and therefore will not converge on the same opinion. However, connectedness of the initial condition is not sufficient for the emergence of a consensus as the dynamics do not necessarily preserve connectedness between two agents (see for instance figure 13-left); the dynamics (3.1) must be modified in some manner for connectedness of the initial condition to be sufficient for consensus.

3.2 No one left behind - enforcing consensus

3.2.1 Critical region

We now modify the dynamics (3.1) with the aim of preserving connectivity between agents. We will distinguish between the 1-dimensional and d-dimensional cases. The

intuitive idea is to introduce a control on the bounded confidence dynamics that causes an agent to alter its trajectory if it is close to disconnecting with a neighbor. With this aim in mind we introduce the notion of a *critical region* associated with each agent (see also an illustration in figure 12).

Definition 10 *Fix* $0 \leq r_* \leq 1$ *and let* $\{\mathbf{x}_1, ..., \mathbf{x}_N\} \subseteq \mathbb{R}^d$ *be a configuration of agents. The* ***critical region*** *associated with agent* i *is given by:*

$$\mathcal{B}_i = \{\mathbf{x} \in \mathbb{R}^d \mid 1 - r_* \leq |\mathbf{x} - \mathbf{x}_i| \leq 1 \text{ and } \langle \overline{\mathbf{x}}_i - \mathbf{x}_i, \mathbf{x} - \mathbf{x}_i \rangle \leq 0\}. \tag{3.4}$$

where $\overline{\mathbf{x}}_i$ *is given by* (3.3).

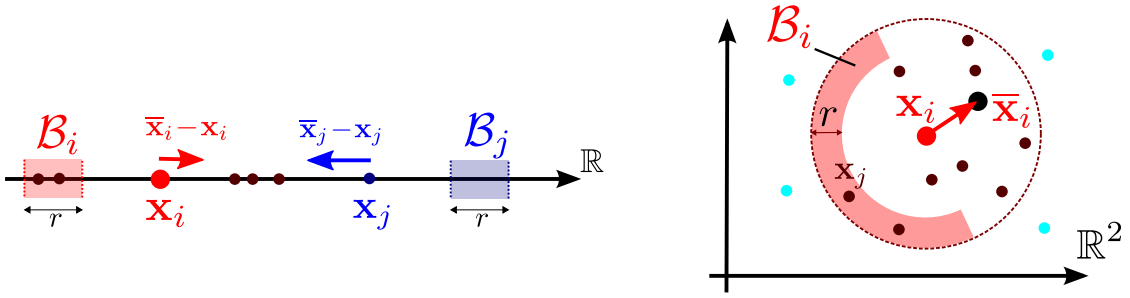

Figure 12: Illustration of the critical regions (3.4) in $\mathbb{R}$ (interval *behind* $\mathbf{x}_i$) and $\mathbb{R}^2$ (semi-annulus region). The opinion $\mathbf{x}_i$ is attracted toward the local average $\overline{\mathbf{x}}_i$, hence moves with velocity $\overline{\mathbf{x}}_i - \mathbf{x}_i$. In the "No-left behind dynamics" (1), $\mathbf{x}_i$ can only move only if there is no one in its critical region $\mathcal{B}_i$. Thus, $\mathbf{x}_i$ freezes whereas $\mathbf{x}_j$ is free to move in the left illustration.

3.2.2 No one left behind dynamics in $\mathbb{R}$

Notice that the critical region of any agent depends on the local average that the agent will move its opinion towards, (3.3); the critical region of an agent is always "behind" the agent in the sense that it is always in the opposite direction of the direction of movement of

the agent. The critical region is the main tool used to enforce connectivity preservation in the bounded confidence dynamics (3.1). We first illustrate the main idea in one dimension.

Model 1 (1D NOLB) *Consider a collection of agents with opinions* $\{ x_1, ..., x_N \}$ *in* $\mathbb{R}$. *The* ***1-D No one left behind*** *dynamics are given by:*

$$\dot{x}_i = \mu_i\big(\overline{x}_i - x_i\big), \quad \mu_i = \begin{cases} 0 & \textit{if there exists } x_j \in \mathcal{B}_i \\ 1 & \textit{otherwise} \end{cases} \tag{3.5}$$

where $\overline{x}_i$ *is the local average defined in* (3.3).

The scalar μ_i can be interpreted as a local control on the opinion of the ith agent. Under the dynamics given by Model 1 an agent evolves according to the normal bounded confidence model (3.1) unless there is another agent in its critical region in which case it does not move. In other words if the normal bounded confidence dynamics will cause an agent's opinion to change in such a way that it will become disconnected from one of its neighbors then it will stop moving and not leave its neighbor behind.

Remark 10 In addition to the network defined by the interactions among agents, the critical region allows one to impose an additional network structure on the collection of agents; there exists a link from agent i to agent j if agent j is in the critical region of agent i. We will refer to this network as the *behind graph* and denote it by $G^{\mathcal{B}} = (V, E^{\mathcal{B}})$ where V is the set of agents $\{ 1 \dots N \}$ and $E^{\mathcal{B}}$ is the set of edges:

$$(i, j) \in E^{\mathcal{B}} \text{ if } j \in \mathcal{B}_i. \tag{3.6}$$

In this notation we could write (1) as:

$$\dot{x}_i = \mu_i\big(\overline{x}_i - x_i\big), \quad \mu_i = \begin{cases} 0 & \text{if } (i, j) \in E_{\mathcal{B}} \\ 1 & \text{otherwise} \end{cases}$$

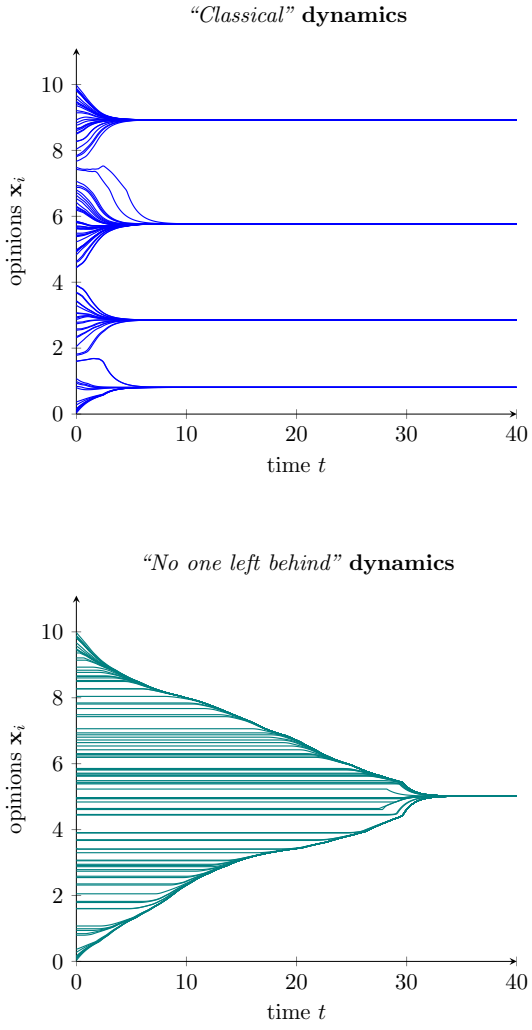

Figure 13: Simulation of the opinion dynamics without and with control (resp. top and bottom figure), e.g. solving resp. (3.3) and Model1 with $r_* = \frac{1}{2}$. With the control (right), the dynamics converge to a consensus.

Note that while the nature of bounded confidence dynamics forces the interaction network to be undirected, the behind graph must be directed as the presence of agent j in the behind region of agent i does *not* imply the opposite. Note also that if one denotes $G = (V, E)$ as the interaction graph (i.e. $(i, j) \in E$ if $a_{ij}, a_{ji} > 0$), then the behind graph $E^{\mathcal{B}}$ is a directed subgraph of E - see Figure 14.

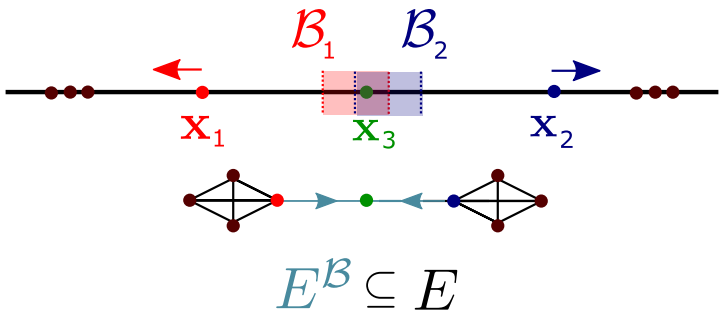

Figure 14: A configuration of agents (top) and the resulting interaction graph (edge set E, black) and behind graph (edge set $E^{\mathcal{B})}$, light blue). Note that the behind graph is a directed subgraph of the interaction graph.

We examine the effect that the augmented dynamics have on the long time behavior of the opinions in Figure 13 and note that for the same initial condition that normal bounded confidence dynamics do not result in a consensus but four "clusters" of opinions whereas the controlled dynamics preserve the connectivity of the agents and result in a consensus. Interestingly, we find that the dynamics introduced in Model (1) are not sufficient to ensure consensus in dimensions larger than one.

3.2.3 No one left behind dynamics in $\mathbb{R}^d$

Given that the dynamics defined in Model (1) are sufficient for convergence to a consensus for a connected initial configuration in one dimension (see Remark 12 and Theorem 2), one might expect that they should be sufficient for consensus in the d - dimensional case as well. Interestingly, this is not the case as the following example illustrates.

Example 1 We illustrate that the dynamics introduced in Model (1) do not ensure consensus in dimension 2 - this example can clearly be generalized to larger dimensions. Consider 6 *clusters* of opinions located on a regular hexagon with equal sides of length $d = 1 - \frac{r_*}{2}$ with $r_* \ll 1$. We denote the vertices of this regular hexagon as $\mathbf{x}_i$ with $i = 1, \ldots, 6$, and the

number of opinions in the cluster $\mathbf{x}_i$ as $\mathcal{N}(\mathbf{x}_i)$. Consider for instance, we have the following distribution of agents (see fig. 15):

$$\mathcal{N}(\mathbf{x}_1) = \mathcal{N}(\mathbf{x}_6) = 1, \quad \mathcal{N}(\mathbf{x}_2) = \mathcal{N}(\mathbf{x}_5) = 10, \quad \mathcal{N}(\mathbf{x}_3) = \mathcal{N}(\mathbf{x}_4) = 100.$$

In this setting, all the agents have another agent in their critical region. Thus, if one uses the same dynamics as in the one dimensional setting (1), we find $\mu_i = 0$ for all i and therefore the agents are stuck in this initial configuration. Thus, the "naive" control fails to achieve consensus.

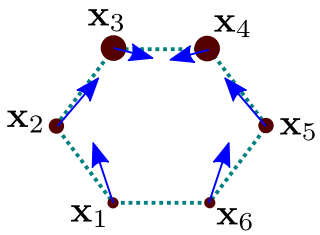

Figure 15: Counter-example in multi-dimension. Blue arrow is the velocity of each cluster. In this setting, every agent has someone in its critical region $\mathcal{B}_i$. Thus, the *naive* control in Model 1 would prevent anyone from moving.

Clearly, in the d-dimensional case, the condition that an agent may not move if another agent is in its critical region must be weakened in order to achieve a consensus. However, we still want to maintain the property that an agent may not move away (and possibly disconnect) from agents in its critical region as we know that connectivity is necessary for consensus. With this in mind we introduce the notion of *admissible velocity*.

Definition 11 *Let* $\{\mathbf{x}_1, ..., \mathbf{x}_N\} \subseteq \mathbb{R}^d$ *be a configuration of agents. The* ***cone of admissible velocity*** *associated with agent i is given by:*

$$C_i = \{\mathbf{v} \in \mathbb{R}^d \mid \langle \mathbf{v}, \mathbf{x}_j - \mathbf{x}_i \rangle \geq 0 \quad \textit{for all } \mathbf{x}_j \in \mathcal{B}_i\}. \tag{3.7}$$

where $\mathcal{B}_i$ *is the critical region* (3.4) *associated to agent i. If the critical region* $\mathcal{B}_i$ *is empty, then* $C_i = \mathbb{R}^d$.

Remark 11 We note that we can define the cone of admissible velocities in terms of the behind graph introduced in Remark 10. We just have to replace $\mathbf{x}_j \in \mathcal{B}_i$ by $(i, j) \in E^{\mathcal{B}}$ in the definition of C_i.

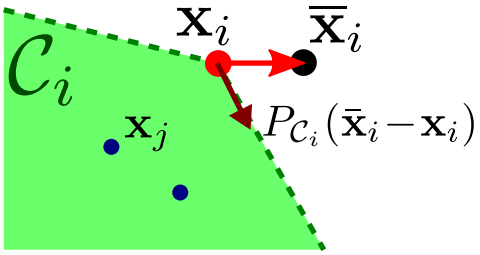

Figure 16: The velocity of agent i is the projection of the *desired* velocity $\overline{\mathbf{x}}_i - \mathbf{x}_i$ onto the cone of admissible velocity C_i.

We can now weaken the dynamics introduced in Model 1 by merely enforcing that the velocity of an agent belong to its cone of admissible velocity via a projection operator [60]. Intuitively, instead of forcing an agent to stop whenever its critical region is nonempty it can "take care" of agents in its critical region by moving closer to those agents *and* its local average (if possible).

Model 2 (NOLB) *Let* $\{\mathbf{x}_1, ..., \mathbf{x}_N\} \subseteq \mathbb{R}^d$ *be a configuration of agents. The* ***no one left behind (NOLB)*** *dynamics are given by:*

$$\mathbf{x}_i' = P_{C_i}\left(\overline{\mathbf{x}}_i - \mathbf{x}_i\right) \tag{3.8}$$

where $\overline{\mathbf{x}}_i$ *is the average velocity defined in* (3.3) *and* $P_{C_i} : \mathbb{R}^d \to C_i$ *is the projection operator associated to the cone of admissible velocities* C_i (3.7).

Remark 12 We note that the 1-D NOLB dynamics introduced in (1) are a special case of the general NOLB dynamics introduced in Model 2. Indeed, in one dimension the cone of admissible velocity is given by

$$C_i = \{\mathbf{v} \in \mathbb{R} \,|\, (\mathbf{v}, \mathbf{x}_j - \mathbf{x}_i) \geq 0 \quad \forall \mathbf{x}_j \in \mathcal{B}_i\}.$$

Now, if $\mathcal{B}_i$ is empty then we must have that $C_i = \mathbb{R}$ and therefore the projection operator P_{C_i} must be the identity, i.e

$$\mathbf{x}_i' = P_{C_i}\left(\overline{\mathbf{x}}_i - \mathbf{x}_i\right) = \overline{\mathbf{x}}_i - \mathbf{x}_i.$$

On the other hand, if there exists $x_j \in \mathcal{B}_i$ (and assuming without loss of generality that $\overline{\mathbf{x}}_i \geq \mathbf{x}_i$) we must have that $\mathbf{x}_j \leq \mathbf{x}_i$ which implies that:

$$C_i = \{\, \mathbf{v} \in \mathbb{R} \mid \mathbf{v} \leq 0 \,\}.$$

Since $\overline{\mathbf{x}}_i - \mathbf{x}_i \geq 0$ we therefore must have that

$$\mathbf{x}_i' = P_{C_i}\left(\overline{\mathbf{x}}_i - \mathbf{x}_i\right) = 0.$$

We illustrate the dynamics (3.8) and its long term behavior in Figure 17. As in the 1D case, a consensus is reached after some time whereas the classical dynamics generate multiple clusters.

We now rigorously show that augmenting the bounded confidence dynamics in this manner is sufficient to ensure consensus in the case that the initial configuration of the agents is connected.

3.3 No one left behind: Convergence to a consensus

3.3.1 Preservation of connectivity

A unifying feature of the classical bounded confidence dynamics and the NOLB dynamics is that due to the local consensus assumption they both result in a configuration of agents $\{\, \mathbf{x}_i(t) \,\}$ *contracting* in space. More specifically let us denote the convex hull of a

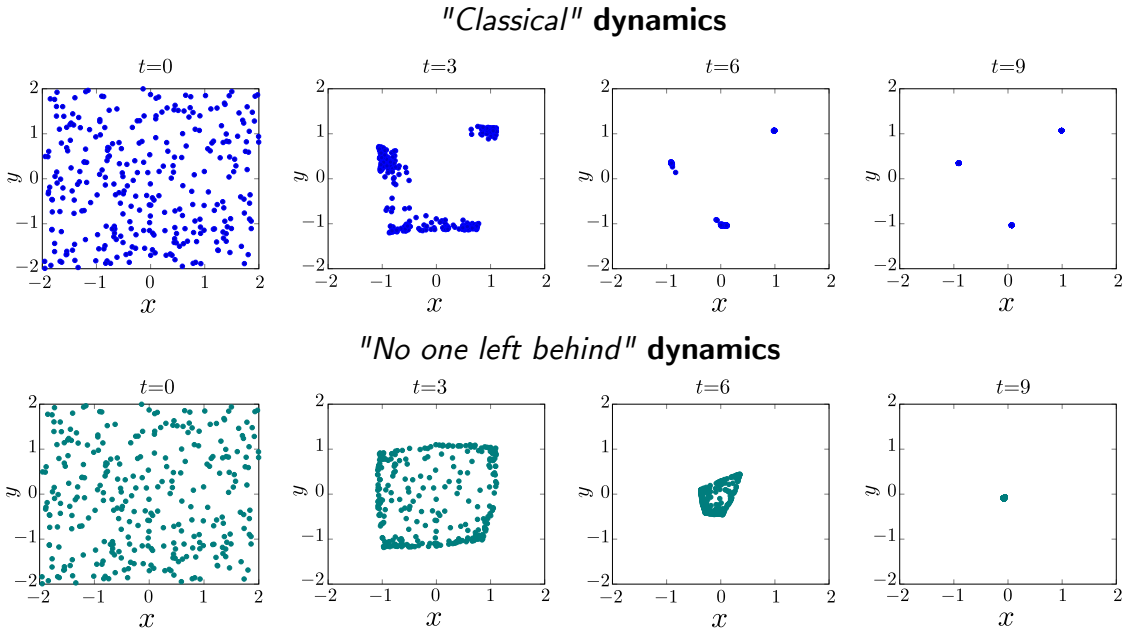

Figure 17: 2D simulation of opinion dynamics without and with control (resp, top and bottom figure), e.g. solving resp. (1) and (3.8) with $r_* = \frac{1}{2}$. With the control (bottom), the dynamics converge to a consensus.

configuration $\{\mathbf{x}_i(t)\}$ by $\Omega(t)$, i.e.

$$\Omega(t) = \text{Conv}\{\mathbf{x}_1(t), ..., \mathbf{x}_N(t)\}.$$

The agents contract in the following sense (see [63] for a proof).

Proposition 5 *If $\{x_i\}_i$ evolves according to the bounded confidence dynamics or the NOLB dynamics then the convex hull Ω satisfies:*

$$\Omega(t_1) \subset \Omega(t_0) \quad \textit{for any} \quad t_1 \geq t_0 \tag{3.9}$$

The contractive nature of the dynamics implies that for at least a subsequence of times the configuration approaches a limiting configuration thanks to Bolzano-Weierstrass theorem.

Corollary 4 *There exists a limiting configuration $\{\mathbf{x}_i^\infty\}_i$ and a sequence of times $(t_n)_n$ such that $t_n \to \infty$ and $\mathbf{x}_i(t_n) \to \mathbf{x}_i^\infty$ as $n \to \infty$.*

Another consequence is that the so-called diameter of the configuration $d(t)$ is decaying.

Corollary 5 *Denote the diameter* $d(t) = \max_{1\leq i,j\leq N} |\mathbf{x}_i(t) - \mathbf{x}_j(t)|$. *The diameter is non-increasing:*

$$d(t_2) \leq d(t_1) \quad \textit{for any } t_2 \geq t_1.$$

So in both the bounded confidence model and the NOLB model, agents remain close to each other in the sense of Proposition 5. However, we have already seen that the bounded confidence dynamics do not converge to a consensus from a connected initial condition; contractiveness alone is not sufficient for consensus.

The fundamental property of the NOLB dynamics (3.8) that distinguish them from the classical bounded confidence dynamics is that they rule out the possibility of agents who are connected later becoming disconnected as illustrated in Figures 13 and 17. The mechanism through which the local control accomplishes this can be seen in the following result.

Proposition 6 *Fix* $0 \leq r_* \leq 1$. *Suppose that at time t the opinions of agents i and j evolve according to* (3.8) *and satisfy:* $1 - r_* \leq |\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 1$ *(i.e. agent i and agent j are within the critical distance of each other), then:*

$$\frac{d}{dt}|\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 0. \tag{3.10}$$

Proof

As agent i and agent j are within the critical distance of each other we can assume that either agent i is in the critical region of agent j or vice versa, otherwise we would be finished. Without loss of generality assume that agent j is in the critical region of agent i. First notice that:

$$\frac{d}{dt}|\mathbf{x}_i - \mathbf{x}_j|^2 = -2\langle (P_{C_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i), \mathbf{x}_j - \mathbf{x}_i)\rangle - 2\langle (P_{C_j}(\overline{\mathbf{x}}_j - \mathbf{x}_j), \mathbf{x}_i - \mathbf{x}_j)\rangle.$$

Denote $\mathbf{v}_i = P_{C_i}(\bar{\mathbf{x}}_i - \mathbf{x}_i)$. Since j is in the critical region of agent i by assumption, it must satisfy $\langle \mathbf{v}_i, \mathbf{x}_j - \mathbf{x}_i\rangle \geq 0$, therefore $\langle P_{C_i}(\bar{\mathbf{x}}_i - \mathbf{x}_i), \mathbf{x}_j - \mathbf{x}_i\rangle \geq 0$. We can prove similarly that

$\langle P_{C_j}(\bar{\mathbf{x}}_j - \mathbf{x}_j), \mathbf{x}_i - \mathbf{x}_j \rangle \geq 0$ and therefore,

$$\frac{d}{dt}|\mathbf{x}_i - \mathbf{x}_j|^2 \ \leq \ 0.$$

■

The critical region acts as a "trap". If the distance between two particles is less than or equal to 1 but starts to increase they will eventually be within the critical distance of each other and their distance cannot increase any longer.

Corollary 6 *Suppose* $|\mathbf{x}_i(t_0) - \mathbf{x}_j(t_0)| \leq 1$. *Then*

$$|\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 1 \quad \textit{for all} \quad t \geq t_0.$$

Proof

Suppose for the sake of contradiction that there exists $T > t_0$ such that $|\mathbf{x}_i(T) - \mathbf{x}_j(T)| > 1$. Since the dynamics are continuous there must exist an exit time t_1 satisfying $|\mathbf{x}_i(t_1) - \mathbf{x}_j(t_1)| = 1$ and $|\mathbf{x}_i(t) - \mathbf{x}_j(t)| < 1$ for $t < t_1$. Thus there must exist some $\delta > 0$ such that $1 - r \leq |\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 1$ for $t_1 - \delta < t < t_1$. Therefore, by Proposition 6 $|\mathbf{x}_i(t) - \mathbf{x}_j(t)|$ is decaying on this time interval, a contradiction as $|\mathbf{x}_i(t_1) - \mathbf{x}_j(t_1)| = 1$. ■

So, adding the control indeed prevents situations like the ones presented in Figures 13 and 17 and preserves connectivity.

Corollary 7 *The dynamics* (3.8) *preserve connectivity, i.e. if the configuration* $\{\mathbf{x}_i(t_0)\}_i$ *is connected, then* $\{\mathbf{x}_i(t)\}_i$ *will be connected for any* $t \geq t_0$.

We will now examine how the preservation of connectivity enforced by (3.8) leads to a consensus.

3.3.2 Emergence of a consensus

We will now see that under the NOLB dynamics defined in (3.8), connectedness of the initial condition is sufficient for convergence to a consensus. In fact, this demonstrates that connectedness of the initial condition is equivalent to emergence of a consensus as we have previously noted that it is at least necessary. We will examine the convergence in two cases. In the general multidimensional case we are prevented from employing traditional ODE methods due to the discontinuous nature of the flow of the NOLB dynamics. We find that the interplay between the contractive nature of the dynamics and the preservation of connectivity allows us to circumvent this difficulty and deduce that a connected initial configuration is sufficient for convergence to a consensus. However we aren't able to say anything in general about the rate at which the dynamics converge to a consensus. In the one dimensional case this information is available and we derive explicit rates of convergence.

3.3.2.1 Multi Dimensional Case

We first examine the general multi dimensional case. Before proceeding to the main results we note that in the special case that $r_* = 1$ that connectivity of the initial condition is *not* sufficient for convergence to a consensus; the large value of r_* prevents some agents who are connected from exerting influence over each other.

Example 2 For simplicity we examine the one dimensional case. Suppose that $r_* = 1$ and consider the initial configuration given by:

$$\mathbf{x}_1(0) = 1,\ \mathbf{x}_2(0) = 2,\ \mathbf{x}_3(0) = 3,\ \mathbf{x}_4(0) = 4.$$

Here, $\mathbf{x}_1$ will move towards $\mathbf{x}_2$ and $\mathbf{x}_4$ will move towards $\mathbf{x}_3$. However, since $r_* = 1$, $\mathbf{x}_2$ and $\mathbf{x}_3$ may never move towards each other despite their connection as $\mathbf{x}_1$ and $\mathbf{x}_4$ will

always be in their respective critical regions - see Figure 18. Additionally we note that this example provides an illustration that the NOLB dynamics are *not* continuous as the dynamics converge to a state that is not an equilibrium of the dynamics.

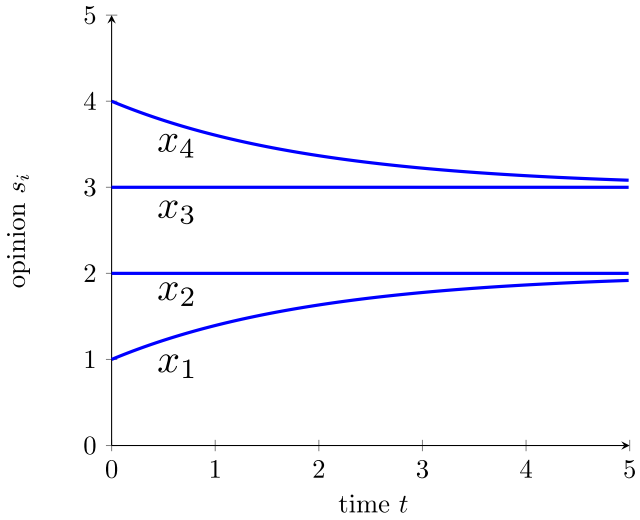

Figure 18: Preserving connectivity does not imply the convergence to a consensus. Here, when $r_* = 1$, the extreme points x_1 and x_4 will converge towards x_2 and x_3 respectively. However, x_2 and x_3 cannot move since x_1 and x_4 are always in their respective critical regions.

We will now show that the situation described in Example 2 is indeed a special case. As long as $0 < r_* < 1$ we will see that the control is sufficient to guarantee consensus given that the initial configuration is connected. Our main obstacle in this argument is discontinuities in the flow caused by the definition of the critical region and the discontinuity in the interaction function, Φ. Since this rules out many of the standard tools in ODE theory, our argument will rely on the interplay between the contractive nature of the dynamics and the fact that they preserve connectivity of the configuration of agents. Our strategy to show that the NOLB dynamics result in a consensus will be to show that the limiting configuration provided by Corollary 4, $\mathbf{x}^\infty$, is a consensus. In general this isn't sufficient to conclude

that $\mathbf{x}(t)$ even converges, much less to a consensus. However, the contractive nature of the dynamics allows us to say more.

We will proceed by contradiction and show that if $\mathbf{x}^\infty$ is *not* a consensus then it is possible to find a term in the sequence $\{\,\mathbf{x}(t_n)\,\}$ provided by Corollary 4 that when taken as initial condition of the dynamics results in points of Ω^∞ being outside of the convex hull of the configuration at a finite time - a contradiction of Proposition 5. We now prove the main result.

Theorem 5 *Assume $r_* < 1$ and that $\{\,\mathbf{x}_i(0)\,\}_i$ is connected. Then if $\{\,\mathbf{x}_i(t)\,\}_i$ evolves according to the NOLB dynamics* (3.8) *it will converge to a consensus.*

Proof

If $\{\,\mathbf{x}_i\,\}_i$ evolves according to the NOLB dynamics then by Corollary 4 (contractiveness) we know there exists a limiting configuration $\mathbf{X}^\infty = \{\,\mathbf{x}_i^\infty\,\}_i$ and a sequence of times $(t_n)_n$ such that $t_n \to \infty$ and $\mathbf{x}_i(t_n) \to \mathbf{x}_i^\infty$ as $n \to \infty$. Note that since by assumption $\mathbf{X}(0)$ is connected we must have by Corollary 6 (preservation of connectivity) that $\mathbf{X}^\infty$ is connected as well. We denote by p and q the extreme points such that:

$$\|\mathbf{x}_p^\infty - \mathbf{x}_q^\infty\| = \max_{ij} \|\mathbf{x}_i^\infty - \mathbf{x}_j^\infty\|.$$

Denote $\mathcal{N}_p^\infty$ as the set of neighbors of p. The main difficulty in the proof is to handle neighbors of $\mathbf{x}_p^\infty$ at a distance exactly 1. We call them *extreme* neighbors and denote them by $\mathcal{E}_p^\infty$:

$$\mathcal{N}_p^\infty = \{j \mid \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| \leq 1\} \tag{3.11}$$

$$\mathcal{E}_p^\infty = \{j \mid \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| = 1\}. \tag{3.12}$$

We are going to investigate 3 cases detailed in figure 19.

- **Case 1: no extreme neighbors $\mathcal{E}_p^\infty = \emptyset$.**

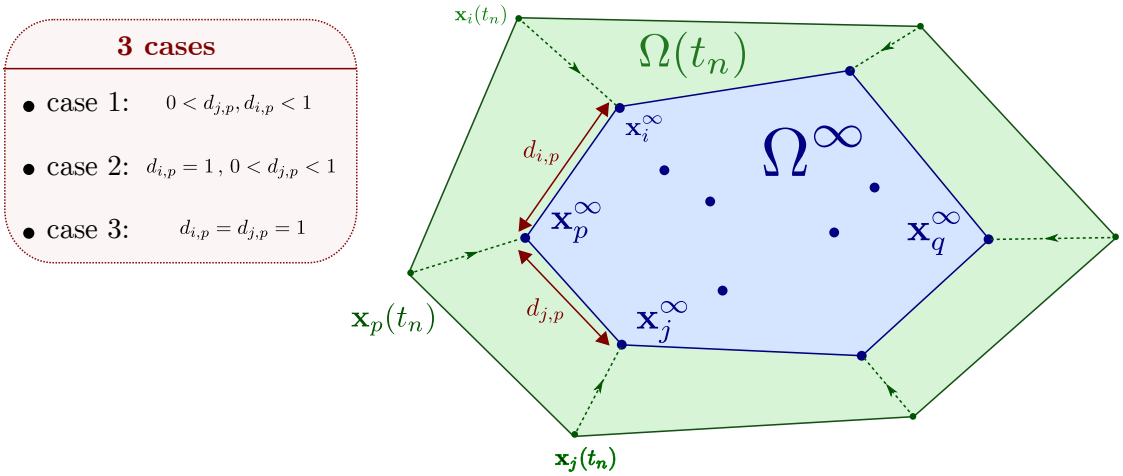

Figure 19: The convex hull $\Omega(t_n)$ has to converge to a limit configuration Ω^∞. The dynamics converge to a consensus if Ω^∞ is reduced to a single point which we prove by contradiction. We distinguish three cases of limit configuration Ω^∞ depending if the extreme point $\mathbf{x}_p^\infty$ has a so-called *extreme* neighbor j, i.e. $|\mathbf{x}_p^\infty - \mathbf{x}_j^\infty| = 1$.

If all the neighbors of $\mathbf{x}_p^\infty$ are at distance 0, then the dynamics converge to a consensus. Thus, we need to look at the case where there exists $j \in \mathcal{N}_p^\infty$ neighbor of p at distance $0 < \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| < 1$ (the distance 1 is excluded in Case 1).

Denote $\Omega(t)$ and Ω^∞ the convex hull of $\{\,\mathbf{x}_i(t)\,\}_i$ and $\{\,\mathbf{x}_i^\infty\,\}$ respectively. Since the dynamics is contracting (Proposition 5), we must have $\Omega^\infty \subseteq \Omega(t)$ for any t. Take now a supporting hyperplane that is tangent to Ω^∞ at the extreme point $\mathbf{x}_p^\infty$ (see figure 20). More specifically, we parametrize this supporting hyperplane using $\varphi_p : \mathbb{R}^d \to \mathbb{R}$ affine function

$$\varphi_p(\mathbf{x}) = \langle \mathbf{u}_p, \mathbf{x} \rangle + b$$

where the vector $\mathbf{u}_p$ and constant b are such that $\varphi_p(\mathbf{x}_p^\infty) = 0$ and $\varphi_p(\mathbf{x}) > 0$ for all $\mathbf{x}$ in Ω^∞ where $\mathbf{x} \neq \mathbf{x}_p^\infty$.

If there exists j such that $0 < \|\mathbf{x}_p^\infty - \mathbf{x}_j^\infty\| < 1$, then the coefficient a_{pj}^∞ satisfies:

$$a_{pj}^\infty = \frac{\Phi_{pj}^\infty}{\sum_{k=1}^N \Phi_{pk}^\infty} > 0.$$

Therefore, the local average $\overline{\mathbf{x}}_p^\infty = \sum_{k=1}^N a_{pk}^\infty \mathbf{x}_j^\infty$ is different from $\mathbf{x}_p^\infty$. Moreover since $\overline{\mathbf{x}}_p^\infty \in \Omega^\infty$, we deduce $\varphi_p(\overline{\mathbf{x}}_p^\infty) > 0$.

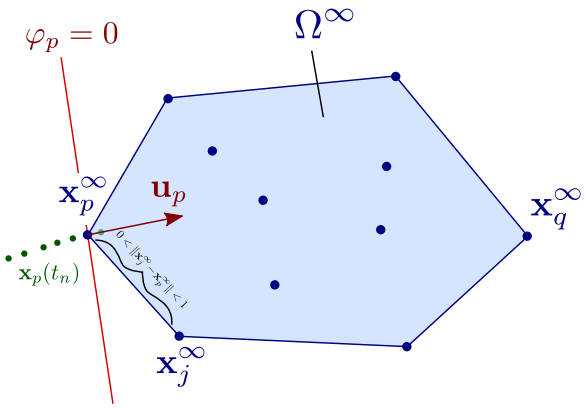

Figure 20: If the limit configuration $\{\mathbf{x}_k^\infty\}_k$ is not a consensus, the extreme point $\mathbf{x}_p(t_n)$ will eventually get inside the convex hull Ω^∞ which gives a contradiction.

We will get a contradiction if we can show that the sequence $\mathbf{x}_p(t_n)$ once closed to $\overline{\mathbf{x}}_p^\infty$ will *cross* the hyperplane $\{\varphi_p = 0\}$. Thus, we investigate the values of $\varphi_p(\mathbf{x}_p(t_n))$. Its time evolution is given by:

$$\frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] = \langle \mathbf{u}_p, P_{C_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)) \rangle. \tag{3.13}$$

To get rid of the projection operator, we notice that the projection is actually increasing the scalar product *if* $\mathbf{u}_p$ is in $C_p(t_n)$ thanks to the following lemma (the proof is in appendix).

Lemma 6 *Let* $\{\mathbf{x}_i\}_i \subseteq \mathbb{R}^d$ *and consider* $C = \{\mathbf{v} \mid \langle \mathbf{v}, \mathbf{x}_i \rangle \geq 0$ *for all* $i\}$. *If* $\mathbf{u}$ *is in* C *then:*

$$\langle P_C(\mathbf{x}), \mathbf{u} \rangle \geq \langle \mathbf{x}, \mathbf{u} \rangle \quad \textit{for all } \mathbf{x} \in \mathbb{R}^d. \tag{3.14}$$

It remains to show that $\mathbf{u}_p \in C_p(t_n)$. Notice that $\langle \mathbf{x} - \mathbf{x}_p^\infty, \mathbf{u}_p \rangle \geq 0$ for all $\mathbf{x}$ in Ω^∞ and therefore $\mathbf{u}_p \in C_{x_p^\infty}$. Eventually this is true for the approximating sequence $\mathbf{x}_p(t_n)$ as well: $\mathbf{u}_p \in C_{x_{p^*(t_n)}}$ for t_n large enough. Indeed, take any $k \neq p$. There are two cases. First case, $\mathbf{x}_k(t_n) \to \mathbf{x}_p^\infty$

and therefore $\|\mathbf{x}_k(t_n) - \mathbf{x}_p(t_n)\| \overset{n\to+\infty}{\longrightarrow} 0$, meaning that $\mathbf{x}_k(t_n)$ will not be in the critical region of agent p, i.e. $k \notin \mathcal{B}_{\mathbf{x}_p(t_n)}$ ($r_* < 1$). In the second case, $\mathbf{x}_k(t_n) \to \mathbf{x}_k^\infty \neq \mathbf{x}_p^\infty$. Then

$$\begin{aligned}\langle \mathbf{u}_p, \mathbf{x}_k(t_n) - \mathbf{x}_p(t_n)\rangle &= \langle \mathbf{u}_p, \mathbf{x}_k(t_n) + \mathbf{x}_k^\infty - \mathbf{x}_p^\infty + \mathbf{x}_p^\infty - \mathbf{x}_k^\infty - \mathbf{x}_{p^*}(t_n)\rangle \\ &= \langle \mathbf{u}_p, \mathbf{x}_k(t_n) - \mathbf{x}_k^\infty\rangle + \langle \mathbf{u}_p, \mathbf{x}_p^\infty - \mathbf{x}_{p^*}(t_n)\rangle + \langle \mathbf{u}_p, \mathbf{x}_k^\infty - \mathbf{x}_p^\infty\rangle \\ &\geq -2\Big\|\mathbf{u}_p\Big\|\delta_n + \langle \mathbf{x}_k^\infty - \mathbf{x}_p^\infty, \mathbf{u}_p\rangle\end{aligned}$$

using Cauchy-Schwarz with $\delta_n = \max(\|\mathbf{x}_k(t_n) - \mathbf{x}_k^\infty\|, \|\mathbf{x}_p(t_n) - \mathbf{x}_p^\infty\|)$. Since $\delta_n \overset{n\to+\infty}{\longrightarrow} 0$ and $\langle \mathbf{x}_k^\infty - \mathbf{x}_p^\infty, \mathbf{u}_p\rangle > 0$, we conclude that: $\langle \mathbf{u}_p, \mathbf{x}_k(t_n) - \mathbf{x}_p(t_n)\rangle > 0$ for t_n large enough, i.e. $\mathbf{u}_p \in C_p(t_n)$.

We can now continue our computation (3.13):

$$\begin{aligned}\frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] &= \langle \mathbf{u}_p, P_{C_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n))\rangle \\ &\geq \langle \mathbf{u}_p, \overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)\rangle \\ &= \varphi_p(\overline{\mathbf{x}}_p(t_n) - \varphi_p(\mathbf{x}_p(t_n)) \quad \to \quad \varphi_p(\overline{\mathbf{x}}_p^\infty) - \varphi_p(\mathbf{x}_p^\infty) > 0\end{aligned} \tag{3.15}$$

by continuity of φ_p. Indeed, $\overline{\mathbf{x}}_p(t_n) \overset{n\to+\infty}{\longrightarrow} \to \overline{\mathbf{x}}_p^\infty$ since Φ has only a discontinuity at distance 1 and the case 1 assumption avoids this possibility for any coefficients a_{pk}.

We deduce a contradiction since we have two conflicting properties:

$$i) \quad \varphi_p(\mathbf{x}_p(t_n)) \to \varphi_p(\mathbf{x}_p^\infty) = 0 \quad \text{since } \mathbf{x}_p(t_n) \to \mathbf{x}_p^\infty, \tag{3.16}$$

$$ii) \quad \frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] \geq \varphi_p(\overline{\mathbf{x}}_p^\infty) > 0 \text{ for large } t_n. \tag{3.17}$$

• **Case 2: there exists an extreme neighbor (i.e. $\mathcal{E}_p^\infty \neq \emptyset$) AND there exists j such that $0 < \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| < 1$.**

In this situation, agent p might connects with a neighbor k only at infinity and thus the local average $\overline{\mathbf{x}}_p(t_n)$ might not converge to $\overline{\mathbf{x}}_p^\infty$ as $t_n \to +\infty$. But thanks to the non-extreme neighbor j, we are going to have a contradiction as in Case 1.

The coefficient a_{pj} is lower bounded at infinity since:

$$\liminf_{n\to\infty} a_{pj}(t_n) \geq \liminf_{n\to\infty} \frac{\Phi(\|\mathbf{x}_p(t_n) - \mathbf{x}_j(t_n)\|)}{N \cdot M} \geq \frac{m}{N \cdot M} > 0$$

where m and M are respectively the minimum and maximum of Φ on the interval $[0, 1]$. This is enough to show that, as in case 1, the derivative $\frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))]$ is bounded below by a positive constant (for large t_n) leading to a contradiction. Indeed,

$$\begin{aligned}\liminf_{t_n\to+\infty} \langle \mathbf{u}_p, \overline{\mathbf{x}}_p(t_n)\rangle &= \liminf_{t_n\to+\infty} \left(\sum_{k=1..N, k\neq j} a_{pk}\langle \mathbf{u}_p, \mathbf{x}_k(t_n)\rangle \;+\; a_{pj}\langle \mathbf{u}_p, \mathbf{x}_j(t_n)\rangle \right) \\ &\geq 0 + \frac{m}{N \cdot M}\langle \mathbf{u}_p, \mathbf{x}_j^\infty\rangle > 0.\end{aligned}$$

Thus, we conclude that:

$$\liminf_{n\to\infty} \left(\varphi_p(\overline{\mathbf{x}}_p(t_n)) - \varphi_p(\mathbf{x}_p(t_n))\right) \geq c > 0$$

and we may apply the same argument as in Case 1.

- **Case 3: there exists an extreme neighbor (i.e. $\mathcal{E}_p^\infty \neq \emptyset$) AND for all $j \in \mathcal{N}_p^\infty$, $\|\mathbf{x}_p^\infty - \mathbf{x}_j^\infty\| = 0$ or $\|\mathbf{x}_p^\infty - \mathbf{x}_j^\infty\| = 1$.**

This is the most delicate case since the neighbors at distance *exactly* one might appear only asymptotically (i.e. at "$t = \infty$"). However, the assumption is also helping: a neighbor j of p *must* converge to p as we will see. More precisely, by the assumption of Case 3, any neighbor j of $\mathbf{x}_p(t_n)$ must satisfy one of the two following scenario:

1. $\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\| \overset{n\to+\infty}{\longrightarrow} 0$
2. $\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\| = 1$ for all n

Indeed, if there exists a time t_n such that $\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\| < 1$, then it is impossible that $\|\mathbf{x}_j(t) - \mathbf{x}_p(t)\| = 1$ at a later time $t > t_n$ (Proposition 6). Since the limit $\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\|$ cannot be in $(0, 1)$ due to the assumption of Case 3, it must converge to zero.

To prove that a consensus emerges, we have to rule out scenario 2: neighbors of $\mathbf{x}_p(t_n)$ cannot stay at a distance exactly 1. Notice that this is precisely what is happening in the counter-example of figure 18: $\|\mathbf{x}_3(t) - \mathbf{x}_2(t)\| = 1$ for all t. But in this counter-example $r_* = 1$ which is not the case in the present scenario.

Let's proceed once again by contradiction and assume that there exists j such that $\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\| = 1$ for all t_n. Denote $\mathcal{E}_p$ all the neighbors j of p satisfying this property. Under such assumption, we must have: $\frac{d}{dt}\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\|^2 = 0$ and therefore:

$$\langle P_{C_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)), \mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\rangle = 0. \tag{3.18}$$

A key observation is to notice that the desired velocity $\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)$ converges to the average over the neighbors in $\mathcal{E}_p$. Indeed, we rewrite:

$$\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n) = \sum_{j=1}^{N} a_{pj}(t_n)\Big(\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\Big).$$

If $j \notin \mathcal{E}_p$, either $a_{pj}(t_n) = 0$ (j not a neighbor of p) or $\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n) \overset{n\to+\infty}{\longrightarrow} 0$. Thus,

$$\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n) \overset{n\to+\infty}{\longrightarrow} \sum_{j\in\mathcal{E}_p} a^\infty_{pj}\Big(\mathbf{x}^\infty_j - \mathbf{x}^\infty_p\Big). \tag{3.19}$$

Moreover, $a^\infty_{pj} = c > 0$ for all $j \in \mathcal{E}_p$ since $\Phi(\|\mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\|) = \Phi(1) > 0$.

We can now pass to the limit in (3.18):

$$\langle P_{C^\infty_p}\Big(\sum_{k\in\mathcal{E}_p} c(\mathbf{x}^\infty_k - \mathbf{x}^\infty_p)\Big), \mathbf{x}^\infty_j - \mathbf{x}^\infty_p\rangle = 0, \tag{3.20}$$

where C^∞_p is the critical region defined by:

$$C^\infty_p = \{\mathbf{v} \in \mathbb{R}^d \mid (\mathbf{v}, \mathbf{x}^\infty_j - \mathbf{x}^\infty_p) \geq 0 \quad \forall j \in \mathcal{E}_p\}. \tag{3.21}$$

Notice that the definition of C^∞_p only includes extreme neighbors (i.e. $j \in \mathcal{E}_p$). Indeed, the other neighbor $\mathbf{x}_k$ converges to $\mathbf{x}_p$ as $t_n \to +\infty$. Since $r_* < 1$, $\mathbf{x}_k(t_n)$ is no longer in the critical region $\mathcal{B}_{\mathbf{x}_p(t_n)}$ for t_n large enough.

Summing the previous expression over the neighbors j in $\mathcal{E}_p$ gives:

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{v}\rangle = 0 \qquad \text{with} \quad \mathbf{v} = \sum_{j\in\mathcal{E}_p}(\mathbf{x}_j^\infty - \mathbf{x}_p^\infty). \tag{3.22}$$

Since $C_{C_p^\infty}$ is a convex cone, we deduce that:

$$\langle P_{C_p^\infty}(\mathbf{v}), P_{C_p^\infty}(\mathbf{v})\rangle = \langle P_{C_p^\infty}(\mathbf{v}), \mathbf{v}\rangle, \tag{3.23}$$

and therefore $P_{C_p^\infty}(\mathbf{v}) = 0$.

To get a contradiction, we use once again $\mathbf{u}_p$ define from the supporting hyperplane at $\mathbf{x}_p^\infty$ (see Figure 20), we have:

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{u}\rangle \geq \langle \mathbf{v}, \mathbf{u}\rangle \tag{3.24}$$

since $\mathbf{u}$ is in the cone C_p^∞. We deduce:

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{u}\rangle \geq \sum_{j\in\mathcal{E}_p}\langle \mathbf{x}_j^\infty, \mathbf{u}\rangle > 0, \tag{3.25}$$

since $\mathbf{x}_j^\infty$ is strictly inside the convex hull Ω^∞ (p being an extreme point). This proves that $P_{C_{\mathbf{x}_p^\infty}}(\mathbf{v}) \neq 0$ and concludes the proof. ■

3.3.2.2 1-D Case

We now investigate consensus in the special case that the dynamics are one dimensional. We know from the previous section that consensus occurs however in this case we will also be able to quantify the rate at which this consensus emerges. Our main tool will be estimates on the diameter $d(t)$. Recall from Remark 12 that in one dimension Model (2) reduces to Model (1) - we will use the notation in the latter in the following. We will see that the control μ_i causes the convergence to occur in two stages. Clearly if a consensus emerges there must exist a time τ after which all agents are directly interacting with each

other, that is $|x_i(t) - x_j(t)| \leq 1$ for any i and j and $t > \tau$; in this case the network on which agents interact is *fully connected*. Before this time there necessarily exist pairs of agents who do not interact but are merely connected by a path. We will first examine the case where all agents are interacting - in this case the dynamics converge towards a consensus at an exponential rate that depends on the extreme values of the interaction function.

Proposition 7 *Suppose* $d(0) \leq 1$. *Then:*

$$d(t) \leq d(0)e^{-\frac{m}{M} \cdot t} \tag{3.26}$$

where $m = \min_{x \in [0,1]} \Phi(x)$ *and* $M = \max_{x \in [0,1]} \Phi(x)$.

Proof

Fix t and denote p and q such that $d(t) = |x_p - x_q|$. Notice that since p and q are the two agents with *extreme* opinions we must have that $\mu_p = \mu_q = 1$ as they cannot have any agents in their critical regions. We aim to get a bound on $(d^2)'$ in terms of d^2 in order to apply Gronwall's lemma. By the Cauchy-Schwarz inequality we have that:

$$\begin{aligned} \frac{d}{dt}[d^2(t)] = 2(\dot{x}_p - \dot{x}_q, x_p - x_q) &= 2\Big((\overline{x}_p - \overline{x}_q)(x_p - x_q) - (x_p - x_q)^2\Big) \\ &\leq 2\Big(|\overline{x}_p - \overline{x}_q||x_p - x_q| - |x_p - x_q|^2\Big). \end{aligned} \tag{3.28}$$

To obtain the bound we desire all that remains is to bound $|\overline{x}_p - \overline{x}_q|$ above by a constant multiple of $|x_p - x_q|$. We will exploit the fact that the local averages $\overline{x}_p$ and $\overline{x}_q$ must be inside the convex hull of opinions and therefore since agents p and q are the agents with the most extreme opinions the difference between their local averages must be smaller than the difference between their opinions. Denote $\eta_i = \min(a_{pi}, a_{qi})$ and notice that since Φ is bounded and $d(0) \leq 1$ we must have that $\eta_i \geq \frac{m}{MN}$ where m and M are given by (3.2). Notice

that:

$$\begin{aligned}\left|\overline{x}_p - \overline{x}_q\right| &= \left|\sum_{i=1}^{N} a_{pi}x_i - \sum_{i=1}^{N} a_{qi}x_i\right| = \left|\sum_{i=1}^{N}(a_{pi}-\eta_i)x_i - \sum_{i=1}^{N}(a_{qi}-\eta_i)x_i\right| \\ &=: \left|\sum_{i=1}^{N}\widetilde{a}_{pi}x_i - \sum_{i=1}^{N}\widetilde{a}_{qi}x_i\right|.\end{aligned} \tag{3.30}$$

Let $\boldsymbol{\eta} = \sum_{i=1}^{N}\eta_i$. We have by (3.30) that:

$$\left|\overline{x}_p - \overline{x}_q\right| = (1-\boldsymbol{\eta})\left|\sum_{i=1}^{N}\frac{\widetilde{a}_{pi}x_i}{1-\boldsymbol{\eta}} - \sum_{i=1}^{N}\frac{\widetilde{a}_{qi}x_i}{1-\boldsymbol{\eta}}\right| := (1-\boldsymbol{\eta})\left|\widetilde{x}_p - \widetilde{x}_q\right| \tag{3.31}$$

Notice that

$$\sum_{i=1}^{N}\frac{\widetilde{a}_{qi}}{1-\boldsymbol{\eta}} = \sum_{i=1}^{N}\frac{\widetilde{a}_{pi}}{1-\boldsymbol{\eta}} = 1, \tag{3.32}$$

and therefore we must have that $\widetilde{x}_p$ and $\widetilde{x}_q$ are in the convex hull of $\{x_i\}_{i=1}^{N}$ so necessarily we have that $|\widetilde{x}_p - \widetilde{x}_q| \leq |\overline{x}_p - \overline{x}_q|$. Therefore by (3.31) we have that $|\overline{x}_p - \overline{x}_q| \leq (1-\boldsymbol{\eta})|x_p - x_q|$ which by (3.28) implies:

$$\begin{aligned}\frac{d}{dt}[d^2(t)] &\leq 2\Big((1-\boldsymbol{\eta})|x_p - x_q||x_p - x_q| - |x_p - x_q|^2\Big) \\ &= -2\boldsymbol{\eta}|x_p - x_q|^2.\end{aligned} \tag{3.34}$$

Finally, since by definition of the weights a_{iq}, a_{ip} and $\boldsymbol{\eta}$ we have that $\boldsymbol{\eta} \geq \frac{m}{M}$ we can conclude using (3.34) that:

$$\frac{d}{dt}[d^2(t)] = 2d(t)d'(t) \leq -\frac{m}{M}d^2(t). \tag{3.35}$$

An application of Gronwall's lemma provides the final result. ■

So, in the case that all agents are interacting, i.e. that $|x_i(t) - x_j(t)| \leq 1$ for any i and j, a consensus is reached exponentially fast at a rate that depends on the maximum and minimum values of the interaction function. However, starting from an initial condition that is connected does not mean that all agents are directly interacting; any two agents are merely connected by a path. We now examine the rate of convergence for $t < \tau$, i.e. *before* all agents are interacting.

Theorem 6 *Suppose the initial condition* $\{x_i(0)\}_i$ *is connected and let* $\eta = \frac{m}{M \cdot N}$. *There exist* $\delta > 0$ *and* $T > 0$ *such that while* $d(t) \geq 1$, *we have:*

$$d(t) \leq d(0) + \eta\delta(N \cdot T - t). \tag{3.36}$$

Thus, after $t \geq N \cdot T + \frac{d_0 - 1}{\eta\delta}$, *the diameter is converging exponentially fast toward zero.*

Proof

Denote p and q such that $d(t) = x_q(t) - x_p(t)$. Suppose $d(t) > 1$, the case $d(t) \leq 1$ has been treated in proposition 7. We analyze the behavior of p and q separately as they do not affect each other. We fix the two constant $\delta > 0$ and $T > 0$ satisfying the two technical conditions:

$$\delta + 2T \leq \min(r_*, 1 - r_*) \tag{3.37}$$

$$\frac{T^2}{2N^2}\eta\Big(\eta(1 - \delta - 2T) - 2N(\delta + 2T)\Big) \geq \delta. \tag{3.38}$$

It is always possible to find such constant (see lemma 11).

We split the study of $\dot{x}_p$ at a given time t_* in two cases.

• **Case 1**: suppose there exists p_1 such that $\delta \leq |x_p(t_*) - x_{p_1}(t_*)| \leq 1$.

In this case, we can easily deduce a lower bound for the speed of p:

$$\dot{x}_p = \sum_{j=1}^{N}(x_j - x_p)a_{jp} \geq \ (x_{p_1} - x_p)a_{p,p_1} \geq \delta\eta$$

with $\eta = \frac{m}{M \cdot N}$. Thus, since we have on the other end $\dot{x}_q \leq 0$, we deduce that the diameter is decaying at a minimum speed $\delta\eta$.

• **Case 2**: $|x_p(t_*) - x_i(t_*)| < \delta$ or $|x_p(t_*) - x_i(t_*)| > 1$ for any $i \neq p$.

In other words, all the neighbors of p are at a distance less than δ. We cannot find a lower bound for $\dot{x}_p$ anymore. Instead, we show that a *neighbor of a neighbor* (denoted p_2) will become a new neighbor of p in a finite time less than T.

Since connectivity is preserved, there exists p_1 and p_2 such that: $p \sim p_1$, $p_1 \sim p_2$ and

$p \nsim p_2$. Therefore, we deduce (see figure)

$$|x_p(t_*) - x_{p_1}(t_*)| \leq \delta, \quad |x_{p_1}(t_*) - x_{p_2}(t_*)| \leq 1, \quad |x_p(t_*) - x_{p_2}(t_*)| > 1.$$

By triangular inequality, we deduce that $|x_{p_1}(t_*) - x_{p_2}(t_*)| \geq 1 - \delta$. Let us show that after time T, we have $|x_p(t_* + T) - x_{p_2}(t_* + T)| \leq 1$.

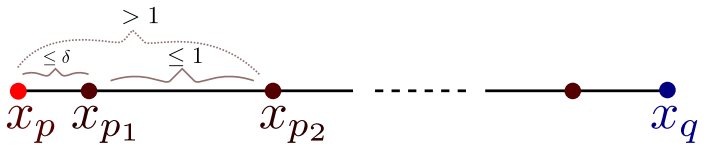

Figure 21: Situation in the case 2. The extreme point x_p needs x_{p_2} the neighbor of its neighbor x_{p_1} to be pushed further to the right.

First, we show that $\dot{x}_{p_2} \leq 0$ thanks to x_{p_1}. During the time interval $t \leq T$, as the velocity $|\dot{x}_i| \leq 1$ for any i, we have:

$$|x_{p_1}(t_* + t) - x_{p_2}(t_* + t)| \geq |x_{p_1}(t_*) - x_{p_2}(t_*)| - 2 \cdot t \geq 1 - \delta - 2T \geq 1 - r_*,$$

by the assumption (3.37). As the consequence, x_{p_1} is always in the critical region of the agent p_2. Thus, x_{p_2} can only move left, which implies $\dot{x}_{p_2} \leq 0$.

Second, we show that x_p increases by at least δ during the period T which will imply that $p \sim p_2$. The main idea is to show that x_{p_1} is going to *pull out* x_p. To prove it, we compute

$$\begin{aligned} \dot{x}_p &= \mu_p(\bar{x}_p - x_p) = 1 \cdot \sum_k a_{p,k}(x_k - x_p) \\ &= a_{p,p_1}(x_{p_1} - x_p) + \sum_{k \neq p_1} a_{p,k}(x_k - x_p) \quad \geq \quad \frac{\eta}{N}(x_{p_1} - x_p) + 0, \end{aligned}$$

since $x_k - x_p \geq 0$. We now need to find a lower bound for $x_{p_1} - x_p$. With this aim, we compute the time derivative:

$$\begin{aligned} \dot{x}_{p_1} - \dot{x}_p &= \mu_{p_1}(\bar{x}_{p_1} - x_{p_1}) - \mu_p(\bar{x}_p - x_p) \\ &\geq (\bar{x}_{p_1} - x_{p_1}) - (\delta + T). \end{aligned}$$

Here, we use that $\dot{x}_{p_1}$ is necessarily positive as p_2 is in its critical region, thus we have a lower bound by replacing μ_{p_1} by 1. We now also suppose that x_p has not connected with p_2 (otherwise there is no need to go further) and therefore its neighbors are at a distance bounded by $\delta + T$ on the time interval $[t_*, t_* + T]$.

Following the same inequality as for $\dot{x}_p$, we deduce that:

$$\begin{aligned}\bar{x}_{p_1} - x_{p_1} &= a_{p_1,p_2}(x_{p_2} - x_{p_1}) + \sum_{k \neq p_2} a_{p_1,k}(x_k - x_{p_1}) \\ &\geq \frac{\eta}{N}(1 - \delta - 2T) + (x_p - x_{p_1}) \geq \frac{\eta}{N}(1 - \delta - 2T) - (\delta + T).\end{aligned}$$

Therefore,

$$\dot{x}_{p_1} - \dot{x}_p \geq \frac{\eta}{N}(1 - \delta - 2T) - 2(\delta + T).$$

We conclude that:

$$\begin{aligned}x_{p_1}(t_* + t) - x_p(t_* + t) &\geq x_{p_1}(t_*) - x_p(t_*) + \int_0^t \frac{\eta}{N}(1 - \delta - 2T) - 2(\delta + T)\,\mathrm{d}t \\ &\geq 0 + \left(\frac{\eta}{N}(1 - \delta - 2T) - 2(\delta + T)\right)t.\end{aligned}$$

Coming back to x_p, we obtain:

$$\begin{aligned}x_p(t_* + T) - x_p(t_*) &\geq \int_0^T \frac{\eta}{N}(x_{p_1}(t_* + t) - x_p(t_* + t))\,\mathrm{d}t \\ &\geq \frac{\eta}{N}\left(\frac{\eta}{N}(1 - \delta - 2T) - 2(\delta + T)\right)\frac{T^2}{2} \geq \delta,\end{aligned}$$

using (3.38). Therefore, at time $t_* + T$, we have $x_{p_2}(t_* + T) - x_p(t_* + T) \leq 1$, and thus $p \sim p_2$.

To conclude, since there is only a finite number of particles N, situations as in *case 2* can only appear a finite number of time (less than N times) and thus spread over a period less than $N \cdot T$. Thus, outside these periods, the decay of $d(t)$ satisfies $\dot{d} \leq -\eta\delta$ leading to the upper-bound (3.36) which concludes the proof. ■

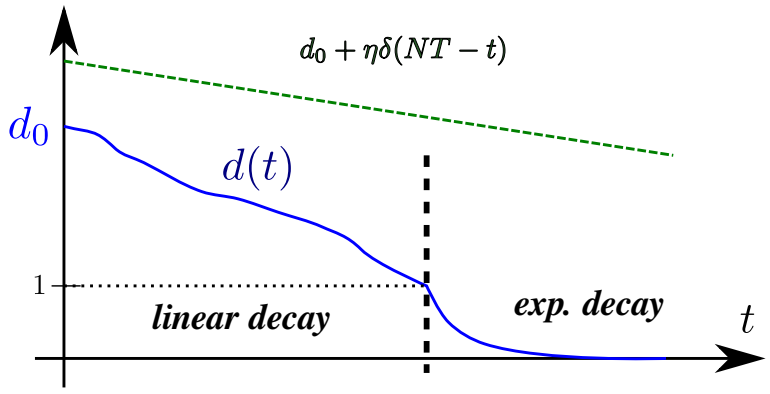

Figure 22: The decay of the diameter $d(t)$ is first linear and then exponential after the diameter $d(t)$ becomes less than 1.

Remark 13 Notice that we could not derive an explicit decay rate in the multi-dimension case. Indeed, in one dimension, we exploited the property that the left behind dynamics preserve the ordering of the agents and in particular that the diameter forming agents are the same agents for all time (i.e. p and q are time independent). In multiple dimensions it is possible for the diameter forming agents to change - this prevents us from applying the techniques used in one dimension to multiple dimensions.

3.3.3 Diameter decay: numerical experiments

From the previous results, we can see that convergence to a consensus occurs in two stages. Before the diameter d becomes less than 1 the convergence is at least linear, after d is less than 1 the convergence becomes exponential. The estimation of the convergence rates provided by Theorem 6 is fairly rough. We would like to explore numerically if in practice the decay of the diameter is faster.

We perform 100 realizations of the *NOLB* dynamics with initial conditions for the configuration $\mathbf{x}_i$ taken from a uniform distribution on the interval $[0, 10]$ (1D simulation).

We then compute the evolution of the diameter $d(t) = \max_{i,j} |x_i(t) - x_j(t)|$ for all the realizations. In Figure 23-left, we plot the decay of the 'median' of the diameter (red) along with the slowest and fastest decay (dashed blue). To measure the disparity of $d(t)$, we also plot the 5% and 95% quantile. We observe two phases in the decay of the diameter: initially $d(t)$ decays quickly and then starts to slow down until it reaches the distance 1 when it decays exponentially fast. We notice that there are a large variation between the different realizations. Indeed, if we denote τ the stopping time at which d reaches 1:

$$\tau = \min_{t \geq 0} \{d(t) \leq 1\}, \tag{3.39}$$

then we observe that τ varies between 22 time units (fastest realization) and 50 time units(slowest realization). Thus, finding a sharp decay rate for the NOLB dynamics seems challenging.

Additionally, we would like to explore how the the radius, r_*, of the critical region affects the convergence. Naively one might expect that lower values of r_* result in faster convergence as agents are more free to move. However, we find that this is not the case; intuitively since lower values of r^* allow for more freedom of movement agents can have a higher degree of clustering which hinders the emergence of a consensus. For each value of r_* ranging from 0 to 1, we run 100 simulations of the NOLB dynamics and estimate the stopping time τ (3.39). The results are plotted in Figure 23-right. We find that indeed τ is lower for higher values of r_*. Interestingly, this effect seems to become less prominent for $r_* \geq 0.2$.

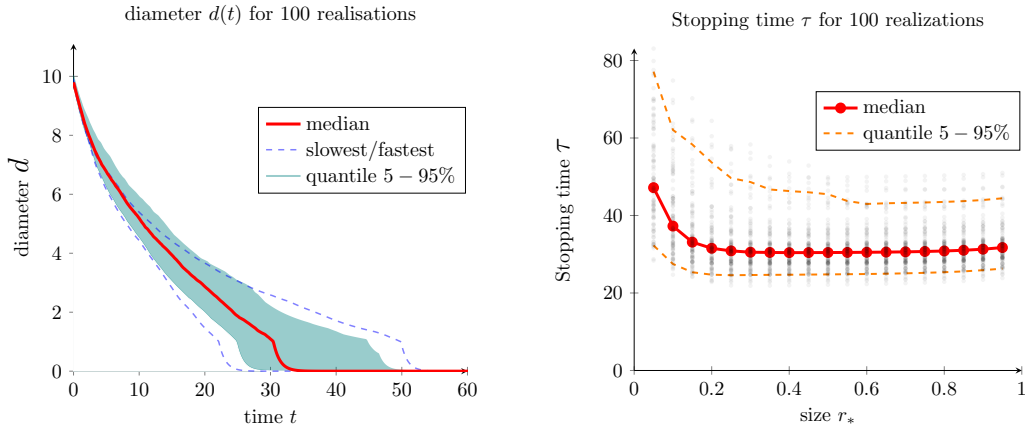

Figure 23: Left: diameter $d(t)$ over time for 100 realizations (quantile representation). **Right:** stopping time τ (3.39) depending on the size of the critical region r_*.

3.4 Relaxed no one left behind

In this section we will investigate whether it is possible to weaken the constraints imposed by the NOLB dynamics and still maintain convergence to a consensus. The critical ingredient in the argument used to show the convergence to consensus of the NOLB dynamics was the preservation of connectivity of the entire configuration of agents. However, the dynamics introduced in (3.8) preserve connectivity between individual *agents* by Proposition 6 - once two agents begin interacting they continue to do so throughout the evolution of the dynamics as each agent "takes care of" every agent in its critical region. However while this is clearly sufficient to preserve connectivity of the whole configuration, it isn't necessary. Instead of maintaining direct connectivity between agents we need only to maintain a *path* between them.

Instead of preventing all disconnections as in the NOLB dynamics, we could allow individual agents to disconnect as long as they remain connected to a mutual neighbor - an agent does not have to "take care" of an agent in its behind region if one of its neighbors is

already doing so. This intuition can be made rigorous via a description using the behind graph. We name the resulting dynamics *relaxed* no one left behind as we remove constraints while maintaining (global) connectivity. It is however unclear whether the new dynamics will lead to a faster convergence to consensus.

3.4.1 Model introduction

Before introducing the model, we formally define what we mean by a relaxed behind graph. We recall that we denote by $G = (V, E)$ and $G^{\mathcal{B}} = (V, E^{\mathcal{B}})$ respectively the interacting graph and the behind graph. $V = \{1 \dots N\}$ is the set of all N opinions while E and $E^{\mathcal{B}}$ are the edge sets defined as:

- $(i, j) \in E$ if $\|\mathbf{x}_i - \mathbf{x}_j\| \leq 1$,
- $(i, j) \in E^{\mathcal{B}}$ if $j \in \mathcal{B}_i$.

The behind graph $E^{\mathcal{B}}$ is a directed subgraph of E (see remark 10). To define the relaxed behind graph, we identify unnecessary edges. If two agents i and j are neighbors then it is possible that their behind regions overlap and therefore possible that a third agent, k, might be in both behind regions. In terms of the behind graph this means that there is an edge from agent i to agent k *and* from agent j to agent k. However since i and j are neighbors, only one of those edges needs to be present in the behind graph for the interaction graph of the configuration of agents to remain connected; i does not need to take care of k if j is already doing so (or vice versa). Therefore, we could remove one of those edges from the behind graph creating a new edge set $\widetilde{E}^{\mathcal{B}}$ (see Figure 24); if the configuration of agents evolves according to the NOLB dynamics in terms of the relaxed edge set $\widetilde{E}^{\mathcal{B}}$ the interaction graph will still remained connected as paths between agents are maintained.

We now define formally a relaxed behind graph.

Definition 12 *Given a configuration of agents* $\{\mathbf{x}_i\}_i \subseteq \mathbb{R}^d$*, their corresponding behind graph* $G^{\mathcal{B}} = (V, E^{\mathcal{B}})$ *and interaction graph* $G = (V, E)$*, we say that* $\widetilde{G}^{\mathcal{B}} = (V, \widetilde{E}^{\mathcal{B}})$ *is a* ***relaxed behind graph*** *of* $G^{\mathcal{B}}$ *if:*

- $\widetilde{G}^{\mathcal{B}}$ *is a subgraph of* $G^{\mathcal{B}}$
- *for any* $(i, j) \in E$*, if* $(i, k) \in \widetilde{E}^{B}$ *then* $(j, k) \notin \widetilde{E}^{B}$.

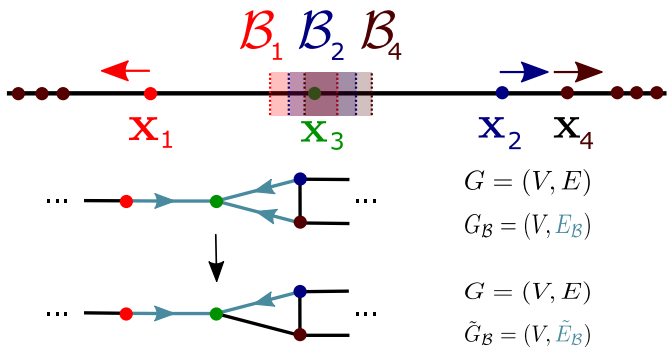

Figure 24: An example of how the behind graph can be relaxed while still ensuring that the interaction graph remains connected. The interaction graph is represented by undirected and directed edges, the behind graph is represented by only the blue directed edges. Agent 3 is in the behind region of both agent 2 and agent 4 and agents 2 and 4 are connected in the interaction graph therefore we may remove the edge from agent 4 to agent 3.

Note that the relaxed behind graph is not unique as one has degrees of freedom in which edges are removed ((i, k) is removed *or* (j, k)). However, given any full behind graph any two relaxed behind graphs will have the same number of edges. We can now define a relaxed version of the NOLB dynamics. Intuitively this new model is exactly the NOLB dynamics however instead of using the full behind graph it is defined in terms of a relaxed behind graph.

Model 3 (RNOLB) *Let* $\{\mathbf{x}_1, ..., \mathbf{x}_N\} \subseteq \mathbb{R}^d$ *be a configuration of agents with behind graph* G^B *and let* $\widetilde{G}^{\mathcal{B}} = (V, \widetilde{E}^{\mathcal{B}})$ *be a relaxed behind graph corresponding to* G^B*. The* ***relaxed no***

one left behind (RNOLB) *dynamics are given by:*

$$\mathbf{x}_i' = P_{\widetilde{C}_i}\left(\overline{\mathbf{x}_i} - \mathbf{x}_i\right) \tag{3.40}$$

where $P_{\widetilde{C}_i} : \mathbb{R}^d \to \widetilde{C}_i$ *is the projection operator associated to the cone of velocities* $\widetilde{C}_i$ *given by:*

$$\widetilde{C}_i = \{\mathbf{v} \in \mathbb{R}^d \mid \langle \mathbf{v}, \mathbf{x}_j - \mathbf{x}_i \rangle \geq 0 \quad \forall\, j \text{ such that } (i,j) \in \widetilde{E}^{\mathcal{B}}\}. \tag{3.41}$$

We now present an algorithm to easily calculate the relaxed behind graph. Intuitively, at each time, t, an order is randomly computed for the agents. Then, according to the order each agent projects its velocity towards agents in its behind region that have not already been projected towards by neighboring agents earlier in the order; each agent "takes care of" agents in its behind region that have not already been taken care of by one of its neighbors. In Figure 25 we demonstrate that under these dynamics that individual agents are indeed allowed to disconnect however the connection of the whole configuration is maintained.

Algorithm 2 Compute relaxed behind graph

```
1:  procedure COMPUTE RELAXED BEHIND GRAPH
2:      Choose an order σ ~ Unif(permutations of {1, ..., N})
3:      for i ∈ {1, ..., N} do
4:          for j ∈ {1, ..., N} do
5:              if (i, j) ∈ E^B and there is no k such that (i, k) ∈ E and (k, j) ∈ Ẽ^B then
6:                  Add (i, j) to Ẽ^B
7:              end if
8:          end for
9:      end for
10: end procedure
```

Notice that in the example of the NOLB dynamics in Figure 25 that the connectivity of the whole configuration of agents is maintained and further that once any two agents connect they remain connected. In particular this is true for the agents corresponding to the red and blue trajectories. However in the RNOLB dynamics these two agents become immediately

disconnected. Nonetheless, the connectivity of the whole configuration is maintained as the green agent "takes care" of the blue agent and preserves the existence of a path between the red and blue agents. Indeed, it is clear that two agents i and j connected by a path cannot become disconnected.

Proposition 8 *The RNOLB dynamics maintain connectivity of the whole configuration of agents.*

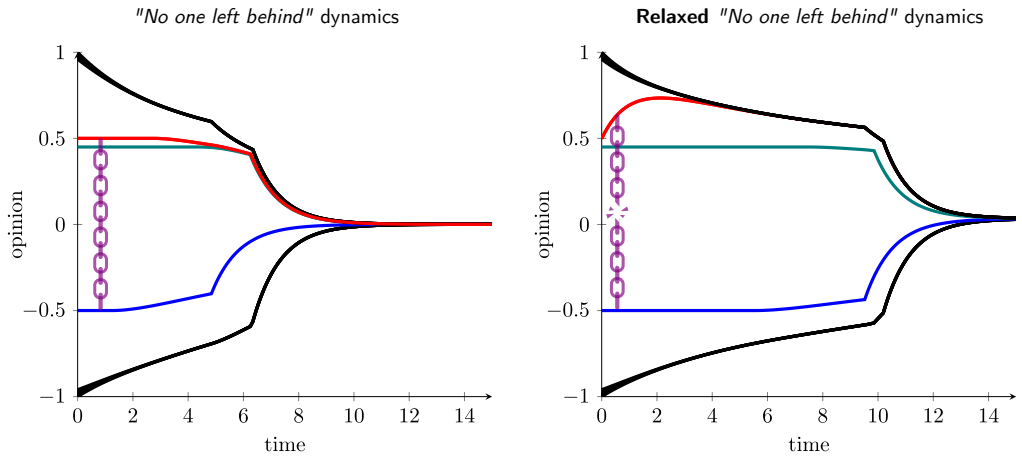

Figure 25: The NOLB dynamics do not allow the red agent to disconnect from the blue agent (illustrated with a purple chain). The RNOLB dynamics allow this disconnection to occur but maintain connectivity of the whole configuration.

3.4.2 RNOLB as an interpolation between NOLB and bounded confidence

In this section we conduct a numerical experiment in one dimension to demonstrate that, in a sense, the RNOLB dynamics can be seen as an "interpolation" between the Hegselmann Krause dynamics (3.1) and the NOLB dynamics defined in Model 2. One of the hallmarks of the bounded confidence dynamics is the formation of clusters of opinions. The requirement of the NOLB dynamics that agents not move if there is another agent in their critical region

prevents the formation of clusters (see for example, Figures 13 and 17). Agents with opinions on the interior of the convex hull of opinions always have another agent in their critical region and are prevented from moving until the boundary of the convex hull has contracted sufficiently close to them. We will see that the weaker conditions of the RNOLB model allow cluster formation to occur initially while still maintaining convergence to a consensus (for a connected initial condition).

To measure the amount of clustering in a configuration of agents we introduce a simple metric. Let $R = L/N$ where L is the length of the range of possible opinions and N is the number of agents. Given a configuration of agents $\{ \mathbf{x}_i \}_i$ and an agent in the configuration $\mathbf{x}_j$, we count the number of agents who are within R of $\mathbf{x}_j$. We then take the average over all agents in the configuration. We will refer to this metric as the *clustering number*. If the configuration of agents is initially uniformly distributed the clustering number is 2 (in one dimension). If agents begin to cluster the clustering number should increase as agents begin to collect more neighbors within R. Clearly, the maximum clustering number for any configuration is the number of agents in the configuration and if a consensus is reached that maximum will be attained.

In the top three plots of Figure 26 we show the long term behavior of the RNOLB, NOLB, and bounded confidence dynamics for the same initial condition. As expected, the bounded confidence dynamics do not reach a consensus and the NOLB and RNOLB dynamics do. Additionally, we qualitatively observe the formation of distinct clusters in the bounded confidence dynamics and the lack of clusters in the NOLB dynamics. Interestingly, in the RNOLB dynamics we initially observe the formation of clusters that are qualitatively very similar to those observed in the bounded confidence dynamics. However, instead of remaining distinct (as in the bounded confidence model) these clusters eventually merge and a consensus is reached.

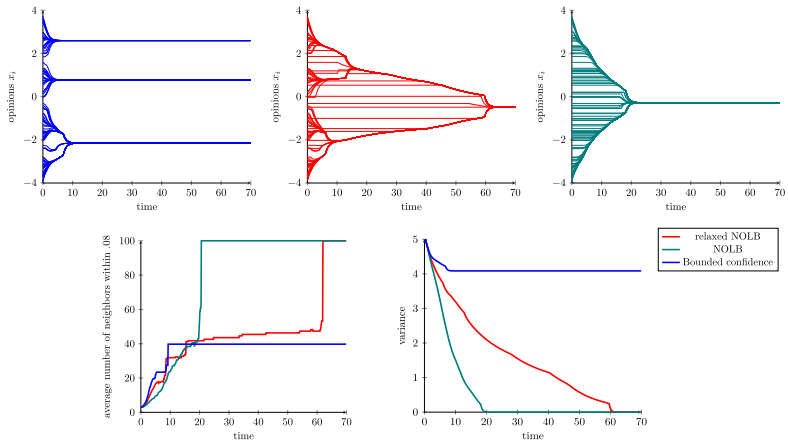

Figure 26: The RNOLB dynamics can be seen as an interpolation between NOLB and bounded confidence.

This qualitative observation is supported by measuring the clustering number of each configuration as time evolves - this is plotted in the bottom left plot of Figure 26. Notice that in the initial period of cluster formation (roughly before $t = 10$) that the clustering number of the bounded confidence dynamics is the fastest to increase whereas the NOLB dynamics is the slowest which reflects the strong formation of clusters in the bounded confidence dynamics and weak cluster formation in NOLB. Notably, in this period the clustering number of the RNOLB dynamics increases more slowly than the clustering number of the bounded confidence dynamics but faster than the clustering number of the NOLB dynamics. This supports our qualitative observation that RNOLB is in a sense an interpolation of bounded confidence and NOLB as it has (at least initially) more clustering than NOLB as it allows for free movement of many agents on the interior of the convex hull, but less clustering than bounded confidence as it forces some "moderating" agents to maintain their position in order to preserve connectivity and eventually reach consensus. We can also observe this interpolation in the evolution of the variance of each configuration which is shown in the bottom right plot of Figure 26. Counter-intuitively, strong clustering causes the decay speed of the variance to reduce as agents in individual clusters can (at least

initially) move away from each other. This effect is observed in the example in Figure 26 as the NOLB dynamics have the fastest decay in variance, the bounded confidence dynamics have the slowest, and the variance decay of RNOLB is faster than bounded confidence but slower than NOLB which reflects the varying amount of clustering observed in the three different models. We also observe that due to the higher degree of clustering, the RNOLB dynamics are slower to converge to a consensus then the NOLB dynamics. Despite this, we still observe exponential convergence once all agents are within the interaction range of each other.

3.4.3 Diameter decay in RNOLB

In our previous results concerning the NOLB dynamics we show rigorously that convergence to a consensus occurs in two stages; when the diameter of the configuration of agents is greater than one the rate is linear, afterwards it spontaneously becomes exponential. Our estimates of these rates were fairly rough and in an attempt to discover whether the rates are faster in practice we simulated 100 realizations of the NOLB dynamics and found a large disparity in stopping times which suggests that our estimates are unlikely to be improved. Here, we repeat this experiment for the RNOLB dynamics in order to investigate whether there is a similar phase transition in the convergence and disparity in stopping times.

We perform 100 realizations of the RNOLB dynamics with initial conditions drawn from a uniform distribution on the interval [0, 10]. We then again compute the evolution of the diameter of the configuration $d(t)$ for all realizations and plot the decay of the median of the diameter, the slowest and fastest decay, and the 5% and 95% quantiles. Here, we again observe two phases of convergence with clear exponential convergence again emerging once the diameter reaches 1. However, compared to the NOLB dynamics convergence is much

slower and there is a much larger disparity between stopping times. The fastest evolution reaches a diameter of 1 at ≈ 50 time units whereas the slowest evolution takes greater than 200 time units. This suggests that it will be difficult to obtain tight convergence rates in the case of the RNOLB dynamics as well.

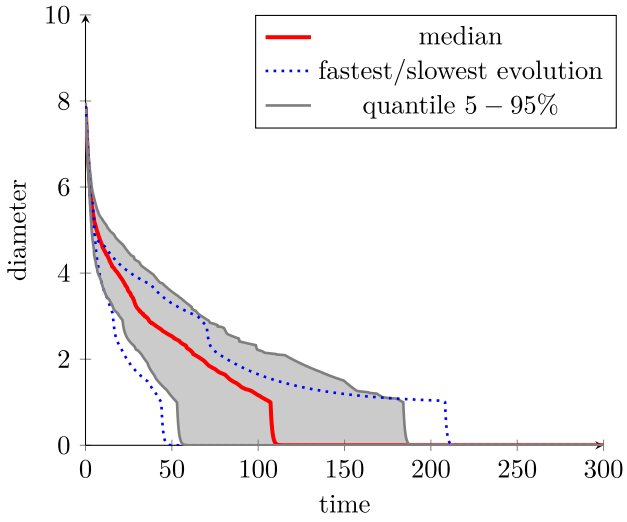

Figure 27: Diameter, $d(t)$ over time for 100 realizations (quantile representation).

Chapter 4

THE FOOD SEEKING BEHAVIOR OF SLIME MOLD: A MACROSCOPIC APPROACH

This chapter closely follows [52]. In this chapter we study a model hierarchy inspired by the food seeking behavior of a slime mold. Starting from a particle model who's interaction term shares many features with the agent based models discussed in previous chapters, we derive a macroscopic aggregation-diffusion equation for the evolution of slime mold under the assumption of propagation of chaos in the large particle limit. We analyze properties of the macroscopic model in the stationary case and study the behavior of the slime mold between food sources. The efficient numerical simulation of the aggregation-diffusion equation allows for a detailed analysis of the interplay between the different regimes drift, interaction and diffusion.

4.1 Slime mold model derivation

We now turn our attention to a model inspired by the movement of the slimemold organism. The main features of slime mold food seeking behavior are aggregation of mass on food sources while maintaining a connected mass and retraction in the absence of food sources. To model these behaviors we first introduce an agent-based model. Here, the slime mold is represented by a swarm of N agents who's trajectories, $\mathbf{X}_i(t) \in \mathbb{R}^d$, are modeled as a system of stochastic differential equations. In the following we will refer to this model as the *microscopic model* in order to reinforce the intuition that the agents represent slime mold "particles".

Definition 13 (Microscopic Model) *Consider a collection of N agents with (random) po-*

sitions

$(\mathbf{X}_1(t), ..., \mathbf{X}_n(t)) \in (\mathbb{R}^d)^N$. *The* ***microscopic model*** *is given by the evolution:*

$$d\mathbf{X}_i(t) = A\nabla V(\mathbf{X}_i(t))dt + B\frac{1}{N-1}\sum_{j\neq i}^{N} \nabla W(\mathbf{X}_j(t) - \mathbf{X}_i(t))dt + \sqrt{C}d\mathbf{B}_i(t) \tag{4.1}$$

where $\{\mathbf{B}_i\}_i$ *are independent, standard* $\mathbb{R}^d$*-valued Brownian motions,* $V, W : \mathbb{R}^d \to \mathbb{R}$ *are* C^1 *functions,* $\{\mathbf{X}_i(0)\}_i$ *are independent and identically distributed random variables independent of* $\{\mathbf{B}_i\}$ *with probability density function given by* $\rho_0 : \mathbb{R}^d \to \mathbb{R}$*, and* $A, B, C > 0$.

We note that if we assume that ∇V and ∇W are globally Lipschitz and further if there exists $K_1, K_2 > 0$ such that for $x \in \mathbb{R}^d$:

$$|\nabla V(x)| \leq K_1(1 + |x|) \quad \text{and } |\nabla W(x)| \leq K_2(1 + |x|),$$

then it can be shown [124] that (4.1) has a unique strong solution.

Intuitively, V represents the density of chemo-attractants given off by food sources. For this reason, we usually think of V as a sum of radially symmetric, positive functions. Each term represents the attractants emanating from a single food source. A prototypical example would be a sum of Gaussians centered at different positions. W represents the interaction between agents. Notice that, roughly, if $\nabla W(\mathbf{X}_j(t)-\mathbf{X}_i(t)) \approx \mathbf{X}_j(t)-\mathbf{X}_i(t)$, then the interaction causes agents to *attract* each other. Likewise if $\nabla W(\mathbf{X}_j(t) - \mathbf{X}_i(t)) \approx \mathbf{X}_i(t) - \mathbf{X}_j(t)$, then the agents will *repel* each other. As we have noted, *P. polycephalum* tends to maintain a connected mass, even when aggregating around disparate food sources. Therefore, this interaction term should cause agents to remain locally close, i.e. agents should not repel each other at long ranges. The noise term, $Cd\mathbf{B}_i(t)$, models *P. polycephalum* foraging behavior. In the absence of chemo-attractants, *P. polycephalum* forages by diffusing outward from its initial position while still maintaining a connected mass. Here, in the absence of

a food source ($V = 0$) the noise term should cause particles to spread from their initial positions while the interaction term causes them to still have some propensity to remain locally "together".

We include the parameters A, B and C for the case that simulations of (4.1) demonstrate the need to scale the terms' influence on each particles trajectory to produce evolutions that resemble *P. polycephalum*. However, any simulation of (4.1) that would give an accurate picture of how (4.1) models *P. polycephalum* would have to use a very large number of particles; this becomes computationally intractable. Therefore, we must find a way to examine the behavior or (4.1), for a large number of particles, without computing the trajectory of every particle explicitly. We will show that in the large particle limit that the evolution of the marginal distribution of any particle is given by an *aggregation-diffusion equation*. There exist computationally efficient methods for simulating such equations, therefore we can examine the behavior of (4.1) by simulating the evolution of the marginal density. The technique we will employ to derive an appropriate model is the assumption of *propagation of chaos*; i.e. if the number of agents agents is large, the trajectories of any two agents can be assumed to be independent as they are not likely to interact strongly. This is true in many cases, see for example [30, 31].

Theorem 7 *In the $N \to \infty$ limit and for a given, $A, B, C > 0$, $W, V \in C^1(\mathbb{R}^d)$ and ρ_0, the evolution of the marginal distribution of any agent evolving according to* (4.1) *is a solution of the aggregation-diffusion equation*

$$\partial_t \rho(x,t) = A\nabla \cdot \Big[\nabla V(x)\rho(x,t)\Big] + B\nabla \cdot \Big[\rho(x,t)\nabla W * \rho(x,t)\Big] + C\Delta\rho(x,t),$$
$$\rho(x,0) = \rho_0(x)\ ,\ \ x \in \mathbb{R}^d$$

under the assumption of propagation of chaos; that is for any finite group of M agents there

exist distributions $\rho_1, ..., \rho_M$ on $\mathbb{R}^d$ such that the joint distribution $\rho^M(x_1, ..., x_M, t)$ satisfies:

$$\rho^M(x_1, ..., x_M, t) \to \prod_{i=1}^{M} \rho_i(x_i, t) \text{ as } N \to \infty$$

for any time.

Proof

Consider a finite group of N agents, $(X_i(t), t \geq 0) \subseteq \mathbb{R}^d, i = 1, ..., N$ evolving according to (4.1). Let μ^N be the empirical distribution of the configuration of agents. That is:

$$\mu^N(t) = \frac{1}{N} \sum_i \delta_{X_i(t)}$$

where $\delta_{X_i(t)}$ is the Dirac measure with unit mass at $X_i(t)$. Let F be a test function in $C_c^\infty((\mathbb{R}^d)^N)$, therefore by definition of the Dirac measure we deduce:

$$\langle \mu^N, F \rangle = \frac{1}{N} \sum_i F_i(X_i(t))$$

we first derive an expression for the time evolution of the expectation of this quantity. Let $\mathbf{X}(t) = (X_1(t), ..., X_N(t))$. By conditioning we find

$$\frac{\mathbb{E}[F(\mathbf{X}(t + \Delta t)] - \mathbb{E}[F(\mathbf{X}(t))]}{\Delta t} = \int_{(\mathbb{R}^d)^N} \frac{\mathbb{E}[F(\mathbf{X}(t + \Delta t)|\mathbf{X}(t) = \mathbf{x}] - F(\mathbf{x})}{\Delta t} \rho^N(\mathbf{x}, t) d\mathbf{x} \quad (4.2)$$

where $\rho^N(\mathbf{x}, t)$ is the joint distribution of the collection of agents. Using Ito's Lemma (see e.g. [124]) we can further expand the conditioned term in the above

$$\begin{aligned} \mathbb{E}\Big[F(\mathbf{X}(t + \Delta t))|\mathbf{X}(\mathbf{t}) = \mathbf{x}\Big] = F(\mathbf{x}) \\ + \mathbb{E}\Big[\sum_{j=1}^{N} \int_t^{t+\Delta t} \partial_{x_j}[F(\mathbf{X}(t))] dX_j(s)|\mathbf{X}(t) = \mathbf{x}\Big] \\ + \mathbb{E}\Big[\frac{1}{2} \sum_{j,k}^{N} \int_t^{t+\Delta t} \partial_{x_j}\partial_{x_k}[F(\mathbf{X}(t))] d[X_i(s), X_j(s)]|\mathbf{X}(t) = \mathbf{x}\Big]. \end{aligned} \quad (4.3)$$

We first focus on the third term of (4.3). Notice that as each X_i evolves according to (4.1), since the noise terms for disparate agents are independent, we have that:

$$d[X_i(s), X_j(s)] = \begin{cases} 0 & \text{if} \quad i \neq j \\ Cds & \text{if} \quad i = j \end{cases}.$$

Therefore, the third term of (4.3) simplifies as:

$$\mathbb{E}\Big[\frac{1}{2}\sum_{j,k}^{N}\int_t^{t+\Delta t}\partial_{x_j}\partial_{x_k}[F(\mathbf{X}(t))]d[X_i(s), X_j(s)]|\mathbf{X}(t) = \mathbf{x}\Big] = \mathbb{E}\Big[C\int_t^{t+\Delta t}\Delta F(\mathbf{X}(s))ds|\mathbf{X}(t) = \mathbf{x}\Big]. \tag{4.4}$$

Now, turning to the second term of (4.3) we find by appyling (4.1) that:

$$\begin{aligned}
&\mathbb{E}\Big[\sum_{j=1}^{N}\int_t^{t+\Delta t}\partial_{x_j}\Big[F(\mathbf{X}(t))]dX_j(s)|\mathbf{X}(t) = \mathbf{x}\Big] \\
&= \mathbb{E}\Big[\sum_{j=1}^{N}\int_t^{t+\Delta t}\partial_{x_j}[F(\mathbf{X}(s))]\Big(A\nabla V(X_j(s))\Big)ds|\mathbf{X}(t) = \mathbf{x}\Big] \\
&+ \mathbb{E}\Big[\sum_{j=1}^{N}\int_t^{t+\Delta t}\partial_{x_j}[F(\mathbf{X}(s))]\Big(\frac{B}{N-1}\sum_{k\neq j}\nabla W(X_j(s) - X_k(s))\Big)ds|\mathbf{X}(t) = \mathbf{x}\Big] \\
&+ \mathbb{E}\Big[\sum_{j}\int_t^{t+\Delta t}\partial_{x_j}[F(\mathbf{X}(s))]\Big(CdB_j(s)\Big)|\mathbf{X}(t) = \mathbf{x}\Big].
\end{aligned} \tag{4.5}$$

As the stochastic integral is a martingale, we have that the last term of (4.5) is equal to 0. As a result, by combining (4.4) and (4.5), we have that:

$$\begin{aligned}
\mathbb{E}\Big[F(\mathbf{X}(t+\Delta t))|\mathbf{X}(\mathbf{t}) = \mathbf{x}\Big] &= F(\mathbf{x}) \\
&+ \mathbb{E}\Big[\sum_{j=1}^{N}\int_t^{t+\Delta t}\partial_{x_j}[F(\mathbf{X}(s))]\Big(A\nabla V(X_j(s))\Big)ds|\mathbf{X}(t) = \mathbf{x}\Big] \\
&+ \mathbb{E}\Big[\sum_{j=1}^{N}\int_t^{t+\Delta t}\partial_{x_j}[F(\mathbf{X}(s))]\Big(\frac{B}{N-1}\sum_{k\neq j}\nabla W(X_j(s) - X_k(s))\Big)ds|\mathbf{X}(t) = \mathbf{x}\Big] \\
&+ \mathbb{E}\Big[C\int_t^{t+\Delta t}\Delta F(\mathbf{X}(s))ds|\mathbf{X}(t) = \mathbf{x}\Big].
\end{aligned} \tag{4.6}$$

So, by (4.2) and the continuity of the trajectories, (4.6) implies P almost-surely that:

$$\begin{aligned}\partial_t \int_{(\mathbb{R}^d)^N} F(\mathbf{x})\rho^N(\mathbf{x},t)d\mathbf{x} = &\int_{(\mathbb{R}^d)^N} C\Delta F(\mathbf{x})\rho^N(\mathbf{x},t)d\mathbf{x} \\ &+ \int_{(\mathbb{R}^d)^N} \sum_{j=1}^N \partial_{x_j}[F(\mathbf{x})]\Big(A\nabla V(X_j(t))\Big)\rho^N(\mathbf{x},t)d\mathbf{x} \\ &+ \int_{(\mathbb{R}^d)^N} \sum_{j=1}^N \partial_{x_j}[F(\mathbf{x})]\Big(\frac{B}{N-1}\sum_{k\neq j}\nabla W(X_j(t)-X_k(t))\Big)\rho^N(\mathbf{x},t)d\mathbf{x}.\end{aligned} \tag{4.7}$$

Therefore, ρ^N is a weak solution to the following initial value problem:

$$\begin{aligned}&\partial_t\rho^N(\mathbf{x},t) - C\Delta\rho^N(\mathbf{x},t) - \sum_{j=1}^N \partial_{x_j}\Big[A\nabla V(x_j)\rho^N(\mathbf{x},t)\Big] \\ &\quad - \sum_{j=1}^N \partial_{x_j}\Big[\rho^N(\mathbf{x},t)\frac{B}{N-1}\sum_{k\neq j}\nabla W(x_j-x_k(t))\Big] = 0, \\ &\rho^N(\mathbf{x},0) = \prod_{j=1}^N \rho_0(x_j).\end{aligned} \tag{4.8}$$

Using (4.8), we now compute the time evolution of the marginal distribution of one agent in the $N \to \infty$ limit under the assumption of *propagation of chaos*, that is for any *finite* collection of M agents we have that their joint distribution satisfies (in a limiting sense):

$$\rho^M(x_1,...,x_N,t) = \prod_{i=1}^M \rho_i(x_i,t). \tag{4.9}$$

In the following we consider a finite group of M particles in the $N \to \infty$ limit. Without loss of generality we will consider the marginal distribution, $\rho(x_1,t)$ of the first agent which is given by:

$$\rho(x_1,t) = \int_{(\mathbb{R}^d)^{M-1}} \rho^M(x_1,...,x_M)d(x_2,...,x_M)$$

and therefore, by (4.8), we have that $\rho(x_1, t)$ satisfies (weakly):

$$\begin{aligned}\partial_t \rho(x_1,t) = &\sum_{j=1}^{M} \int_{(\mathbb{R}^d)^{M-1}} \partial_{x_j}\Big[A\nabla V(x_j)\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M)\\ &+ \sum_{j=1}^{M} \int_{(\mathbb{R}^d)^{M-1}} \partial_{x_j}\Big[\frac{B}{M-1}\sum_{k\neq j}\nabla W(x_j-x_k)\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M)\\ &+ \int_{(\mathbb{R}^d)^{M-1}} C\Delta\rho^M(\mathbf{x},t)d(x_2,...,x_M).\end{aligned} \tag{4.10}$$

We analyze (4.10) term by term. First, by applying (4.9) we find that

$$\int_{(\mathbb{R}^d)^{M-1}} C\Delta\rho^M(\mathbf{x},t)d(x_2,...,x_M) = C\Delta\rho_1(x_1,t). \tag{4.11}$$

Next, if we again apply (4.9) and integrate by parts in each term of the sum where $j \neq 1$ we find that

$$\sum_{j=1}^{M} \int_{(\mathbb{R}^d)^{M-1}} \partial_{x_j}\Big[A\nabla V(X_j(t))\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M) = \partial_{x_1}\Big[A\nabla V(x_1)\rho(x_1,t)\Big]. \tag{4.12}$$

Similarly we deduce

$$\begin{aligned}&\sum_{j=1}^{M} \int_{(\mathbb{R}^d)^{M-1}} \partial_{x_j}\Big[\frac{B}{M-1}\sum_{k\neq j}^{M}\nabla W(x_j-x_k)\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M) =\\ &\qquad\int_{(\mathbb{R}^d)^{M-1}} \partial_{x_1}\Big[\frac{B}{M-1}\sum_{k=2}^{M}\nabla W(x_1-x_k)\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M).\end{aligned} \tag{4.13}$$

We can simplify (4.13) further by integrating under the divergence operator and applying (4.9), notice that:

$$\begin{aligned}&\int_{(\mathbb{R}^d)^{M-1}} \sum_{k=2}^{N}\nabla W(x_1-x_k)\rho^M(\mathbf{x},t)d(x_2,...,x_M)\\ &= \sum_{k=2}^{M}\int_{(\mathbb{R}^d)^{M-1}} \nabla W(x_1-x_k)\Big(\rho_1(x_1,t)\rho_2(x_2,t),...,\rho_M(x_M,t)\Big)d(x_2,...,x_M)\\ &= \sum_{k=2}^{M}\rho(x_1,t)\int_{\mathbb{R}^d}\nabla W(x_1-x_k)\rho_k(x_k,t)dx_k\\ &= (M-1)\rho(x_1,t)\nabla W * \rho(x_1,t).\end{aligned} \tag{4.14}$$

Hence:

$$\int_{(\mathbb{R}^d)^{M-1}} \partial_{x_1}\Big[\frac{B}{M-1}\sum_{k=2}^{M}\nabla W(x_1-x_k)\rho^M(\mathbf{x},t)\Big]d(x_2,...,x_M) = \partial_{x_1}\Big[B\rho(x_1,t)\nabla W * \rho(x_1,t)\Big]. \tag{4.15}$$

Therefore, combining (4.11), (4.12) and (4.15) we have by (4.10) that the marginal of any agent satisfies:

$$\begin{aligned} \partial_t\rho(x,t) &= A\partial_x\Big[\nabla V(x)\rho(x,t)\Big] + B\partial_x\Big[\rho(x,t)\nabla W * \rho(x,t)\Big] + C\Delta\rho(x,t) \\ \rho(x,0) &= \rho_0(x) \end{aligned} \tag{4.16}$$

as desired. ■

As the equation derived in Theorem 7 describes how the entire collection of particles evolves from a density standpoint, in the following we will refer to it as the *macroscopic model*.

Definition 14 (Macroscopic Model) *The **macroscopic model** is given by the evolution of the aggregation-diffusion equation*

$$\begin{aligned} \rho_t(t,x) &= A(\nabla\cdot(\nabla V\rho)) + B(\nabla\cdot((\nabla W * \rho)\rho)) + C(\Delta\rho) \quad x\in\Omega\subseteq\mathbb{R}^d, \\ \rho(0,x) &= \rho_0 \end{aligned}. \tag{4.17}$$

for $W, V \in C^1(\mathbb{R}^d)$ and $\rho_0 \in H^1(\mathbb{R}^d)$ with $\int_\Omega \rho_0(x)dx = 1$.

Here, V and W continue to represent the chemoattractant of food sources and the slime mold's propensity to aggregate respectively. In the following we will study (4.17) from the viewpoint of using it to model *P. polycephalum* food-seeking behavior.

4.2 Slimemold model properties

Many questions remain about how to utilize the macroscopic equation derived in the previous section to model a slime mold. In this section we present some analysis of

stationary states of the equation in order to inform some of these modeling choices. First, we investigate stationary states in the case where there is no food present as a way to gain some information about what choice of kernel function might be reasonable. Inspired by [88], we believe that Gaussians represent a reasonable model for stationary states of a slime mold - however we find that Gaussian stationary states are only possible in the case that the kernel function, W, is quadratic. Next, we investigate stationary states of the scaled equation. We find through a fixed point argument that if the diffusion scaling parameter, C, is sufficiently higher than A and B then that the only stationary state of the equation is 0 (even in the presence of a food source). This suggests that when tuning the scaling parameters that there cannot be "too much" diffusion.

4.2.1 Zero-food stationary states

In the case that there are no food sources a slime mold will "retract" to a more compact configuration [92, 90, 91]. Often, this configuration is roughly radially symmetric. Given that and inspired by the central limit theorem, we believe that a reasonable model for the configuration of the slime mold in the case that there is no food could be a Gaussian. We now investigate if there are conditions on the kernel function that are imposed by the assumption that stationary states are Gaussian. For simplicity we work in one dimension however the calculations are analogous in higher dimensions. In the case that there are no food sources (4.17) becomes:

$$\rho_t = C\Delta\rho + B\nabla \cdot ((\nabla W * \rho)\rho). \tag{4.18}$$

Therefore a stationary state satisfies (in the following we will ignore the scaling parameters, A and B, as they do not change the computation):

$$-\Delta\rho = \nabla \cdot ((\nabla W * \rho)\rho)$$

which implies that

$$-\nabla\rho = (\nabla W * \rho)\rho + D$$

for some constant D. Finally we have that:

$$\nabla W * \rho = -\nabla(\ln(\rho)). \tag{4.19}$$

Note that (4.19) represents a general condition that zero food stationary states must satisfy. If we assume that stationary states are Gaussian, i.e. that:

$$\rho(x) = \beta \exp\left(-x^2\tau\right),\ \beta, \tau > 0$$

then by (4.19) we have that:

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{-\infty}^{\infty} W(x-y)\beta \exp\left(-y^2\tau\right)\mathrm{d}y\right] = -\frac{\mathrm{d}}{\mathrm{d}x}\left[\ln(\beta \exp\left(-x^2\tau\right)\right] \\ \implies &\int_{-\infty}^{\infty} W'(y)\beta \exp\left(-(y-x)^2\tau\right)\mathrm{d}y = 2x\tau. \end{aligned}$$

The above can only hold if $W'(y) = ay$ for some $a \in \mathbb{R}$. Therefore, Gaussians are zero food stationary states only if the interaction kernel is quadratic. Further, by plugging in ay for $W'(y)$ in the above we can see that the following must hold:

$$\beta a = \frac{2\tau^{\frac{3}{2}}}{\sqrt{\pi}} \tag{4.20}$$

So, for a given quadratic kernel there are a family of Gaussians that are potential zero food stationary states (depending on the mass of the initial profile). For an illustration of

this fact see Figure 28. Here, we simulate (4.17) in the case of no drift term using the "blob method" for aggregation-diffusion equations introduced in [28]. We will discuss the numerical method in more detail in the Numerics section of the paper. We choose an initial condition representative of a slime mold aggregated around two food sources at $x = 1$ and $x = -1$. We find in accordance with our calculation above that the mass profile appears to converge to a Gaussian configuration. This also illustrates a modeling property of (4.17); in the case that food sources "run out" the profile will "retract" to a central configuration.

Effectively, by starting from information about the zero food stationary state, we have "solved" for the correct kernel. This analysis suggests an empirical method for determining the correct kernel function from experimental data. Instead of our assumption that zero food stationary states are Gaussian, statistical analysis of slime mold configurations in the presence of zero food sources could provide information about the stationary state configuration. This information could then be used to "solve" for the kernel that results in this "correct" stationary state.

4.2.2 High diffusion stationary states

We now turn to considering stationary states of the scaled equation:

$$\rho_t = A(\nabla \cdot (\nabla V \rho)) + B(\nabla \cdot ((\nabla W * \rho)\rho)) + C(\Delta \rho). \tag{4.21}$$

We will see in a later section that the "fair" regime ($A = B = C$) does not result in food-source aggregation. Therefore we are motivated to study different scalings of the terms in (4.17). Here, we will rule out a large regime of scaling parameters by employing a fixed point argument to show that if the diffusion parameter, C, is sufficiently larger then the drift and interaction parameters then the only stationary state of (4.17) is 0 on bounded domains.

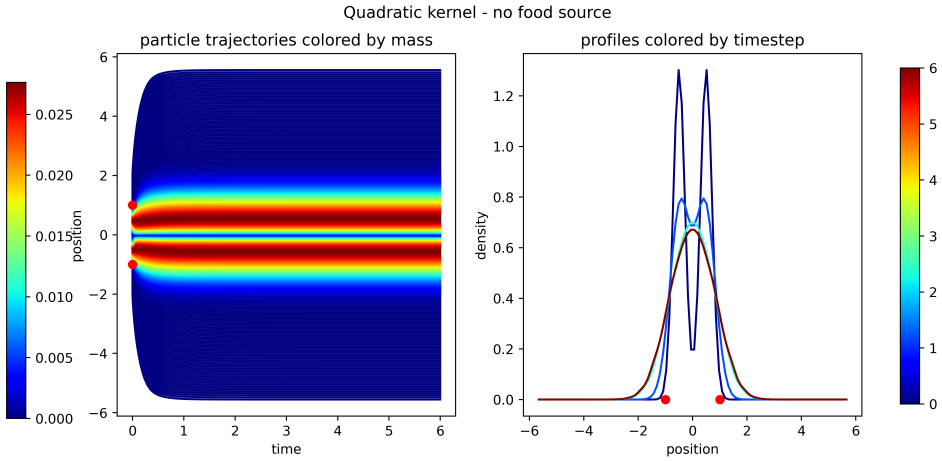

Figure 28: The evolution of (4.17) in the case of a quadratic kernel with no drift term. The initial condition is chosen to represent the state of a slime mold after it has aggregated around two food sources at $x = 1$ and $x = -1$. Notice that the profiles appear to converge on a central Gaussian configuration - an illustration of the fact that Gaussians are stationary states of the equation in the case of a quadratic kernel. Here the kernel is given by $W'(x) = x$, in particular $a = 1$ and therefore, by (4.18), if the final configuration is gaussian its scaling parameters, β and τ, must satisfy $\sqrt{\pi} = \frac{2\tau^{\frac{3}{2}}}{\beta}$.

Theorem 8 *Let $\Omega \subseteq \mathbb{R}^d$ be compact and connected. For a given $A, B > 0$, $V, W \in H^2(\Omega) \cap L^\infty(\Omega)$ and $\rho_0 \in H_0^1(\Omega)$ then for any sufficiently large $C > 0$, 0 is the unique stationary state of:*

$$\rho_t(t,x) = A(\nabla \cdot (\nabla V \rho)) + B(\nabla \cdot ((\nabla W * \rho)\rho)) + C(\Delta \rho^m), \quad x \in \Omega \subseteq \mathbb{R}^d,$$

$$\rho(0,x) = \rho_0(x)$$

in $H_0^1(\Omega)$.

Proof

Recall that ρ is a stationary state of (4.17) if:

$$-\Delta\rho = \frac{A}{C}[\nabla V \cdot \nabla\rho + \Delta V \rho] + \frac{B}{C}[\nabla(W * \rho) \cdot \nabla\rho + \Delta(W * \rho)\rho].$$

We define:

$$G(\rho) := \frac{A}{C}[\nabla V \cdot \nabla\rho + \Delta V \rho] + \frac{B}{C}[\nabla(W * \rho) \cdot \nabla\rho + \Delta(W * \rho)\rho].$$

Given $\rho \in H_0^1(\Omega)$ we know that $G(\rho) \in L^2(\Omega)$ and that the equation:

$$\Delta\tilde{\rho} = G(\rho) \tag{4.22}$$

has a unique solution in $H_0^1(\Omega)$. Given $K > 0$, define $H_0^1(\Omega, K) := \{\, \rho \in H_0^1(\Omega) : |\, |\rho\|_{H_0^1} \leq K \,\}$. It can be shown (see, for example, Theorem 4 in section 6.3 of [44]) that if $\tilde{\rho}$ is the unique solution of (4.22) then the following estimate holds for some $D > 0$:

$$\|\tilde{\rho}\|_{H^2} \leq D\|G(\rho)\|_{L^2}. \tag{4.23}$$

Therefore, by applying the triangle inequality and Cauchy-Schwarz inequality we can see that for a sufficiently large choice of C if $\rho \in H_0^1(\Omega, K)$ then we must also have that the solution to (4.22), $\tilde{\rho}$, belongs to $H_0^1(\Omega, K)$ as well.

Define $\tilde{A} : H_0^1(\Omega, K) \to H_0^1(\Omega, K)$ via:

$$\tilde{A}[\rho] = \tilde{\rho} \quad \text{where } \tilde{\rho} \text{ solves (4.22).}$$

We will show that $\tilde{A}$ is a contraction and thus has a unique fixed point by Banach's fixed point theorem. Since 0 is trivially a fixed point of $\tilde{A}$ and K is arbitrary this implies that 0 is the only fixed point of $\tilde{A}$ and thus the only stationary state of (4.17) . Let $\tilde{\rho}_1 = \tilde{A}[\rho_1]$ and $\tilde{\rho}_2 = \tilde{A}[\rho_2]$. We will show that there exists a $0 < \gamma(A, B, C) < 1$ such that:

$$\|A[\rho_1] - A[\rho_2]\|_{H_0^1} \leq \gamma(A, B, C)\|\rho_1 - \rho_2\|_{H_0^1}.$$

Since Ω is bounded by assumption we have by Poincare's inequality that:

$$\|\tilde{\rho_1} - \tilde{\rho_2}\|_{H_0^1} \leq \alpha\|D(\tilde{\rho_1} - \tilde{\rho_2})\|_{L^2(U)}. \tag{4.24}$$

for some $\alpha > 0$. Therefore since $\tilde{\rho_1}$ and $\tilde{\rho_2}$ are both solutions of (4.22) we must have by

another application of the Poincare inequality that:

$$\begin{aligned}
\|D(\tilde{\rho}_1 - \tilde{\rho}_2)\|^2_{L^2(U)} &\leq |\int_\Omega D(\tilde{\rho}_1 - \tilde{\rho}_2) \cdot D(\tilde{\rho}_1 - \tilde{\rho}_2) dx| \\
&= |-\int_\Omega (G(\rho_1) - G(\rho_2))(\tilde{\rho_1} - \tilde{\rho_2}) dx| \\
&\leq \|G(\rho_1) - G(\rho_2)\|_{L^2(U)} \|\tilde{\rho}_1 - \tilde{\rho}_2\|_{L^2(U)} \\
&\leq \alpha \|G(\rho_1) - G(\rho_2)\|_{L^2(U)} \|D(\tilde{\rho}_1 - \tilde{\rho}_2)\|_{L^2(U)}.
\end{aligned}$$

Therefore we have that:

$$\|D(\tilde{\rho}_1 - \tilde{\rho}_2)\|_{L^2(U)} \leq \alpha \|G(\rho_1) - G(\rho_2)\|_{L^2(U)}$$

and by (4.24) that:

$$\|\tilde{\rho}_1 - \tilde{\rho}_2\|_{H^1_0} \leq \alpha \|G(\rho_1) - G(\rho_2)\|_{L^2(U)}. \tag{4.25}$$

We claim that

$$\|G(\rho_1) - G(\rho_2)\|_{L^2(U)} \leq \gamma(A, B, C) \|\rho_1 - \rho_2\|_{L^2(U)}. \tag{4.26}$$

for some $\gamma > 0$ that depends only on A, B and C for a given W and V. Expanding we find that:

$$\begin{aligned}
\|G(\rho_1) - G(\rho_2)\|_{L^1(U)} \leq & \frac{A}{C} \|\nabla V \cdot \nabla(\rho_1 - \rho_2) + \nabla V(\rho_1 - \rho_2)\|_{L^1(U)} \\
& + \frac{B}{C} \|\nabla(W * \rho_1) \cdot \nabla \rho_1 - \nabla(W * \rho_2) \cdot \nabla \rho_2\|_{L^1(U)} \\
& + \frac{B}{C} \|\Delta(W * \rho_1)\rho_1 - \Delta(W * \rho_2)\rho_2\|_{L^1(U)}.
\end{aligned} \tag{4.27}$$

We will continue our analysis term by term. Starting with the first term and applying the Cauchy-Schwarz inequality we find that

$$\frac{A}{C} \|\nabla V \cdot \nabla(\rho_1 - \rho_2) + \nabla V(\rho_1 - \rho_2)\|_{L^1(U)} \leq \frac{A}{C} \Big[\|\nabla V\|_{L^2(U)} \|\nabla(\rho_1 - \rho_2)\|_{L^2(U)} + \|\Delta V\|_{L^2(U)} \|\rho_1 - \rho_2\|_{L^2(U)} \Big]. \tag{4.28}$$

Now, turning to the second term of (4.27) and applying the Cauchy-Schwarz inequality we find:

$$\frac{B}{C}\|\nabla(W*\rho_1)\cdot\nabla\rho_1-\nabla(W*\rho_2)\cdot\nabla\rho_2\|_{L^1(U)}\leq\frac{B}{C}\Big[\|\nabla W*\rho_1\|_{L^2(U)}\|\nabla(\rho_1-\rho_2)\|_{L^2(U)} + \|\nabla W*(\rho_1-\rho_2)\cdot\nabla\rho_2\|_{L^1(U)}\Big]. \tag{4.29}$$

We will further expand the last term of (4.29). Notice that:

$$\frac{B}{C}\|\nabla W*(\rho_1-\rho_2)\cdot\nabla\rho_2\|_{L^1(U)}\leq\frac{B}{C}\|\rho_1-\rho_2\|_{L^2(U)}\sum_{i=1}^{d}\|W_{x_i}\|_{L^2(U)}\|\rho_{2,x_i}\|_{L^1(U)}. \tag{4.30}$$

Now, combining (4.29) and (4.30) we find

$$\frac{B}{C}\|\nabla(W*\rho_1)\cdot\nabla\rho_1-\nabla(W*\rho_2)\cdot\nabla\rho_2\|_{L^1(U)}\leq\frac{B}{C}\|\nabla W*\rho_1\|_{L^2(U)}\|\nabla(\rho_1-\rho_2)\|_{L^2(U)} + \frac{B}{C}\|\rho_1-\rho_2\|_{L^2(U)}\sum_{i=1}^{d}\|W_{x_i}\|_{L^2(U)}\|\rho_{2,x_i}\|_{L^1(U)}. \tag{4.31}$$

Finally, we turn to the third term of (4.27). Notice that:

$$\begin{aligned}\frac{B}{C}\|\Delta(W*\rho_1)\rho_1-\Delta(W*\rho_2)\rho_2\|_{L^1(U)} &= \frac{B}{C}\int_\Omega|\Delta(W*\rho_1)\rho_1-\Delta(W*\rho_2)\rho_2|dx\\ &=\frac{B}{C}\int_\Omega|\Delta(W*\rho_1)\rho_1-\Delta(W*\rho_2)\rho_1+\Delta(W*\rho_2)\rho_1-\Delta(W*\rho_2)\rho_2|dx\\ &=\frac{B}{C}\int_\Omega|\Delta(W*(\rho_1-\rho_2)\rho_1)+\Delta(W*\rho_2)(\rho_1-\rho_2)|dx\\ &\leq\frac{B}{C}\|\Delta W*\rho_2\|_{L^2(U)}\|\rho_1-\rho_2\|_{L^2(U)}\\ &+\frac{B}{C}\|\Delta W\|_{L^2(U)}\|\rho_1-\rho_2\|_{L^2(U)}\|\rho_1\|_{L^1(U)}.\end{aligned} \tag{4.32}$$

Now, combining our results in (4.28), (4.31) and (4.32) and applying the definition of the

norm in $H_0^1(\Omega)$ we find that:

$$\begin{aligned}
\|G(\rho_1) - G(\rho_2)\|_{L^1(U)} \leq & \frac{A}{C}\Big[\|\nabla V\|_{L^2(U)}\|\rho_1 - \rho_2\|_{H_0^1} + \|\Delta V\|_{L^2(U)}\|\rho_1 - \rho_2\|_{H_0^1}\Big] \\
& + \frac{B}{C}\Big[\|\nabla W * \rho_1\|_{L^2(U)}\|\rho_1 - \rho_2\|_{H_0^1} + \|\rho_1 - \rho_2\|_{H_0^1}\sum_{i=1}^{d}\|W_{x_i}\|_{L^2(U)}\|\rho_{2,x_i}\|_{L^1(U)}\Big] \\
& + \frac{B}{C}\Big[\|\Delta W * \rho_2\|_{L^2(U)}\|\rho_1 - \rho_2\|_{H_0^1} + \|\Delta W\|_{L^2(U)}\|\rho_1 - \rho_2\|_{H_0^1}\|\rho_1\|_{L^1(U)}\Big] \\
& := \gamma(A, B, C)\|\rho_1 - \rho_2\|_{H_0^1}
\end{aligned} \tag{4.33}$$

where $\gamma \geq 0$ and is independent of ρ_1 and ρ_2 as we have assumed a uniform bound on their norms (they are members of $H(\Omega, K)$). Therefore, if W and V are given, γ can be made arbitrarily small if C is sufficiently larger than A and B. Additionally we have that

$$\|G(\rho_1) - G(\rho_2)\|_{L^2(U)} \leq \beta\|G(\rho_1) - G(\rho_2)\|_{L^1(U)}$$

for some $\beta > 0$ and therefore

$$\begin{aligned}
\|G(\rho_1) - G(\rho_2)\|_{L^2(U)} &\leq \beta\|G(\rho_1) - G(\rho_2)\|_{L^1(U)} \\
&\leq \beta\gamma(A, B, C)\|\rho_1 - \rho_2\|_{H_0^1}.
\end{aligned} \tag{4.34}$$

Combining (4.25) and (4.34) we obtain:

$$\|\tilde{\rho_1} - \tilde{\rho_2}\|_{H_0^1} \leq \alpha\|G(\rho_1) - G(\rho_2)\|_{L^2(U)} \leq \alpha\beta\gamma(A, B, C)\|\rho_1 - \rho_2\|_{H_0^1}.$$

Therefore for C sufficiently larger than A and B, the mapping $\tilde{A}$ is a contraction as desired.■

From a modeling point of view, Theorem 8 demonstrates that a sufficiently strong contribution from the diffusion term will destroy any of the qualitative behaviors we wish to capture. We will see this explicitly in our numerical studies of (4.17) in the next section.

4.3 Numerics: qualitative features of the slimemold macroscopic model

We now present some numerical investigations of the macroscopic model derived in the previous section.

$$\rho_t = A(\nabla \cdot (\nabla V \rho)) + B(\nabla \cdot ((\nabla W * \rho)\rho)) + C\Delta\rho. \tag{4.35}$$

These investigations are aimed at replicating qualitative properties of slime mold movement seen in [92, 90, 91], specifically aggregation around food sources while still maintaining a "connected" mass. Recall that here, the function V models the density of chemoattractants dispersed by food sources. Therefore, V will always take the form of a sum of radially symmetric positive functions where each term is centered on a food source. We simulate (4.35) via the "blob method for aggregation-diffusion equations" introduced in [28] which approximates (4.35) by solving the ODE:

$$\begin{aligned} \dot{x}_i(t) = &- A(\nabla V(X_i)) - B(\sum_{j=1}^{N} \nabla W(x_i(t) - x_j(t))m_j) \\ &- C\left(\sum_{j=1}^{N} \left(\left(\sum_{k=1}^{N} \varphi_\epsilon(x_j - x_k)m_k \right) + \left(\sum_{k=1}^{N} \varphi_\epsilon(x_i - x_k)m_k \right)^{-1} \right) \nabla \varphi_\epsilon(x_i - x_j)m_j \right) \end{aligned} \tag{4.36}$$

for a collection of N particles, $(x_1(t), ..., x_N(t), t \geq 0)$, who's initial positions are a regular grid on the domain on which we'd like to approximate (4.35) and who's masses are given by $m_i = \rho(0, X_i(0))$. To recover an approximation of ρ from the positions of the particles we convolve the particle solution with a mollifier φ_ϵ:

$$\bar{\rho}(x, t) = \sum_i \varphi_\epsilon(x - X_i(t))m_i.$$

We assume mollifiers are always of the form:

$$\varphi_\epsilon(x) = \frac{1}{(4\pi\epsilon^2)^{d/2}} e^{-|x|^2/4\epsilon^2},$$

where d is the dimension of the domain. It is shown in [28] that under some regularity and growth conditions on W and V that $\bar{\rho}$ converges to the solution of (4.35). As a first example we will consider (4.35) in one dimension. We will take the food source to be

$$V(x) = -e^{-(1-x)^2} - e^{-(-1-x)^2}.$$

Therefore we can think of the food sources as being located at $x = -1$ and $x = 1$. For the parameters of the ODE we take:

- $N = 100$ particles initially equally space on the interval $[-2.1, 2.1]$, i.e. with separation $h = .042$
- $\epsilon = h^{.99}$ (Mollifier parameter)
- initial mass profile given by $\rho_0(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$, $\sigma^2 = 0.0625$ as in [28].

Note that as the variance of the initial profile is very small that, at the particle level, slime mold particles have a very low probability of being found outside of $[-2.1, 2.1]$ (initially). In Figure 29 we perform a first simulation of (4.17) in one dimension via the blob method in the so called "fair regime" ($A = B = C = 1$). Here, we choose a quadratic interaction kernel - specifically we impose that $W'(x) = x$.

Qualitatively, we do not observe aggregation of slime mold mass around food sources to the degree present in [92]. Therefore, motivated by this example, in the following we will attempt to modify (4.35) in order to better model slime mold food seeking behavior. In Figure 29 we observe that diffusion appears to be the dominating effect. Therefore, in order to examine the interplay of the diffusion term against the other two terms in (4.35) we will employ the scaling parameters, $A, B, C > 0$, in order to control the contributions from the drift term, interaction term and diffusion term respectively. Roughly we examine two main regimes; "drift dominated" (where $A > B, C$) and "interaction dominated" (where

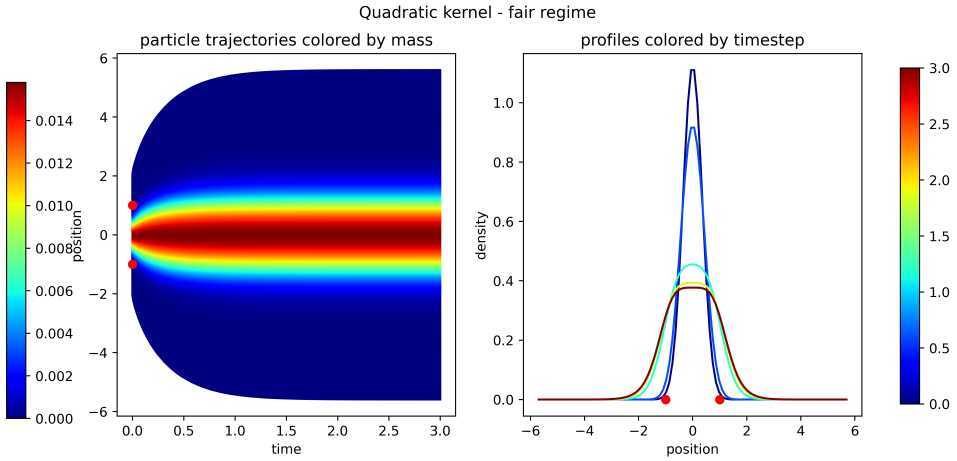

Figure 29: The evolution of the slime mold model (4.35) in the fair regime ($A = B = C$). Food source aggregation is not observed.

$B > A, C$). We also examine so called "competition regimes" where two scaling parameters are equal but dominate the third. We do not examine the diffusion dominated regime as we've previously seen that if $A = B = C$ (we will refer to this as the "fair regime") then diffusion is the dominating effect. In each regime we examine three different interaction kernels - quadratic, polynomial and Gaussian. We choose the same ODE parameters as in the previous fair regime simulation.

4.3.1 Food source dominated regime

We will first examine simulations of (4.17) in the regime where the drift term (which models attraction to the food sources) is the dominating factor. We first examine the case where the kernel is quadratic and given by $W(x) = \frac{x^2}{2}$. Here, the kernel is purely attractive. At the agent based level all agents exert a "pull" on other agents with a force that is proportional to their spacial seperation. The results can be seen in Figure 30. Here, we qualitatively observe that in the case where $A = 10$, $B = C = 1$ that we have strong

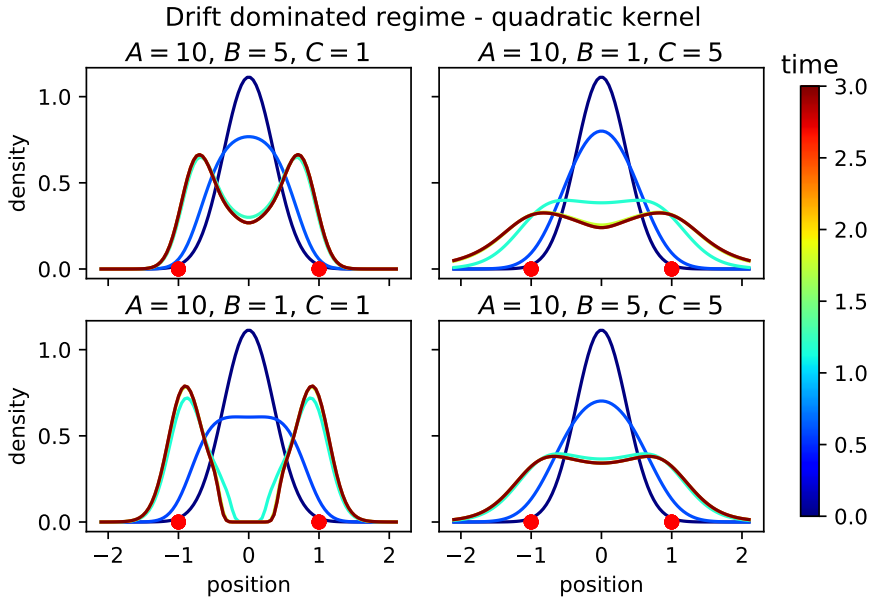

Figure 30: The evolution of (4.17) in the drift dominated regime with kernel given by $W(x) = \frac{x^2}{2}$. Food source aggregation occurred in two cases only one of which maintained a connected mass.

food source aggregation however, the mass profile splits into two distinct bumps - the strong contribution from the drift term prevents a connected mass from being maintained. In the other three cases a connected mass is maintained and food source aggregation is observed. However, in the two cases where $C = 5$ the degree of aggregation is relatively weak likely due to the stronger contribution from the diffusion term.

Next, we examine the case where the kernel is polynomial and given by $W(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Intuitively, at the particle level, this means that this kernel is attractive at long-ranges but includes a short-range repulsion to prevent particles from becoming too close. The results can be seen in Figure 31. Here, we observe results that are qualitatively very similarly to the results obtained in the quadratic case. This is not surprising given that the polynomial kernel is largely attractive and at long ranges the degree of attraction is proportional to the separation between particles - analogous to the quadratic kernel. However in the case where

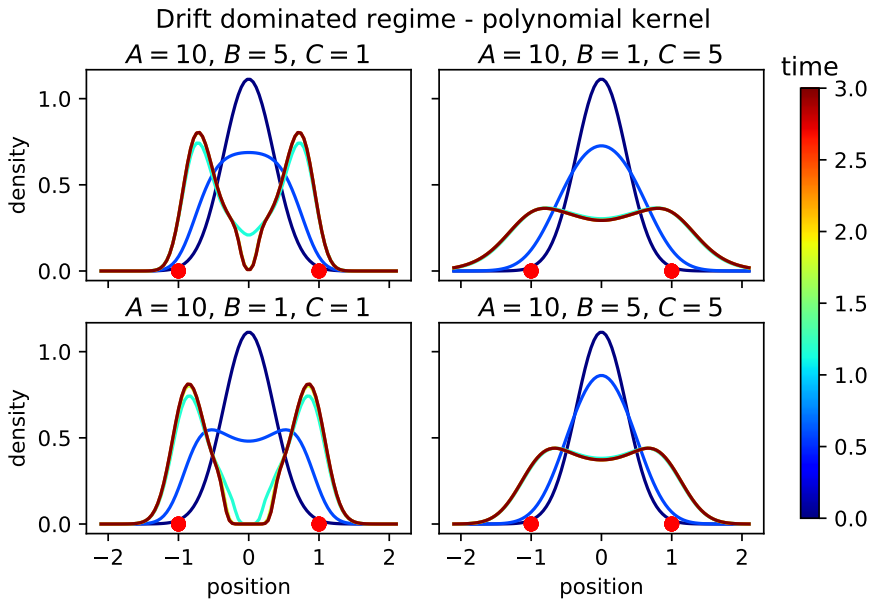

Figure 31: The evolution (4.17) in the drift dominated regime with kernel given by $W(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Strong food source aggregation is observed in two cases however neither maintains a connected mass. Weak food source aggregation is observed in the other cases.

$A = 10, B = 5, C = 1$ the evolution with a polynomial kernel does not maintain a connected mass in contrast to the evolution witha quadratic kernel. This is likely due to the short-range repulsion that is present in the polynomial kernel. At the particle level, particles in the middle of the domain are repelled away from each otherwhile simultaneously being pulled by a food-source - encouraging a separation.

Finally, we examine the case where the kernel is Gaussian and given by $W(x) = \varphi_{.99}(x) = \frac{1}{(4\pi \cdot .99^2)^{1/2}} e^{-x^2/4\cdot .99^2}$. Here, similarly to the quadratic kernel the interaction is purely attractive. However, unlike the polynomial the strength of attraction between two particles is *not* proportional to their separation. Instead, attraction becomes stronger (but not in an unbounded fashion) as two particles become closer and weaker as they become farther apart. This encodes the modeling assumption that interactions should be "local" - particles who

are physically distant from each othershould not interact strongly. The results can be seen in Figure 32.

Here, we observe food source aggregation in every case. similarly to the polynomial and quadratic kernels the aggregation is stronger when the contribution from the diffusion term is small. However, similarly to the polynomial kernel a connected mass is not maintained in either of these cases. Additionally, in the cases where a connected mass is maintained we still observe a sharp drop off in the middle of the profile - indicating formation of a separation. We believe that the tendency to separate in this case is caused by a combination of the strong local interaction and the strong contribution from the food sources; particles that start close to a food source are "trapped" and remain close to the food source as the contribution from the food source and the local interaction reinforce each other. Then, particles towards the center of the domain feel a pull towards food sources form both the food source itself and the "trapped" particles.

4.3.2 Interaction dominated regime

We now examine simulations of (4.17) in the regime where the interaction term is the strongest contribution. Again, we start with the case where the kernel function W is quadratic and given by $W(x) = \frac{x^2}{2}$. The results can be seen in Figure 33. Here, we do not see any food source aggregation. This can possibly be explained by the fact that in these simulations the initial profile is Gaussian. We have previously seen that Gaussians are stationary states for (4.17) in the case of a quadratic kernel and no drift term.

Next we examine the case where the kernel function is polynomial, of the same form as in the food source dominated regime. The results can be seen in Figure 34. Here, analogous

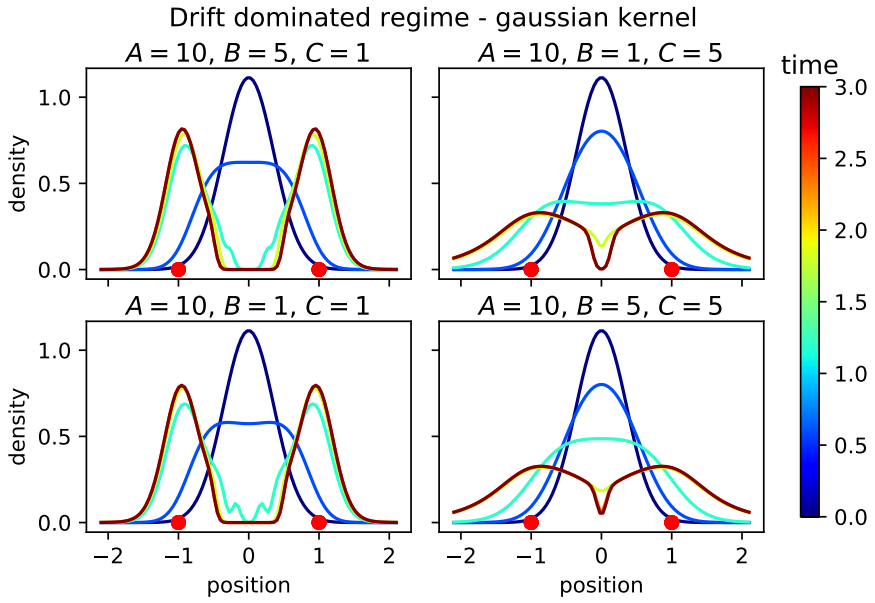

Figure 32: The evolution of (4.17) with kernel given by $W(x) = \varphi_{.99}(x) = \frac{1}{(4\pi\cdot.99^2)^{1/2}} e^{-|x|^2/4\cdot.99^2}$. Food source aggregation occurs in all cases, however in two cases a connected mass is not maintained.

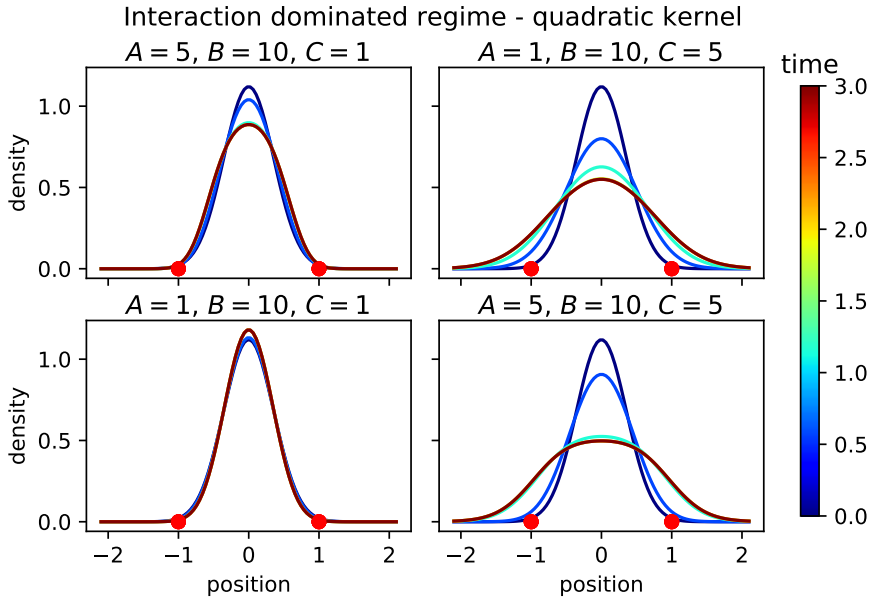

Figure 33: The evolution of (4.17) in the interaction dominated regime with kernel given by $W(x) = \frac{x^2}{2}$. Food source aggregation is not observed.

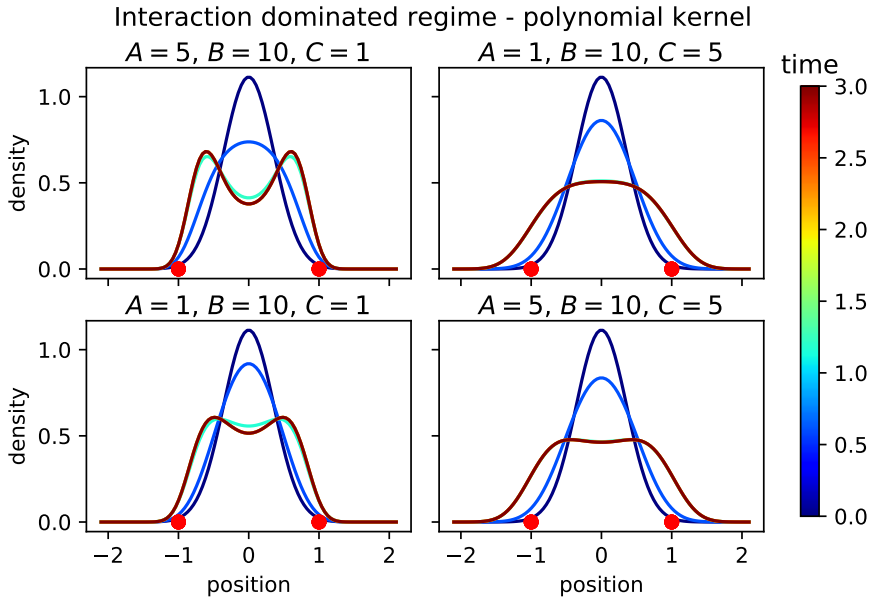

Figure 34: The evolution (4.17) in the interaction dominated regime with kernel given by $W(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Food source aggregation while maintaining a connected mass is observed in two cases.

to the food source dominated regime we see more pronounced aggregation in the cases where the diffusion term has the weakest contribution. However the aggregation is less pronounced than the respective examples in the food source dominated regime due to the stronger contribution from the interaction term. Unlike the analogous cases in the drift dominated regime we do not observe food source aggregation in the cases where $C = 5$, this is possibly due to the stronger contribution from the short range repulsion in the interaction term.

Finally, we examine the evolution in the case of a Gaussian kernel function. The results can be seen in Figure 35. Here, in contrast to the respective cases in the drift dominated regime, we do not observe food source aggregation in the cases where the diffusion term has the second largest contribution. Addtionally, unlike the analogous cases in the drift dominated regime and the polynomial kernel in the interaction dominated regime, we do not observe aggregation in the case where $B = 10, C = 1, A = 1$. We do observe aggregation in the case where $A = 5, B = 10, C = 1$. However, similarly to the polynomial kernel the effect is less pronounced than in the analogous case in the drift dominated regime. Both of these effects are consistent with a stronger contribution from the attractive kernel.

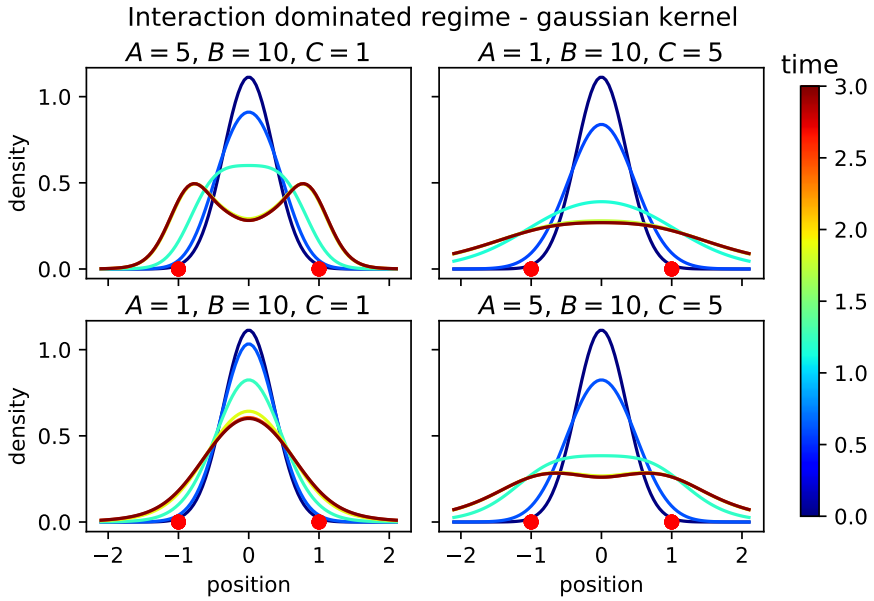

Figure 35: The evolution (4.17) in the interaction dominated regime with kernel given by $W(x) = \varphi_{.99}(x) = \frac{1}{(4\pi\cdot.99^2)^{1/2}} e^{-|x|^2/4\cdot.99^2}$. Food source aggregation while maintaining a connected mass is observed in one case.

4.3.3 Competition regime

We now simulate cases where two parameters dominate in competition. The results can be seen in Figure 36, Figure 37, and Figure 38. Interestingly, across all three interaction kernels, pronounced food source aggregation is only observed in cases where $A = B = 10$. All of these cases except one exhibit the qualitative features we wish to capture as they do maintain a connected mass. All other cases do not result in food source aggregation. This can be explained by the fact that in all of these cases the diffusion term has a significant contribution. The observation that strong contributions from the diffusion term dampen food source agregation is a general pattern that we have also observed in the other regimes and is additionally supported by the statement of Theorem 8.

The cases in the competition regime where we do see food source aggregation also reinforce another pattern observed across all regimes; examples that have the qualitative

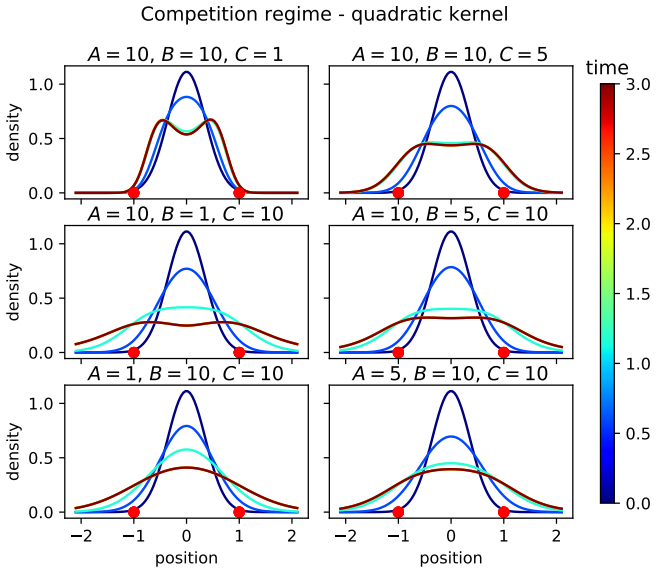

Figure 36: The evolution of (4.17) in the competition regime with kernel given by $W(x) = \frac{x^2}{2}$. Food source aggregation is only observed in cases with a week diffusion contribution.

features we wish to capture have strong contributions from the drift term, the interaction term, or both. However, we also find that cases in which a connected mass was not maintained only occured where the drift term was dominant. Likewise, cases in which food source aggregation did not occur in the absence of strong diffusion only occured in the interaction dominated regime. These observations suggest a general modeling principal; contributions from the drift and diffusion terms should be scaled similarly and larger than the contribution from the diffusion term.

4.3.4 Evidence for asymptotic states and a two dimensional example

We now turn to some further numerical studies aimed at validating observations made in the previous section. First, note that in the previous section we used a uniform simulation

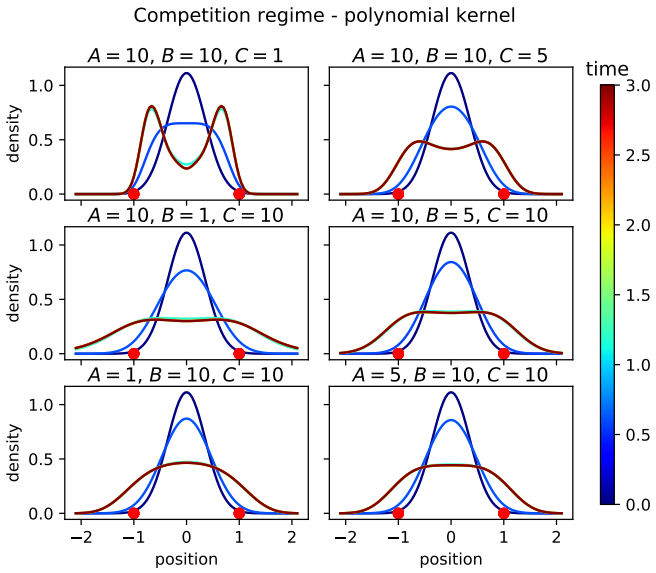

Figure 37: The evolution of (4.17) in the competition regime with kernel given by $W(x) = \frac{x^4}{4} - \frac{x^2}{2}$. Food source aggregation is only observed in cases with a weak diffusion contribution.

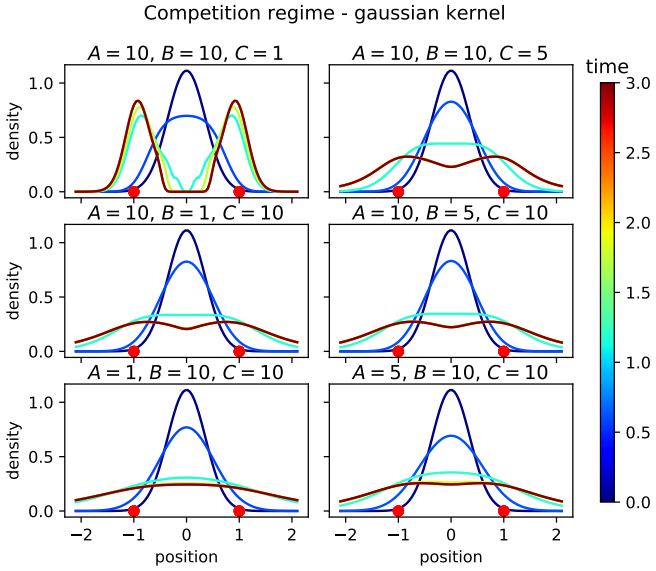

Figure 38: The evolution of (4.17) in the competition regime with kernel given by $W(x) = \frac{1}{(4\pi\cdot.99^2)^{1/2}} e^{-|x|^2/4\cdot.99^2}$. Food source aggregation is only observed in cases with a weak diffusion contribution.

time. Therefore, in theory some of the qualitative observations we have made could merely be artifacts of not simulating the dynamics for long enough; its possible that we are not observing asymptotic states of the dynamics. For example, in Figure 30 we observe what appears to be food source aggregation with a connected mass in the regime $A = 10, B = 5, C = 1$ and mass separation in the regime $A = 10, B = 1, C = 1$ suggesting that the stronger interaction term in the first case helps the profile maintain connectivity. However, if we are not truly observing asymptotic states its possible that we *would* observe mass separation if we waited a longer time. In Figure 39 we perform the same simulation as Figure 30 except the simulation time is doubled and we plot six profiles instead of four. However, most of the six profiles are not distinguishable as they are overlapping in exactly the same profiles as the final profiles plotted in Figure 30. Mass aggregation with a connected profile is observed in exactly analogous fashions in the regime $A = 10, B = 5, C = 1$ and mass separation is observed in both cases in the regime $A = 10, B = 1, C = 1$. This provides strong evidence that we are observing asymptotic states, however to provide further support in this direction (and to show that this phenomena does not depend on the kernel choice) we repeat all simulations of the polynomial kernel with the longer simulation time in Figures 40, 41 and 42. We again observe exactly analogous results to the shorter simulation time in all cases providing further support that asymptotic states are observed.

The aim of this study was to identify parameter regimes in which the derived macroscopic model qualitatively displays slime mold food seeking behavior. For this reason we have limited our simulations to one dimension as they are faster to compute and if a phenomena does not appear in one dimension it will surely not manifest in two. However, for actual application to slime mold movement (and to possibly validate the model with empirical

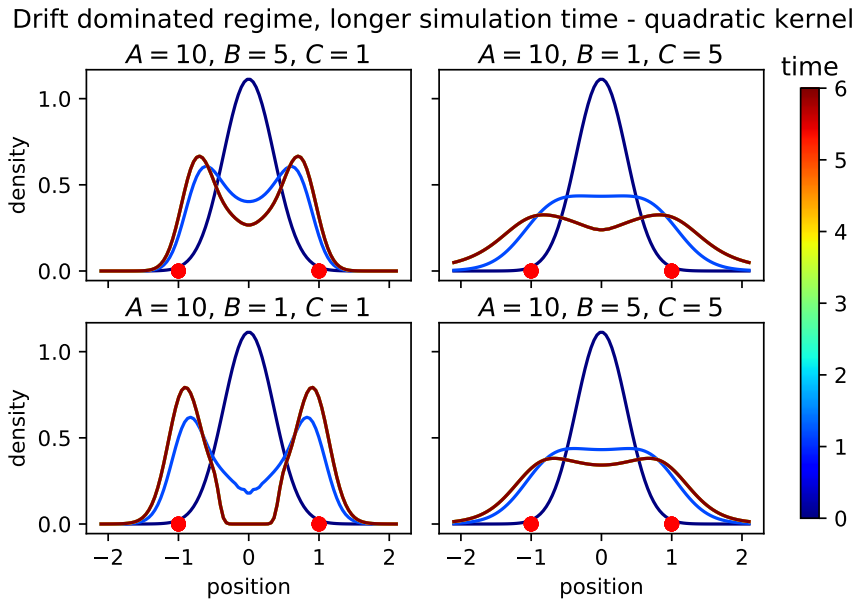

Figure 39: The evolution of (4.17) in the drift dominated regime with the quadratic kernel. The simulation time is doubled from Figure 30 and six profiles are plotted instead of four. The same qualitative phenomena are exhibited providing evidence for the observation of asymptotic states.

experiments) two dimensions is surely the case of interest. To illustrate that observations made in one dimension do manifest in two dimensions we will show that mass separation observed in one dimension is also observed in two dimensions as we feel that this phenomena is easiest to qualitatively evaluate in the two dimensional plots. We simulate the macroscopic model in two dimensions with the gaussian kernel in the regime $A = 10, B = 5, C = 1$ - the results can be found in Figure 43. In this case the dynamics were simulated in the region $[-2.1, 2.1] \times [-2.1, 2.1]$, the food sources are located at the points $(1, 0)$ and $(-1, 0)$, 225 particles were used in the blob method ODE, and the initial profile is given by the two dimensional gaussian centered at the origin with the same variance as that used in one dimension. Analogous to the one dimensional case we observe mass separation.

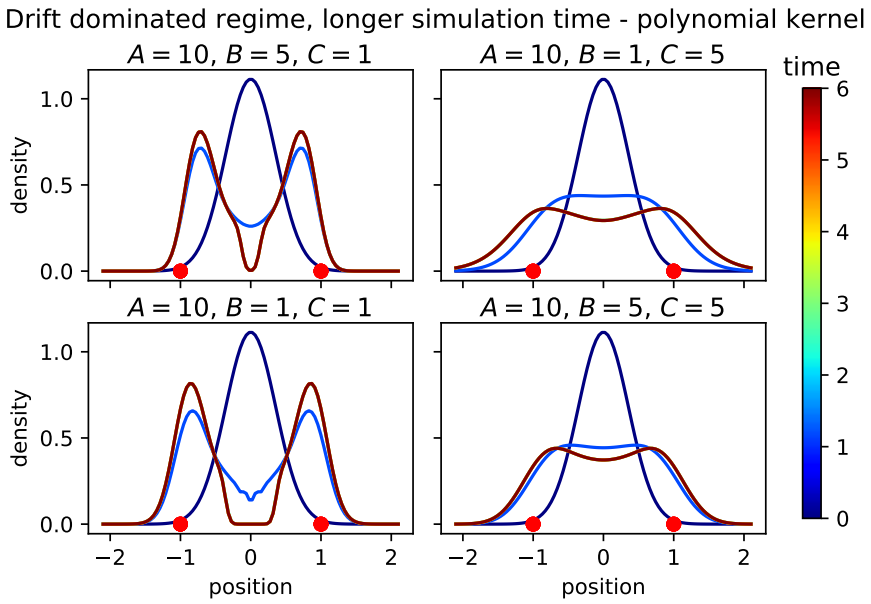

Figure 40: The evolution of (4.17) in the drift dominated regime with the polynomial kernel. The simulation time is doubled and six profiles are plotted instead of four. The same qualitative phenomena are exhibited providing evidence for the observation of asymptotic states.

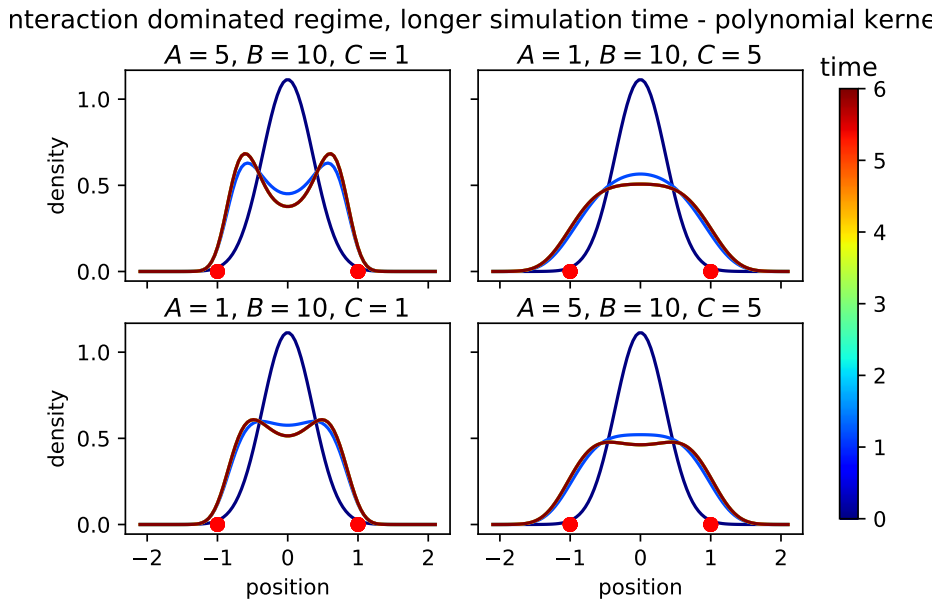

Figure 41: The evolution of (4.17) in the interaction dominated regime with the polynomial kernel. The simulation time is doubled and six profiles are plotted instead of four. The same qualitative phenomena are exhibited providing evidence for the observation of asymptotic states.

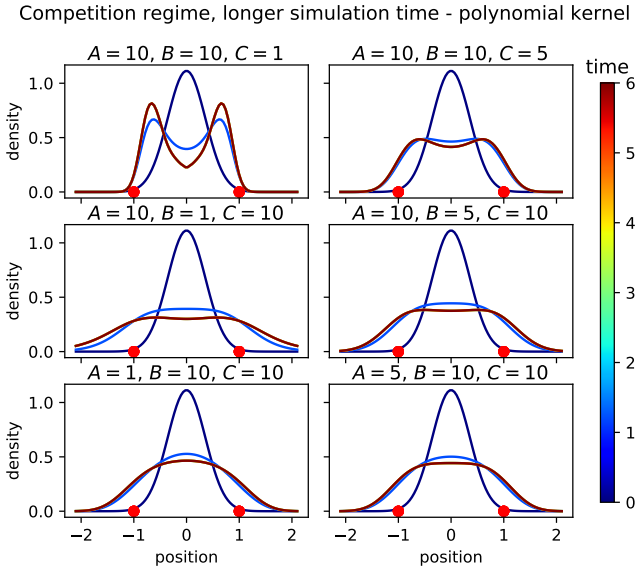

Figure 42: The evolution of (4.17) in the competition regime with the polynomial kernel. The simulation time is doubled and six profiles are plotted instead of four. The same qualitative phenomena are exhibited providing evidence for the observation of asymptotic states.

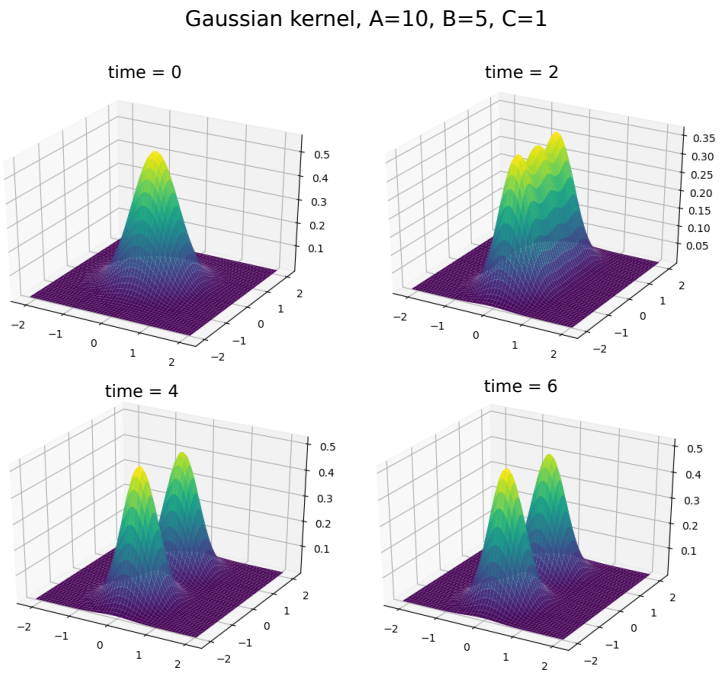

Figure 43: The evolution of (4.17) in a drift dominated regime in two dimensions. Similarly to the one dimensional case we observe mass separation

Chapter 5

FUTURE AND ONGOING WORK

5.1 The Doppelgänger model

Our first aim in extending the work previously described is to introduce models of opinion formation aimed at a better understanding of the polarization of opinions in addition to consensus. As described in the previous section, most analysis of opinion formation models has been conducted on models where only attractive forces are present among the agents. This causes consensus to be a general phenomena. While the use social media has increased connectivity among individuals in a sense, opinions have become more polarized [49, 102, 8, 46, 110, 84]. This phenomenon is usually attributed to the so-called 'echo-chamber' effect. We claim that mathematical models could help better understand the fine-scale details of this dynamic and as a consequence could lead to new strategies to mitigate the effects of polarization.

5.1.1 Doppelgänger dynamics

We aim at introducing repulsion in models of opinion formation. We first aim at extending the linear consensus dynamics (2.1) to include a dynamic of disagreement; specifically we aim to include the possibility of edges in the interaction network with negative weights. Denoting as before the opinion $\mathbf{s}_i(t)$ for $1 \leq i \leq N$, we consider an adjacency matrix A with weights a_{ij} that could be positive and negative. As such the consensus model might lead to an unbounded solution - the difference $|\mathbf{s}_j - \mathbf{s}_i|$ could grow to

infinity when $a_{ij} < 0$. Rather than being repulsed, we propose dynamics where the opinion $\mathbf{s}_i$ is actually attracted to the opposite opinion of $\mathbf{s}_j$, i.e. $-\mathbf{s}_j$.

Definition 15 (pseudo-linear Doppelgänger) *Given a collection of N agents, let $s_i \in \mathbb{R}$ represent the opinion of the i^{th} agent. The **pseudo-linear Doppelgänger model** is defined by the dynamics*

$$\mathbf{s}_i' = \sum_{j=1}^{N} a_{ij}\mathbf{s}_j - \|S\|\mathbf{s}_i. \tag{5.1}$$

with $\|S\| = \|(\mathbf{s}_1, \ldots, \mathbf{s}_N)\|$ the norm of the vector of all opinions.

When $a_{ij} > 0$, opinion $\mathbf{s}_i$ is attracted to $\mathbf{s}_j$ whereas when $a_{ij} < 0$ then $\mathbf{s}_i$ is attracted to $-\mathbf{s}_j$. We dub the dynamics (5.2) the Doppelgänger model as each opinion $\mathbf{s}_i$ has a mirror opinion $-\mathbf{s}_i$ that has an opposite effect. In contrast to the consensus model, the pseudo-linear Doppelgänger model does not necessarily converge and consensus is not always stable. We illustrate the pseudo-linear Doppelgänger dynamics numerically in Figure 44-left with $N = 20$ opinions. Notice the solution does not converge to a consensus but rather to an eigenvector of the adjacency matrix A.

The Doppelgänger model is naturally extended in a fully non-linear version by making the weights $a_{i,j}$ of the adjacency matrix A dependent on the opinions. A natural choice is to assert that the weight a_{ij} is a function of the difference of opinion analogous to models studied in earlier chapters.

Definition 16 (The Doppelgänger model) *Given a collection of N agents, let $s_i \in \mathbb{R}$ represent the opinion of the i^{th} agent. The **Doppelgänger model** is defined by the dynamics*

$$\mathbf{s}_i' = \sum_{j=1}^{N} a_{ij}\mathbf{s}_j - \|S\|\mathbf{s}_i. \tag{5.2}$$

where

$$a_{ij} = \frac{\phi(|\mathbf{s}_j - \mathbf{s_i}|)}{\sum_k |\phi(|\mathbf{s_k} - \mathbf{s_i}|)|} \tag{5.3}$$

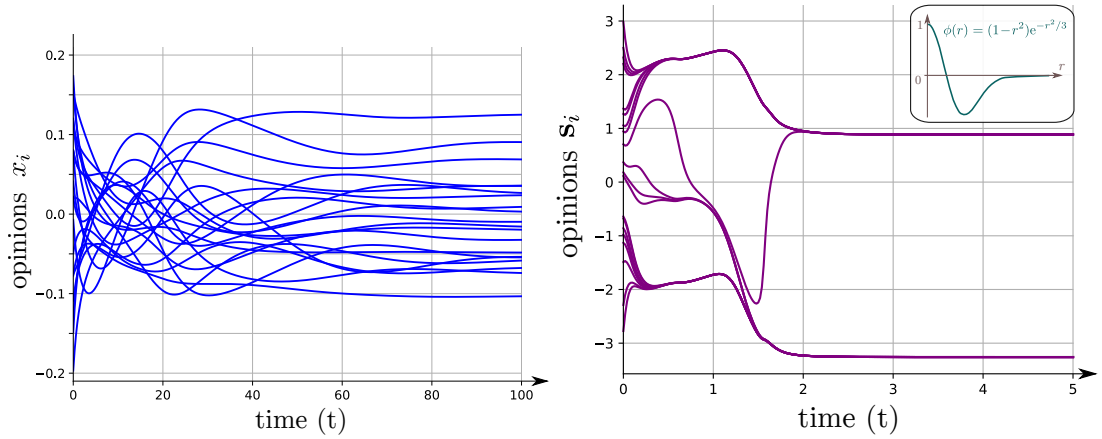

Figure 44: Left: Evolution of the opinions $\mathbf{s}_i(t)$ for the Doppelgänger dynamics (5.1) for a fixed adjacency matrix $A = (a_{ij})$. The dynamics converge to an eigenvector of A. **Right**: Doppelgänger dynamics with evolving adjacency matrix (5.2). The dynamics converge to a polarized configuration.

for some $\phi : \mathbb{R} \to \mathbb{R}$.

Notice that the interaction function ϕ is not always positive. We illustrate the non-linear Doppelgänger dynamics in Figure 44-right. As long as the function ϕ remains bounded, we still have that the opinions $\mathbf{s}_i(t)$ remains bounded using the inequality (5.5). The convergence of the dynamics toward an equilibrium is more delicate, the dynamics could very well oscillate. The investigation of the dynamics will likely mainly focus on the interplay between the support of the function ϕ, the spectrum of the corresponding adjacency matrix A and the convex hull of the configuration of opinions. For example, for the sample choice of interaction function above the model displays a "consensus trap" - if the convex hull of the initial configuration is sufficiently small (such that the difference between all agent opinions lies in the set of the domain of ϕ where ϕ is positive) then consensus is the only outcome possible. Mathematically, it will be challenging to characterize the convergence of the dynamics depending on the matrix A (in the case of the pseudo linear model) or the interaction function ϕ (in the case of the fully nonlinear model). Nonetheless,

it is clear that the solution $\mathbf{s}(t)$ of the Doppelgänger dynamics remains bounded thanks to the following result.

Proposition 9 *If ϕ is bounded then the Doppelgänger dynamics remain bounded.*

Proof

We denote by the matrix $A(t) = (a_{ij}(t))_{ij}$ where:

$$a_{ij} = \frac{\phi(|\mathbf{s}_j - \mathbf{s_i}|)}{\sum_k |\phi(|\mathbf{s_k} - \mathbf{s_i}|)|}. \tag{5.4}$$

As ϕ is bounded by assumption there must exist C such that $\left\|A(t)\right\| \leq C$ for any t. Therefore we have that

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|S\|^2 = \langle AS, S\rangle - \|S\|^3 &\leq \|A(t)\| \cdot \|S\|^2 - \|S\|^3 \\ &\leq C \cdot \|S\|^2 - \|S\|^3 \end{aligned} \tag{5.5}$$

where $\langle , \rangle$ is a standard scalar product and $\|A\|$ the ℓ^2 matrix norm of A. Thus, the norm is decaying when it is sufficiently large. ■

5.1.2 Convergence of the pseudo-linear Doppelgänger model

We now turn our attention to the pseudo-linear Doppelgänger model (5.1). Here, similar to the consensus model, the interaction network is static and we are interested in characterizing the convergence of the model in terms of the adjacency matrix of the interaction network, A.

Theorem 9 *Assume that $\{\lambda_j\}_j$ are the eigenvalues of A indexed such that for any j we have that $Re(\lambda_i) \leq Re(\lambda_j)$ for any $i \leq j$. Then, if $\mathbf{s}(t)$ of the pseudo-linear Doppelgänger dynamics (5.1), $\lambda_N > 0$, and $Re(\lambda_N) > Re(\lambda_{N-1})$ we must have that*

$$\mathbf{s}(t) \to \lambda_N \mu_N \text{ as } t \to \infty$$

where μ_N is the eigenvector of A associated to λ_N.

Proof

For clarity we will work under the assumption that A is diagonalizable. An exactly analogous argument proves the result in the general case via representing A in its Jordan canonical form. In other words, there exists P such that:

$$A = PDP^{-1} \text{ where } D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \text{ and } P = \begin{pmatrix} | & & | \\ \mu_1 & \dots & \mu_2 \\ | & & | \end{pmatrix}$$

then, if $\mathbf{y} = P^{-1}\mathbf{x}$, then the dynamics (5.1) become:

$$\mathbf{y}'(t) = D\mathbf{y} - \|\mathbf{s}\|\mathbf{y}. \tag{5.6}$$

Therefore, it is sufficient to show that $y_i(t) \to 0$ for all $i < N$ and that $y_N(t) \to \lambda_N$. Now, the solution, $\mathbf{y}$ is given by:

$$\begin{aligned} \mathbf{y}(t) &= e^{\int_0^t (D - \|\mathbf{s}(s)\| \mathrm{Id}) ds} \mathbf{y}(0) \\ &= e^{Dt} e^{-t\alpha(t)} \mathbf{y}(0) \text{ where } \alpha(t) := \frac{1}{t} \int_0^t \left\|\mathbf{x}(s)\right\| ds \end{aligned} \tag{5.7}$$

Now, by Proposition 9, we have that for any t, $\left\|\mathbf{y}(t)\right\| \leq C$ for some C. Therefore, by (5.7), we have:

$$e^{-t\alpha(t)} \left\| e^{Dt} \mathbf{y}(0) \right\| \leq C \tag{5.8}$$

additionally we have that:

$$\left\| e^{Dt} \mathbf{y}(0) \right\| \geq e^{\lambda_N t} |y_N(0)|$$

which by (5.8) implies that

$$\alpha(t) \geq \lambda_N - \frac{\ln(\frac{C}{|y_N(0)|})}{t}$$

So, for any $k < N$ there exists T_k such that for all $t \geq T_k$ we have that $\alpha(t) > \mathrm{Re}(\lambda_k)$ by assumption. So, by (5.6), we have that

$$y_k(t) = e^{(\lambda_k - \alpha(t))t} y_k(0) \to 0 \text{ for any } k < N. \tag{5.9}$$

It remains to show that $y_N \to \lambda_N$. By (5.9) we have that

$$\left\|\mathbf{s}(t)\right\| = |y_N(t)| + \epsilon(t) \tag{5.10}$$

where $\epsilon(t) \to 0$. Therefore we have by (5.6) that the evolution of y_N satisfies:

$$y_N' = (\lambda_N - y_N)y_N + \epsilon(t)y_N. \tag{5.11}$$

Consider the solution, $\tilde{y}$, to the following initial value problem

$$\tilde{y}' = (\lambda_N - \tilde{y})\tilde{y}. \tag{5.12}$$

Clearly, for any choice of $\tilde{y}(0)$ we have that $\tilde{y} \to \lambda_N$. We claim that $y_N \to \tilde{y}$ and therefore y_N converges to λ_N as well. First, notice that if $|y_N|$ is sufficiently small then we must have $|y_N|' > 0$. Therefore, we may assume that y_N is positive without loss of generality. Now, we have that:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}[(y_N - \tilde{y})^2] &= 2(y - \tilde{y})(y' - \tilde{y}') \\ &= 2(y_N - \tilde{y})^2(\lambda_N - (\tilde{y} + y_N)) + 2(y - \tilde{y})\epsilon(t)y_N \\ &:= 2(y_N - \tilde{y})^2\gamma(t) + \rho(t).\end{aligned} \tag{5.13}$$

As y_N and $\tilde{y}$ are bounded and $\epsilon(t) \to 0$ we have that $\rho(t) \to 0$. Additionally, as y_N is positive and bounded away from 0 and $\tilde{y} \to \lambda_N$ we have that there exists $C < 0$ and T such that for all $t > T$, $\gamma(t) < C$. Therefore, we deduce by Gronwall's inequality that $y_N \to \tilde{y}$ and therefore $y_N \to \lambda_N$ as desired. ■

5.1.3 Generalizations and extensions

In the non-linear Doppelgänger dynamics (5.1)(5.2), the graph depends only on the difference of opinions $\mathbf{s}_j - \mathbf{s}_i$. However, extreme opinions tend to be less inclined to listen to

other opinions. Thus, it is relevant to adjust the range of interaction depending on opinion. For instance, replacing the difference $\mathbf{s}_j - \mathbf{s}_i$ with the relative difference $(\mathbf{s}_j - \mathbf{s}_i)/|\mathbf{s}_i|$.

The Doppelgänger dynamics proposed in (5.2) is a deterministic model. It would be interesting to define a stochastic analogue and study the correspondences between the deterministic and stochastic framework, much like the correspondence between the linear voter model and the classical consensus model as in [119]. As a possible stochastic analogue we propose that rather than changing the opinion of node i, $\mathbf{s}_i$ gradually to its neighbor's opinion $\mathbf{s}_j$, the stochastic Doppelgänger model could be written as a jump process where opinion $S_i(t,\omega)$ would jump toward $S_j(t,\omega)$ with intensity a_{ij}. If the coefficient a_{ij} is negative, then X_i would jump to opinion $-S_j$ with intensity $|a_{ij}|$. Such stochastic dynamics lead to the following formulation:

$$\mathrm{d}\mathbb{E}[\varphi(S_i(t))] = \sum_{j=1}^{N} |a_{ij}| \cdot \mathbb{E}[\varphi\Big(\mathrm{sign}(a_{ij})S_j(t-)\Big) - \varphi(S_i(t-))]\mathrm{d}t. \tag{5.14}$$

These dynamics necessarily converge as the number of possible opinions is decreasing in each time step. Additionally it could be proven by a martingale argument ($S_i(t)$ remaining uniformly bounded in time). What are conditions that cause the limit configuration to be a consensus or a polarized configuration? Are these conditions comparable to the deterministic model? Can we recover the deterministic model from the stochastic model in any sense? If A is positive and the underlying graph remains connected, then a consensus occurs as proven in [119]. Including negative weights makes the analysis more challenging.

The Twitter dataset to which we have access is very large - the number of individual users is in the millions. Therefore it is relevant to study the Doppelgänger dynamics from a mean field approach [12, 97, 38, 37]. This could present new avenues of analysis through the application of the theory of Partial Differential Equations (PDEs). Denoting $\rho(s,t)$ the density of agents with opinion s at time t, the non-linear Doppelgänger dynamics leads to a

transport equation of the form:

$$\partial_t \rho + \nabla_s \cdot (G[\rho]\rho) = 0 \quad , \quad G[\rho](s) = \int_{y \in \mathbb{R}^d} \phi(|y-s|) y \rho(y)\, \mathrm{d}y \; - \; \left(\int_{y \in \mathbb{R}^d} |y|^2 \rho(y)\, \mathrm{d}y \right) s. \quad (5.15)$$

The completely rigorous derivation of such a PDE from the interacting particle systems (5.2) is challenging as it requires the proof of so-called propagation of chaos [111, 83, 17, 27]. In the context of deterministic dynamics, it boils down to proving the regularity of the transport equation (5.15).

5.2 Future work: Empirical social media data - sentiment analysis and graph clustering

A main takeaway from much of the work previously described is that the interplay between the interaction network and evolution of opinions is central; for example, in bounded confidence dynamics clustering of opinions is equivalent to a fragmentation of the interaction network. We propose a framework to investigate whether such a feedback loop can be observed in real social media data.

Critical to these goals is a large source of social media data. We have access to a large dataset that centers around the gun-control debate in the United States. Gun Debate dataset contains nearly 24 million tweets in English about the gun control debate in the USA between Feb 1, 2018 and Aug 23, 2018 following the Stoneman Douglas High School shooting on February 14, 2018. There are two sides of the debate: pro-gun control (blue camp) and anti-gun control (red camp). Tweets are posted by 3,282,592 individual anonymized Twitter accounts. Users are classified according to a content classifier that is trained based on most tweeted user's labeled positions. 293,046 of the users are classified as anti-gun control, while others are classified as pro-gun control. Almost 5 million tweets carry anti-gun control (red) sentiment and the rest are pro-gun control (blue) represented in

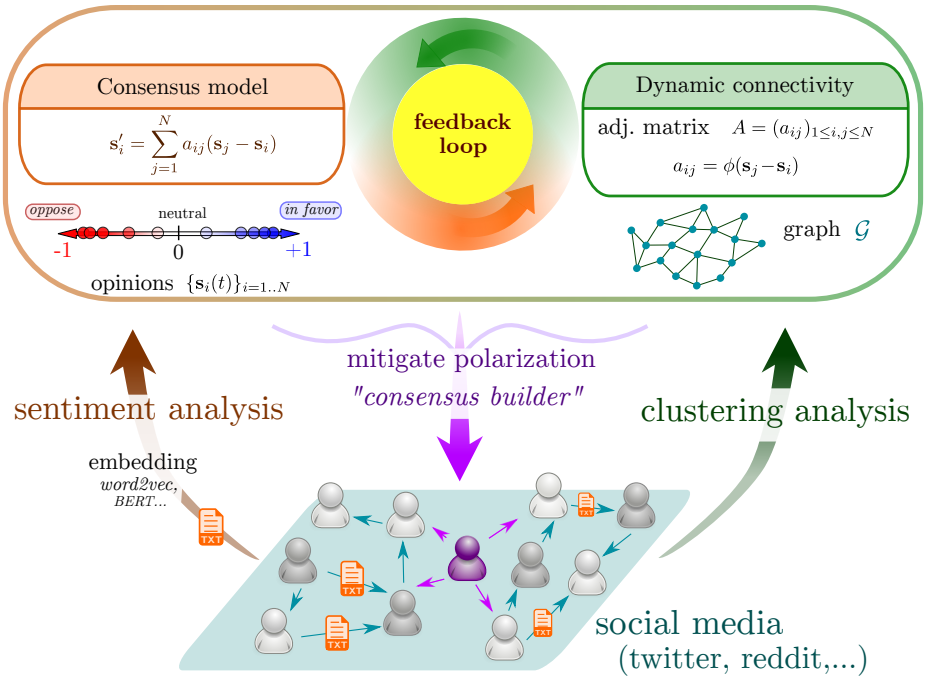

Figure 45: Schematic representation of the work proposed. In analytical work, we study the interplay between the opinions $\{\mathbf{s}_i\}_{1\le i\le N}$ and the interacting graph $\mathcal{G}$ leading to consensus or polarization. In empirical investigations, experimental data are extracted from social media using sentiment analysis (measuring opinions $\mathbf{s}_i$) and graph clustering (extracting the graph $\mathcal{G}$). Finally, empirical data is combined with the analytical framework to investigate strategies to mitigate polarization.

figure 46. Additionally we have access to a dataset that centers around the Brexit debate. UK Brexit Twitter Dataset contains nearly 4 years of Twitter data from UK for the June 1, 2015 - May 12, 2019 period containing the political debates for 2016 United Kingdom European Union membership referendum and the 2017 United Kingdom general election. There are 2.8M Twitters (2,846,262) and 51M (51,862,627) Tweets in this dataset. Both of these datasets are well suited to our goals as they represent issues that are “binary” in a sense; one could place themselves on a linear scale that represents the degree to which they support or oppose gun-control measures or Brexit.

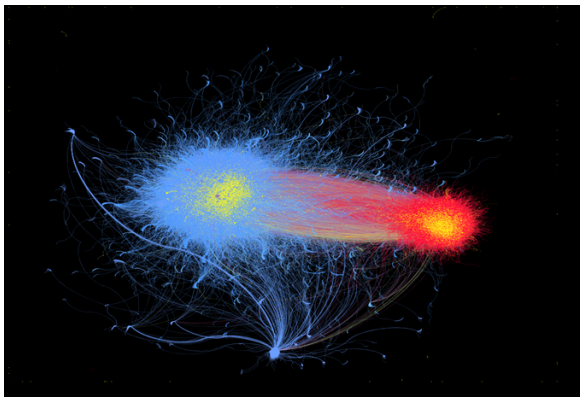

Figure 46: Schematic representation of the gun-control debate in the United States in 2018 using tweets from 293,046 individual anonymized Twitter accounts: pro-gun and anti-gun control are represented in blue and red (respectively).

5.2.1 Measuring opinions $\mathbf{s}_i$

In order to inform modeling pursuits with social media data two quantities must be measured; the structure of the interaction network $\mathcal{G}$ among users and the opinion of those users $\{\mathbf{s}_i\}_{1\leq i\leq N}$ on a given issue. Sentiment analysis aims to measure opinion by training machine learning algorithms to directly classify text generated by users. Given a text written by a user i (e.g. tweet), denoted $\mathcal{T}_i$, the goal is to associate a sentiment $\mathbf{s}_i \in (-1, 1)$ where -1 means 'negative', 0 'neutral' and $+1$ 'positive'. This is accomplished through training a *word embedding* on a *language modeling task* [16, 53, 103, 104, 62, 67]:

$$\mathcal{T}_i \text{ (Text)} \overset{\text{embedding}}{\rightsquigarrow} \mathbf{z}_i \in \mathbb{R}^D \overset{\text{classifier}}{\rightsquigarrow} \mathbf{s}_i \in (-1, 1). \tag{5.16}$$

Learning text representations is an active and open area of research, currently the most successful approach is to allow representations to be context dependent; the representation of a given word changes depending on the words surrounding it in the current training example. The key to this approach is to find a representation of text that preserves its structure and meaning and is amenable to traditional classification techniques.

It has been shown empirically that language models with many parameters trained on

a very large text corpus can be effectively *fine tuned* to a wide array of natural language processing tasks including sentiment analysis [117, 42, 98, 105, 104, 76, 122]. This means that once a sufficiently large language model is trained on a large collection of text it can be further trained to effectively perform sentiment analysis with a relatively small amount of labeled training data (a collection of texts that are labeled as 'positive', negative' or 'neutral').

In order to quantify a user's opinion $\mathbf{s}_i(t)$ on a given issue (e.g. a national election or a noteworthy news event) using raw social media data it is clear that the most relevant part of the data is the text that a given user generates. The issue in question (e.g. Brexit) should be fairly narrowly defined in the sense that one can say that their alignment relative to the issue could be placed on a linear scale $(-1, 1)$ (with -1 meaning "disagree" and $+1$ "agree").

Using sentiment analysis as described in eq. (5.16), each text $\mathcal{T}_i$ generated by a user is assigned a sentiment $\mathbf{s}_i$. The opinion of a user in question can then be computed as the average of the opinion of their texts. In the past sentiment analysis has been done mostly in a static manner - the data remains fixed and one sentiment measurement is done. We propose to analyze the data in a time dependent manner by organizing it into time windows and performing sentiment analysis only on text that occurs in each window (see figure 47). In this way we can uncover the dynamics of changing opinions.

There are several challenges in applying the sentiment analysis pipeline previously discussed. First and foremost to collect a dataset the size of the Brexit dataset necessitates fairly broad scraping parameters. This allows for the possibility of noise in the dataset as it could include tweets that are not on topic. Additionally, the Brexit debate is vast and encompasses many sub issues, it is possible that it will be useful to focus on one or several of these. Both of these problems necessitate developing a topic/entity analysis framework to ensure that the data that is analyzed is as relevant as possible. Second, although we intend

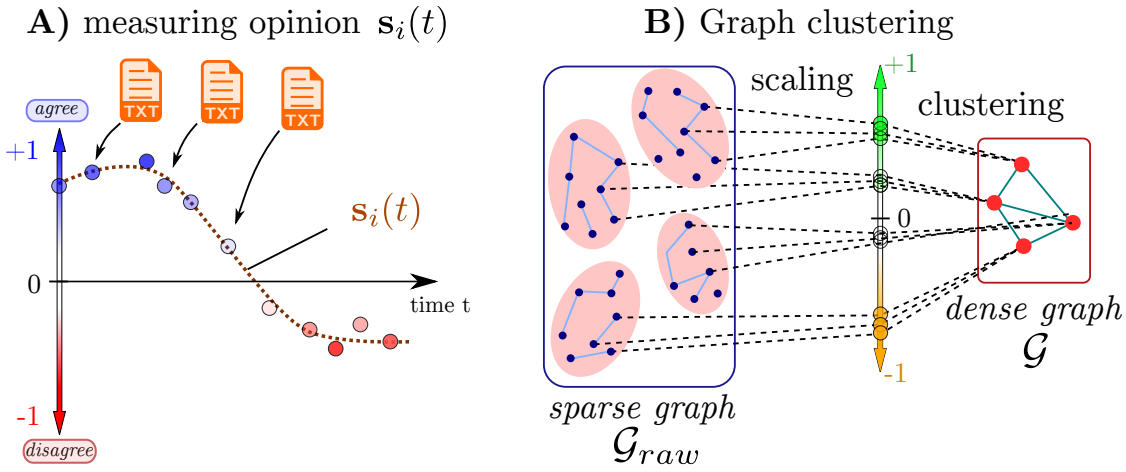

Figure 47: Left: the opinion $\mathbf{s}_i(t)$ of a user i is infered by the text posted online. Using sentiment analysis, a text $\mathcal{T}_i$ (represented in orange) is scaled on a $(-1, 1)$ axis. The opinion $\mathbf{s}_i(t)$ is the time window averaging of these sentiments. **Right**: starting from a sparse graph $\mathcal{G}_{raw}$ of interating users, a dense graph $\mathcal{G}$ is extracted. To do so, each node i is scaled and then a clustering algorithm is performed to create "cluster" node.

to largely use pretrained language models (as is the best practice in the field [105]), these models require fine tuning on the sentiment analysis task. This necessitates the human labeling of a dataset of text (labeling a collection of tweets as pro-Brexit, anti-Brexit, or moderate towards Brexit) or an unsupervised labeling approach. While the amount of labeled data needed is small compared to the original amount of text the language model was pretrained on, this requirement can still be time consuming and challenging. First, care must be taken to ensure the dataset is not biased. Secondly, it must be labeled by as many human volunteers as possible to weed out vague examples.

5.2.2 Extracting graph $\mathcal{G}$

There are many approaches to structuring raw social media data into a network $\mathcal{G}$. Almost all social networks have some sort of friendship mechanic which can be leveraged (although this mechanic is not always mutual, for example in Twitter); an edge exists

between users if they are friends on the platform. Additionally, all social networks have various mechanics to allow users to interact which allows for network structures based on one or several of the different types of interactions. For example, in Twitter there are four main ways for users to interact with each other; they can "follow" each other, they can "like" each other's tweets, they can "retweet" each other's tweets, or they can "reply" to each other's tweets. Any single one of these interactions could be used to define a directed network between a group of users; there is an edge from user i to user j if they interact. In this work, we are particularly interested in "replies" since they indicate, through sentiment analysis, a positive or negative interaction between users, i.e. this definition of interaction leads to an adjacency matrix a_{ij} between users with both positive and negative weights. We denote this *raw* network $\mathcal{G}_{raw}$.

A challenge to this approach is that the network $\mathcal{G}_{raw}$ is usually sparse [41, 10]. Graph clustering can be leveraged to reduce the sparsity while still maintaining the important structures in the network. This is performed in two steps. The first step consists in embedding each node of the graph into a latent space, i.e. a node i is assigned a value $\mathbf{x}_i$. This is the process referred to as scaling [68, 93, 108, 107]. Users who are similar to each other in terms of network structure should be placed geometrically close in the latent space. In the second step, the embedded vectors $\mathbf{x}_i$ are clustered using standard algorithm such as k-means or DBSCAN [1, 7]. Thus, users are placed into discrete clusters. This approach is also often called community detection. Each cluster can can then be regarded as a single node in a "collapsed" network denoted $\mathcal{G}$:

$$i \text{ node of } \mathcal{G}_{raw} \overset{\text{scaling}}{\rightsquigarrow} \mathbf{x}_i \in (-1,1) \overset{\text{clustering}}{\rightsquigarrow} \{C_j\}_j \text{ clusters} \overset{\text{collapse}}{\rightsquigarrow} \mathcal{G}. \tag{5.17}$$

This approach raises a host of research questions. Broadly how does the structure of $\mathcal{G}_{raw}$ depend on the choice of interaction network? Do different choices of interaction network display characteristic structures (eg. the following network in twitter is known

to display a "hub and spoke" structure due to the presence of influencers and celebrities)? How does the choice of scaling algorithm affect the process of collapsing $\mathcal{G}_{raw}$ to $\mathcal{G}$? Are particular algorithms sensitive to network structures that are characteristic of the way we chose to define $\mathcal{G}_{raw}$? Finally, which combination of $\mathcal{G}_{raw}$ definition and scaling algorithm best resolves the dynamics of network fragmentation and coalescence over time (if they exist)?

5.2.3 Uncovering interplay between $\mathbf{s}_i$ and a_{ij}

We now propose a framework to study the interplay between the two measured quantities as it is the key to understand how the distribution of opinions is likely to evolve. For example, in the classical Hegselmann Krause model, fragmentation and coalescence in the opinion distribution is equivalent to fragmentation and coalescence in the interaction network. Can we observe fragmentation and coalescence over time in our measurements of $\mathbf{s}_i$ and a_{ij}? If so, is there a relationship between these phenomena? Does fragmentation or coalescence in the interaction network imply the same in the observed opinion distribution? To study the network's effects on the opinion distribution we propose to examine if and how the change in the difference of opinion between two users who are connected in the interaction network depends on their current difference of opinion. Formally, given two users i and j we look for a relationship ϕ such that:

$$(\mathbf{s}_j - \mathbf{s}_i)' = \phi(\mathbf{s}_j - \mathbf{s}_i). \tag{5.18}$$

To establish the interaction network's influence over the observed opinion distribution we propose to study two quantities. Given that two users i and j are connected in the interaction network how does the likelihood that the link between them is destroyed depend on the current difference between $\mathbf{s}_i$ and $\mathbf{s}_j$? Conversely, given two users that are disconnected in

the interaction network how does the likelihood that a link between them is created depend on the current difference between $\mathbf{s}_i$ and $\mathbf{s}_j$? In this way we can determine whether there is a feedback loop of influence between the opinion distribution and network structure as in the models, or if one has a more dominating effect.

5.3 Future work: mitigating polarization

Learning how opinions become polarized could help us find a strategy to mitigate this tendency. Indeed, understanding of the feedback loop between opinion and graph could lead to new strategy to modify the time evolution. Mathematically, this leads to an optimal control problem similar to the escaping problem in pedestrian dynamics or flocking models [100, 86, 23, 47]. There is however a notable difference: our goal is not to steer opinions in one way or another but rather to maintain interaction among users, i.e. keep the graph $\mathcal{G}$ connected.

As a first "toy problem", we consider the linear consensus model (2.1) and ask where to allocate new moderators, denoted $\mathbf{m}_k$, to enforce the emergence of a consensus (see Figure 48). We illustrate the problem using only one moderator for simplicity but the problem can be generalized for more moderators. Consider a moderator $\mathbf{m}_*$ connected with other agents with link α_i (i.e. $\sum_i \alpha_i \leq 1$). The opinion of a moderator is simply the weighted average of its neighbors, i.e. $\mathbf{m}_* = \sum_i \omega_i \mathbf{s}_i$ with $\sum_i \omega_i = 1$. The consensus model becomes:

$$\mathbf{s}_i' = \sum_{j=1}^{N} a_{ij}(\mathbf{s}_j - \mathbf{s}_i) + \alpha_i(\mathbf{m}_* - \mathbf{s}_i) = \sum_{j=1}^{N}(a_{ij} + \alpha_i\omega_j)(\mathbf{s}_j - \mathbf{s}_i). \tag{5.19}$$

Hence, the moderator changes the interaction graph with a new adjacency matrix $\widetilde{A}$ with: $\widetilde{a}_{ij} = a_{ij} + \alpha_i\omega_j$. The goal is therefore to maximize the spectral gap of the new Laplacian

matrix $\widetilde{L}$ leading to a saddle point problem:

$$\max_{\sum_i \alpha_i \le 1,\, \sum_i \omega_i = 1} \lambda_2(\widetilde{L}) = \max_{\sum_i \alpha_i \le 1,\, \sum_i \omega_i = 1} \; \min_{\|\mathbf{S}\|=1, \mathbf{S} \perp \mathbf{1}} \langle \widetilde{L}\mathbf{S}, \mathbf{S} \rangle. \tag{5.20}$$

Solving this eigenvalue problem gives information as to where to allocate the moderator in order to optimally encourage consensus.

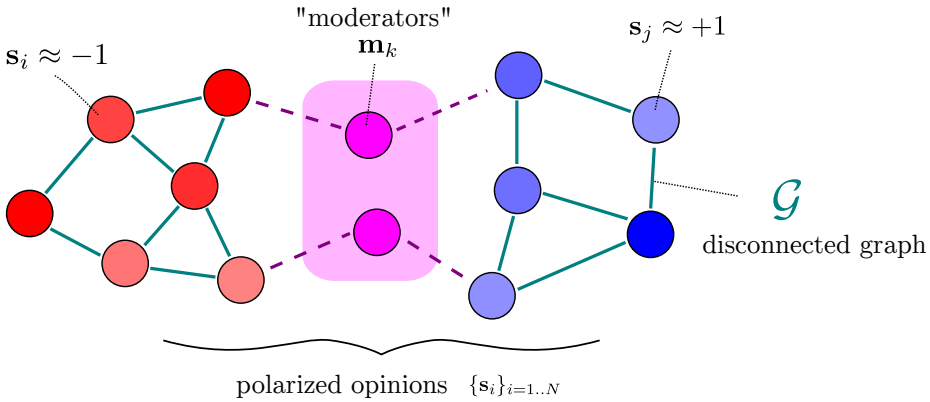

Figure 48: When the graph $\mathcal{G}$ is disconnected, the consensus model (2.1) lead to clustered opinion formation. To mitigate the polarization, "moderators" are introduced to link disconnected clusters of opinions. Mathematically, the challenge is to determine to whom moderators should be linked.

In a non-linear context such as Hegselmann-Krause model (3.1) or the Doppelgänger model (5.2), the optimization problem is more delicate since we cannot directly infer the interaction graph $\mathcal{G}$. Instead, we have to allocate moderators with a given opinion; these moderators act as "informed leaders" such as in a shoal of fish [34] with the goal to ensure the preservation of connectivity. The NOLB dynamics (3.8) previously discussed are aimed precisely at maintaining connectivity by changing the dynamics (i.e. opinions $\mathbf{s}_i$ cannot change if it might disconnected from a neighbor $\mathbf{s}_j$). However, we do not expect to be able to change the dynamics in social-interaction, only to facilitate communication. Instead, we consider the following new dynamics with K mediators with fixed opinions $\mathbf{m}_k$:

$$\mathbf{s}_i' = \sum_{j=1}^{N} a_{ij}(\mathbf{s}_j - \mathbf{s}_i) + \sum_{k=1}^{K} \alpha_{ik}(\mathbf{m}_k - \mathbf{s}_i), \quad a_{ij} = \frac{\Phi_{ij}}{\sum_{\ell=1}^{N} \Phi_{i\ell} + \sum_{k=1}^{K} \widetilde{\Phi}_{ik}} \tag{5.21}$$

with $\Phi_{ij} = \Phi(\left|\mathbf{s}_j - \mathbf{s}_i\right|)$ and $\widetilde{\Phi}_{ik} = \Phi(|\mathbf{m}_k - \mathbf{s}_i|)$. The optimization problem consists of minimizing the variance of opinions over a fixed time horizon T, i.e. minimizing the variance of $\{\mathbf{s}_i(T)\}_{1\leq i\leq N}$.

Chapter 6

CONCLUSION

In this dissertation we presented three contributions to the body of work surrounding network based models of opinion formation. In the first, we discussed consensus models in both a deterministic and stochastic setting. In both cases we completely characterized the convergence of the model. In the deterministic case we find that the model always converges although not necessarily to a consensus; this contrasts with the stochastic case where we find that the model only converges almost surely if and only if the network has only one isolated block. We have also given conditions that are equivalent to unconditional convergence to a consensus - isolated blocks play a critical and identical role in both cases. We find that for both models unconditional convergence to a consensus is equivalent to the network on which the model is defined having only one isolated block. This similarity is explained by the fact that the deterministic dynamics can be recovered from the stochastic dynamics in expectation. Finally, we show in the deterministic case that the rate of convergence depends only on the algebraic connectivity of the network - the stronger the connectivity the faster the model converges. We find a similar result in the stochastic model in the case of symmetric networks however in this case we see that the rate of convergence is much slower as it also depends on the number of agents.

In the second study we examined variants of the Hegselmann-Krause bounded confidence dynamics introduced in [57]. The modifications we introduced were aimed at mitigating the generic cluster forming behavior seen in the bounded confidence dynamics and inducing consensus among the agents. Motivated by the attractive nature of the interaction in bounded confidence dynamics we introduced a variant dubbed *No one left behind* (NOLB) that

maintains connectivity between agents. We rigorously demonstrated that this control is sufficient for unconditional convergence to a consensus regardless of the dimension of agent opinions. Due to the nonlinear and discontinuous nature of the dynamics the argument relies on the interplay between two key properties of the dynamics; contractivity and the preservation of connectivity of the configuration of agents. In one dimension we were able to derive explicit convergence rates that quantify how fast a consensus is reached. Additionally, we conducted numerical experiments that suggest that tighter bounds on the convergence rates are likely not possible. The NOLB dynamics maintain local, pairwise connectivity between agents, however our argument for unconditional convergence to a consensus relied only on global connectivity of the configuration of agents. Motivated by this we introduced a second variant to the bounded confidence dynamics we dubbed *Relaxed no one left behind* (RNOLB) aimed at maintaining the existence of paths between pairs of agents. We find that while this modification still results in unconditional convergence to a consensus, it retains more of the qualitative features of the bounded confidence dynamics; notably the emergence of clusters in the beginning of its evolution. For this reason the RNOLB dynamics can be regarded as an interpolation between the bounded confidence dynamics and the NOLB dynamics. We presented numerical investigations into the variance of agent opinions and a metric we dub the *clustering number* that further support this view.

Finally, in a slight departure from the previous work we presented an investigation of another agent based model inspired by a slimemold. This model shares many features with agent based models of opinion formation and serves as an example of how these models could be extended as well as demonstrates a framework for studying them in a macroscopic setting. We were principally interested in capturing the slime mold's ability to aggregate around disparate food sources while maintaining a connected mass. We first presented a particle based model which includes three main features - a drift term to model

a gradient of chemoattractant produced by food sources, an interaction term to model the slime mold's propensity to maintain a connected mass, and a diffusion term to model slime mold foraging behavior. Additionally, we included scaling parameters for each of these three terms in order to study how varying their contributions could aid in modeling *P. polycephalum*. However simulating enough particles to realistically model the evolution of slime mold would be computationally intractable. Therefore, under the assumption of propagation of chaos we showed that in the large particle limit the evolution of the particle model can be described by a macroscopic aggregation-diffusion equation which can be efficiently simulated. Before embarking on simulating the equation we made some analytical observations via analysis of the equation's stationary states. First, we found that in the case of no drift term that an assumption of Gaussian stationary states implies that the interaction kernel must be quadratic and discussed how this could be used as a heuristic to determine a kernel that realistically models *P. polycephalum* behavior. Then, we showed that if the contribution from the diffusion term is sufficiently larger than the interaction and drift terms that the only possible stationary state is 0, deeming this regime as unsuitable for modeling *P. polycephalum*. We then ran a series of simulations of the macroscopic model in order to investigate how to scale the three contributing terms and choose the interaction kernel in order to model *P. polycephalum* food seeking behavior. We simulated three different parameter regimes; a regime in which the drift term dominated, a regime in which the interaction term dominated and a regime in which the dominating terms were in competition. In each regime we examined three different interaction kernels, an attractive kernel with a local interaction, an attractive kernel with long range interaction, and a kernel that included a short range repulsion. We observed cases that reproduced connected mass food source aggregation across all three regimes with all kernels however several general patterns did emerge. We found that strong contributions from the diffusion term tend to

dampen food source aggregation even in the presence of strong contributions from the other terms. Additionally, we found that in the absence of strong diffusion, dominance of the drift term can result in mass separation while dominance of the interaction term can prevent food aggregation. This suggests a general modeling strategy of scaling contributions from the drift and interaction terms similarly and larger than the diffusion term.

There are many ways in which the work discussed can be extended. First, we propose a class of models aimed at introducing a dynamic of disagreement or repulsion to agent based opinion formation models which previously had mainly only included attractive forces among the agents. We dub this model the Doppelgänger model and demonstrate that the models as defined do manage to include a repulsive dynamic while simultaneously remaining bounded. This observation makes this class of models attractive for further analysis. As a first application of this fact we characterize the convergence of the model in the case of a static interaction network. Secondly, a main takeaway from much of the theory developed in this dissertation is the centrality of the interplay between the interaction network and opinion distribution. We propose an empirical framework aimed at uncovering this interplay in an extremely large social media dataset to which we have access. This framework leverages the techniques of sentiment analysis and graph clustering. Finally, we propose an approach for leveraging modeling efforts to identify strategies to mitigate polarization in social networks. This approach recasts the models into an optimization problem aimed at efficiently distributing moderators with the purpose of maintaining the structure of the interaction network and preventing the formation of echo chambers.

REFERENCES

[1] In:

[2] R. Abelson. *Mathematical models of the distribution of attitudes under controversy*. en. 1964.

[3] A. Adamatzky. "Slime Mold Solves Maze in One Pass, Assisted by Gradient of Chemo-Attractants". In: *IEEE Transactions on NanoBioscience* 11.2 (June 2012), pp. 131–134. DOI: 10.1109/TNB.2011.2181978.

[4] Andrew Adamatzky. "Developing proximity graphs by physarum polycephalum: does the plasmodium follow the toussaint hierarchy?" In: *Parallel Processing Letters* 19.01 (Mar. 2009). DOI: 10.1142/S0129626409000109.

[5] Andrew Adamatzky. "From reaction-diffusion to Physarum computing". en. In: *Natural Computing* 8.3 (Sept. 2009), pp. 431–447. DOI: 10.1007/s11047-009-9120-5.

[6] Andrew Adamatzky and Pedro P.B. de Oliveira. "Brazilian highways from slime mold's point of view". In: *Kybernetes* 40.9/10 (Jan. 2011). DOI: 10.1108/03684921111169440.

[7] Khaled Alsabti, Sanjay Ranka, and Vineet Singh. "An efficient k-means clustering algorithm". In: *Electrical Engineering and Computer Science* (Jan. 1997).

[8] Eytan Bakshy, Solomon Messing, and Lada A. Adamic. "Exposure to ideologically diverse news and opinion on Facebook". en. In: *Science* 348.6239 (June 2015). DOI: 10.1126/science.aaa1160.

[9] Michele Ballerini et al. "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study". In: *Proceedings of the national academy of sciences* 105.4 (2008), pp. 1232–1237. DOI: https://doi.org/10.1073/pnas.0711437105.

[10] Albert-László Barabási and Eric Bonabeau. "Scale-free networks". In: *Scientific American* 288.5 (2003), pp. 60–69.

[11] Werner Baumgarten, Tetsuo Ueda, and Marcus J. B. Hauser. "Plasmodial vein networks of the slime mold Physarum polycephalum form regular graphs". In: *Physical Review E* 82.4 (Oct. 2010). DOI: 10.1103/PhysRevE.82.046113.

[12] Nicola Bellomo and Mario Pulvirenti. *Modeling in applied sciences: A kinetic theory approach*. Springer Science & Business Media, 2013.

[13] E. Ben-Naim et al. "Unity and discord in opinion dynamics". In: *Physica A: Statistical Mechanics and its Applications* 330.1 (2003), pp. 99–106.

[14] V. Blondel, J. Hendrickx, and J. Tsitsiklis. "On Krause's multi-agent consensus model with state-dependent connectivity". In: *IEEE transactions on Automatic Control* 54.11 (2009), pp. 2586–2597.

[15] Vincent D. Blondel, Julien M. Hendrickx, and John N. Tsitsiklis. "On the 2R conjecture for multi-agent systems". In: *2007 European Control Conference (ECC)*. July 2007, pp. 874–881. DOI: 10.23919/ECC.2007.7068885.

[16] Piotr Bojanowski et al. "Enriching Word Vectors with Subword Information". In: *Transactions of the Association for Computational Linguistics* 5 (2017), pp. 135–146. DOI: 10.1162/tacl_a_00051. URL: https://www.aclweb.org/anthology/Q17-1010.

[17] François Bolley, José A Canizo, and José A Carrillo. "Stochastic mean-field limit: non-Lipschitz forces and swarming". In: *Mathematical Models and Methods in Applied Sciences* 21.11 (2011), pp. 2179–2210.

[18] Vincenzo Bonifaci, Kurt Mehlhorn, and Girish Varma. "Physarum can compute shortest paths". In: *Journal of Theoretical Biology* 309 (Sept. 2012), pp. 121–133. DOI: 10.1016/j.jtbi.2012.06.017.

[19] R. Borsche et al. "The scalar keller–segel model on networks". In: *Mathematical Models and Methods in Applied Sciences* 24.02 (Oct. 2013), pp. 221–247. DOI: 10.1142/S0218202513400071.

[20] Gabriella Bretti and Roberto Natalini. "Numerical approximation of nonhomogeneous boundary conditions on networks for a hyperbolic system of chemotaxis modeling the Physarum dynamics". en. In: *Journal of Computational Methods in Sciences and Engineering* 18.1 (Feb. 2018), pp. 85–115. DOI: 10.3233/JCM-170773.

[21] Joseph N. Burchett et al. "Revealing the Dark Threads of the Cosmic Web". en. In: *The Astrophysical Journal* 891.2 (Mar. 2020). DOI: 10.3847/2041-8213/ab700c.

[22] Scott Camazine et al. *Self-organization in biological systems*. Princeton University Press, 2003.

[23] Marco Caponigro et al. "Sparse stabilization and control of alignment models". In: *Mathematical Models and Methods in Applied Sciences* 25.03 (2015), pp. 521–564.

[24] Marco Caponigro et al. “Sparse stabilization and optimal control of the Cucker-Smale model”. In: *Mathematical Control and Related Fields* 3.4 (2013), pp. 447–466.

[25] E. Carlen, P. Degond, and B. Wennberg. “Kinetic limits for pair-interaction driven master equations and biological swarm models”. In: *Mathematical Models and Methods in Applied Sciences* 23.07 (2013), pp. 1339–1376.

[26] E. Carlen et al. “Kinetic hierarchy and propagation of chaos in biological swarm models”. In: *Physica D: Nonlinear Phenomena* 260 (2013), pp. 90–111.

[27] Eric Carlen et al. “Kinetic hierarchy and propagation of chaos in biological swarm models”. In: *Physica D: Nonlinear Phenomena* 260 (2013), pp. 90–111.

[28] José Antonio Carrillo, Katy Craig, and Francesco S. Patacchini. “A blob method for diffusion”. en. In: *Calculus of Variations and Partial Differential Equations* 58.2 (Feb. 2019), p. 53. DOI: 10.1007/s00526-019-1486-3.

[29] C. Castellano, S. Fortunato, and V. Loreto. “Statistical physics of social dynamics”. In: *Reviews of Modern Physics* 81.2 (May 2009), pp. 591–646. DOI: 10.1103/RevModPhys.81.591.

[30] Li Chen, Simone Göttlich, and Stephan Knapp. “Modeling of a diffusion with aggregation: rigorous derivation and numerical simulation”. en. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 52.2 (Mar. 2018). DOI: 10.1051/m2an/2018028.

[31] Li Chen, Simone Göttlich, and Qitao Yin. “Mean Field Limit and Propagation of Chaos for a Pedestrian Flow Model”. en. In: *Journal of Statistical Physics* 166.2 (Jan. 2017), pp. 211–229. DOI: 10.1007/s10955-016-1679-5.

[32] Euijin Choo, Ting Yu, and Min Chi. “Detecting Opinion Spammer Groups Through Community Discovery and Sentiment Analysis”. en. In: *Data and Applications Security and Privacy XXIX*. Ed. by Pierangela Samarati. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 170–187. DOI: 10.1007/978-3-319-20810-7_11.

[33] Fan R K Chung. “Lectures on Spectral Graph Theory”. en. In: (), p. 25.

[34] Iain D Couzin et al. “Effective leadership and decision-making in animal groups on the move”. In: *Nature* 433.7025 (2005), pp. 513–516.

[35] F. Cucker and S. Smale. “Emergent Behavior in Flocks”. In: *IEEE Transactions on Automatic Control* 52.5 (May 2007). DOI: 10.1109/TAC.2007.895842.

[36] G. Deffuant et al. “Mixing beliefs among interacting agents”. In: *Advances in Complex Systems* 03.01n04 (Jan. 2000), pp. 87–98. DOI: 10.1142/S0219525900000078.

[37] Pierre Degond and Sébastien Motsch. “Large Scale Dynamics of the Persistent Turning Walker Model of Fish Behavior”. en. In: *Journal of Statistical Physics* 131.6 (June 2008), pp. 989–1021. DOI: 10.1007/s10955-008-9529-8.

[38] Pierre Degond et al. “Hydrodynamic models of self-organized dynamics: Derivation and existence theory”. EN. In: *Methods and Applications of Analysis* 20.2 (June 2013). DOI: 10.4310/MAA.2013.v20.n2.a1.

[39] Morris H. Degroot. “Reaching a Consensus”. en. In: *Journal of the American Statistical Association* 69.345 (Mar. 1974), pp. 118–121. DOI: 10.1080/01621459.1974.10480137.

[40] William Deitrick and Wei Hu. “Mutually Enhancing Community Detection and Sentiment Analysis on Twitter Networks”. en. In: 2013 (Aug. 2013). DOI: 10.4236/jdaip.2013.13004.

[41] Charo I. Del Genio, Thilo Gross, and Kevin E. Bassler. “All Scale-Free Networks Are Sparse”. In: *Physical Review Letters* 107.17 (Oct. 2011). DOI: 10.1103/PhysRevLett.107.178701.

[42] Jacob Devlin et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. en. In: *arXiv:1810.04805 [cs]* (May 2019).

[43] Inderjit S. Dhillon, Yuqiang Guan, and Brian Kulis. “Kernel k-means: spectral clustering and normalized cuts”. en. In: *Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining - KDD ’04*. Seattle, WA, USA: ACM Press, 2004, p. 551. DOI: 10.1145/1014052.1014118.

[44] Lawrence C. Evans. *Partial Differential Equations*. en. American Mathematical Soc., 2010.

[45] M. Fiedler. “Algebraic connectivity of graphs”. In: *Czechoslovak mathematical journal* 23.2 (1973), pp. 298–305.

[46] Seth Flaxman, Sharad Goel, and Justin M. Rao. “Ideological Segregation and the Effects of Social Media on News Consumption”. en. In: *SSRN Electronic Journal* (2013). DOI: 10.2139/ssrn.2363701.

[47] Massimo Fornasier, Benedetto Piccoli, and Francesco Rossi. “Mean-field sparse optimal control”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 372.2028 (2014), p. 20130400.

[48] John R. P. French Jr. “A formal theory of social power”. In: *Psychological Review* 63.3 (1956). DOI: 10.1037/h0046123.

[49] Kiran Garimella et al. “Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship”. In: *Proceedings of the 2018 World Wide Web Conference*. WWW ’18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 913–922. DOI: 10.1145/3178876.3186139.

[50] S. Gokalp et al. “Partitioning and Scaling Signed Bipartite Graphs for Polarized Political Blogosphere”. In: *2013 International Conference on Social Computing*. Sept. 2013, pp. 168–173. DOI: 10.1109/SocialCom.2013.32.

[51] Pollyanna Gonçalves et al. “Comparing and combining sentiment analysis methods”. en. In: *Proceedings of the first ACM conference on Online social networks - COSN ’13*. Boston, Massachusetts, USA: ACM Press, 2013, pp. 27–38. DOI: 10.1145/2512938.2512951.

[52] Simone Göttlich, Stephan Knapp, and Dylan Weber. “The food seeking behavior of slime mold: a macroscopic approach [J]”. In: *Mathematical Biosciences and Engineering* 17.6 (2020), pp. 6631–6658.

[53] Alex Graves. “Generating Sequences With Recurrent Neural Networks”. In: *arXiv:1308.0850 [cs]* (June 2014).

[54] C.M. Grinstead and J.L. Snell. *Introduction to probability*. Chance Project, 2006.

[55] Yukio-Pegio Gunji et al. “Minimal model of a cell connecting amoebic motion and adaptive transport networks”. In: *Journal of Theoretical Biology* 253.4 (Aug. 2008), pp. 659–667. DOI: 10.1016/j.jtbi.2008.04.017.

[56] Frank Harary. “A criterion for unanimity in French’s theory of social power”. In: *Studies in social power*. Oxford, England: Univer. Michigan, 1959, pp. 168–182.

[57] R. Hegselmann and U. Krause. “Opinion dynamics and bounded confidence: models, analysis and simulation”. In: *Journal of Artificial Societies and Social Simulation* 5.3 (2002).

[58] Rainer Hegselmann, Ulrich Krause, et al. “Opinion dynamics and bounded confidence models, analysis, and simulation”. In: *Journal of artificial societies and social simulation* 5.3 (2002).

[59] Itai Himelboim et al. “Classifying Twitter Topic-Networks Using Social Network Analysis”. en. In: *Social Media + Society* 3.1 (Jan. 2017). DOI: 10.1177/2056305117691545.

[60] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Fundamentals of convex analysis*. Springer Science & Business Media, 2012.

[61] Franca Hoffmann et al. “Spectral Analysis Of Weighted Laplacians Arising In Data Clustering”. en. In: *arXiv:1909.06389 [math, stat]* (Sept. 2019).

[62] Clayton Hutto and Eric Gilbert. “Vader: A parsimonious rule-based model for sentiment analysis of social media text”. In: *Proceedings of the International AAAI Conference on Web and Social Media*. Vol. 8. 1. 2014.

[63] P-E. Jabin and S. Motsch. “Clustering and asymptotic behavior in opinion formation”. In: *Journal of Differential Equations* 257.11 (2014), pp. 4165–4187. DOI: 10.1016/j.jde.2014.08.005.

[64] Jeff Jones. “Approximating the Behaviours of Physarum polycephalum for the Construction and Minimisation of Synthetic Transport Networks”. en. In: *Unconventional Computation*. Ed. by Cristian S. Calude et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2009, pp. 191–208. DOI: 10.1007/978-3-642-03745-0_23.

[65] Jeff Jones. “Characteristics of Pattern Formation and Evolution in Approximations of Physarum Transport Networks”. In: *Artificial Life* 16.2 (Jan. 2010), pp. 127–153. DOI: 10.1162/artl.2010.16.2.16202.

[66] Jeff Jones. “Influences on the formation and evolution of Physarum polycephalum inspired emergent transport networks”. en. In: *Natural Computing* 10.4 (Dec. 2011), pp. 1345–1369. DOI: 10.1007/s11047-010-9223-z.

[67] Armand Joulin et al. “Bag of Tricks for Efficient Text Classification”. In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Valencia, Spain: Association for Computational Linguistics, Apr. 2017, pp. 427–431. URL: https://www.aclweb.org/anthology/E17-2068.

[68] Jon M Kleinberg. “Authoritative sources in a hyperlinked environment”. In: *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*. 1998, pp. 668–677.

[69] T. Kolokolnikov. “Maximizing algebraic connectivity for certain families of graphs”. In: *Linear Algebra and its Applications* 471 (2015), pp. 122–140.

[70] T. Kolokolnikov, B. Osting, and J. Von Brecht. “Algebraic connectivity of Erdös-Rényi graphs near the connectivity threshold”. In: *Manuscript in preparation* (2014).

[71] Ulrich Krause et al. “A discrete nonlinear and non-autonomous model of consensus formation”. In: *Communications in difference equations* 2000 (2000), pp. 227–236.

[72] Tanya Latty and Madeleine Beekman. “Speed–accuracy trade-offs during foraging decisions in the acellular slime mould Physarum polycephalum”. In: *Proceedings of the Royal Society B: Biological Sciences* 278.1705 (Feb. 2011). DOI: 10.1098/rspb.2010.1624.

[73] GuanLin Li, Sebastien Motsch, and Dylan Weber. “Bounded confidence dynamics and graph control: Enforcing consensus”. en. In: *Networks & Heterogeneous Media* 15.3 (2020). DOI: 10.3934/nhm.2020028.

[74] Frank Lin and William W. Cohen. “Power Iteration Clustering”. en. In: (June 2010). DOI: 10.1184/R1/6476252.v1.

[75] Bing Liu. “Sentiment Analysis and Opinion Mining”. In: *Synthesis Lectures on Human Language Technologies* 5.1 (May 2012). DOI: 10.2200/S00416ED1V01Y201204HLT016.

[76] Yinhan Liu et al. “RoBERTa: A Robustly Optimized BERT Pretraining Approach”. In: *arXiv:1907.11692 [cs]* (July 2019).

[77] Yuxin Liu et al. “A new multi-agent system to simulate the foraging behaviors of Physarum”. en. In: *Natural Computing* 16.1 (Mar. 2017), pp. 15–29. DOI: 10.1007/s11047-015-9530-5.

[78] Yuxin Liu et al. “A Physarum Network Evolution Model Based on IBTM”. en. In: *Advances in Swarm Intelligence*. Ed. by Ying Tan, Yuhui Shi, and Hongwei Mo. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2013, pp. 19–26.

[79] Ugo Lopez et al. “From behavioural analyses to models of collective motion in fish schools”. In: *Interface focus* 2.6 (2012), pp. 693–707.

[80] J. Lorenz. “Continuous opinion dynamics under bounded confidence: a survey”. In: *International Journal of Modern Physics C* 18.12 (Dec. 2007), pp. 1819–1838. DOI: 10.1142/S0129183107011789.

[81] Jan Lorenz. *Consensus Strikes Back in the Hegselmann-Krause Model of Continuous Opinion Dynamics Under Bounded Confidence*. en. Jan. 2006.

[82] Eugenio Martínez-Cámara et al. “Sentiment analysis in Twitter”. In: *Natural Language Engineering* 20 (Jan. 2014), pp. 1–28. DOI: 10.1017/S1351324912000332.

[83] Sylvie Méléard. “Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models”. In: *Probabilistic models for nonlinear partial differential equations*. Springer, 1996, pp. 42–95.

[84] Amy Mitchell and Tom Rosenstiel. *Navigating News Online*. en-US. May 2011.

[85] S. Motsch and E. Tadmor. “Heterophilious Dynamics Enhances Consensus”. In: *SIAM Review* 56.4 (Jan. 2014), pp. 577–621. DOI: 10.1137/120901866.

[86] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. “How simple rules determine pedestrian behavior and crowd disasters”. In: *Proceedings of the National Academy of Sciences* 108.17 (2011), pp. 6884–6888.

[87] Mika V. Mäntylä, Daniel Graziotin, and Miikka Kuutila. “The evolution of sentiment analysis—A review of research topics, venues, and top cited papers”. en. In: *Computer Science Review* 27 (Feb. 2018), pp. 16–32. DOI: 10.1016/j.cosrev.2017.10.002.

[88] T. Nakagaki et al. “Obtaining multiple separate food sources: behavioural intelligence in the Physarum plasmodium”. In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271.1554 (Nov. 2004). DOI: 10.1098/rspb.2004.2856.

[89] Toshiyuki Nakagaki. “Smart behavior of true slime mold in a labyrinth”. en. In: *Research in Microbiology* 152.9 (Nov. 2001), pp. 767–770. DOI: 10.1016/S0923-2508(01)01259-1.

[90] Toshiyuki Nakagaki, Hiroyasu Yamada, and Ágota Tóth. “Maze-solving by an amoeboid organism”. en. In: *Nature* 407.6803 (Sept. 2000). DOI: 10.1038/35035159.

[91] Toshiyuki Nakagaki, Hiroyasu Yamada, and Ágota Tóth. “Path finding by tube morphogenesis in an amoeboid organism”. en. In: *Biophysical Chemistry* 92.1-2 (Sept. 2001), pp. 47–52. DOI: 10.1016/S0301-4622(01)00179-X.

[92] Nakagaki T. et al. “Obtaining multiple separate food sources: behavioural intelligence in the Physarum plasmodium”. In: *Proceedings of the Royal Society of*

London. Series B: Biological Sciences 271.1554 (Nov. 2004), pp. 2305–2310. DOI: 10.1098/rspb.2004.2856.

[93] Andrew Y. Ng, Michael I. Jordan, and Yair Weiss. "On Spectral Clustering: Analysis and an algorithm". In: *Advances in Neural Information Processing Systems 14*. Ed. by T. G. Dietterich, S. Becker, and Z. Ghahramani. MIT Press, 2002, pp. 849–856.

[94] Christina Oettmeier, Klaudia Brix, and Hans-Günther Döbereiner. "Physarum polycephalum—a new take on a classic model system". en. In: *Journal of Physics D: Applied Physics* 50.41 (Sept. 2017). DOI: 10.1088/1361-6463/aa8699.

[95] R. Olfati-Saber, J. A. Fax, and R. M. Murray. "Consensus and Cooperation in Networked Multi-Agent Systems". In: *Proceedings of the IEEE* 95.1 (Jan. 2007), pp. 215–233. DOI: 10.1109/JPROC.2006.887293.

[96] Lawrence Page et al. *The PageRank Citation Ranking: Bringing Order to the Web.* Nov. 1999.

[97] Benoit Perthame et al. "Mathematical tools for kinetic equations". In: *Bulletin of the American Mathematical Society* 41.2 (2004), pp. 205–244.

[98] Matthew E. Peters et al. "Deep contextualized word representations". In: *arXiv:1802.05365 [cs]* (Mar. 2018).

[99] Benedetto Piccoli, Nastassia Pouradier Duteil, and Emmanuel Trélat. "Sparse Control of Hegselmann–Krause Models: Black Hole and Declustering". In: *SIAM Journal on Control and Optimization* 57.4 (2019), pp. 2628–2659.

[100] Benedetto Piccoli and Andrea Tosin. "Pedestrian flows in bounded domains with obstacles". In: *Continuum Mechanics and Thermodynamics* 21.2 (2009), pp. 85–107.

[101] N. Privault. *Stochastic Finance: An Introduction with Market Examples*. CRC Press, 2013.

[102] Walter Quattrociocchi, Antonio Scala, and Cass R. Sunstein. *Echo Chambers on Facebook*. en. SSRN Scholarly Paper ID 2795110. Rochester, NY: Social Science Research Network, June 2016.

[103] Alec Radford et al. "Improving Language Understanding by Generative Pre-Training". en. In: (), p. 12.

[104] Alec Radford et al. "Language models are unsupervised multitask learners". In: *OpenAI blog* 1.8 (2019), p. 9.

[105] Sascha Rothe, Shashi Narayan, and Aliaksei Severyn. “Leveraging Pre-trained Checkpoints for Sequence Generation Tasks”. In: *arXiv:1907.12461 [cs]* (Apr. 2020).

[106] R. Saber and R. Murray. “Consensus protocols for networks of dynamic agents”. In: *Proceedings of the 2003 American Control Conference, 2003*. Vol. 2. Piscatway, NJ: IEEE, June 2003, pp. 951–956.

[107] Satu Elisa Schaeffer. “Graph clustering”. en. In: *Computer Science Review* 1.1 (Aug. 2007), pp. 27–64. DOI: 10.1016/j.cosrev.2007.05.001.

[108] Jianbo Shi and Jitendra Malik. “Normalized Cuts and Image Segmentation”. en. In: *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE Intelligence* 22.8 (2000), p. 18.

[109] D. Spanos, R. Olfati-Saber, and R. Murray. “Dynamic consensus on mobile networks”. In: *IFAC world congress*. Citeseer, 2005, pp. 1–6.

[110] Natalie Jomini Stroud. “Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure”. en. In: *Political Behavior* 30.3 (Sept. 2008), pp. 341–366. DOI: 10.1007/s11109-007-9050-9.

[111] Alain-Sol Sznitman. “Topics in propagation of chaos”. In: *Ecole d’été de probabilités de Saint-Flour XIX—1989*. Springer, 1991, pp. 165–251.

[112] Atsushi Tero, Ryo Kobayashi, and Toshiyuki Nakagaki. “A mathematical model for adaptive transport network in path finding by true slime mold”. In: *Journal of Theoretical Biology* 244.4 (Feb. 2007), pp. 553–564. DOI: 10.1016/j.jtbi.2006.07.015.

[113] Atsushi Tero, Ryo Kobayashi, and Toshiyuki Nakagaki. “Physarum solver: A biologically inspired method of road-network navigation”. en. In: *Physica A: Statistical Mechanics and its Applications*. Information and Material Flows in Complex Networks 363.1 (Apr. 2006), pp. 115–119. DOI: 10.1016/j.physa.2006.01.053.

[114] Atsushi Tero et al. “Flow-network adaptation in Physarum amoebae”. en. In: *Theory in Biosciences* 127.2 (May 2008), pp. 89–94. DOI: 10.1007/s12064-008-0037-9.

[115] Nicolas Garcia Trillos, Franca Hoffmann, and Bamdad Hosseini. “Geometric structure of graph Laplacian embeddings”. en. In: *arXiv:1901.10651 [math, stat]* (Jan. 2019).

[116] Michail-Antisthenis I Tsompanas and Georgios Ch Sirakoulis. “Modeling and hardware implementation of an amoeba-like cellular automaton”. en. In: *Bioinspiration & Biomimetics* 7.3 (Sept. 2012), p. 036013. DOI: 10.1088/1748-3182/7/3/036013.

[117] Ashish Vaswani et al. “Attention Is All You Need”. In: *arXiv:1706.03762 [cs]* (June 2017).

[118] Tamás Vicsek and Anna Zafeiris. “Collective motion”. In: *Physics reports* 517.3-4 (2012), pp. 71–140.

[119] Dylan Weber, Ryan Theisen, and Sebastien Motsch. “Deterministic Versus Stochastic Consensus Dynamics on Graphs”. en. In: *Journal of Statistical Physics* 176.1 (July 2019), pp. 40–68. DOI: 10.1007/s10955-019-02293-5.

[120] Yuheng Wu et al. “A new model to imitate the foraging behavior of Physarum polycephalum on a nutrient-poor substrate”. In: *Neurocomputing* 148 (Jan. 2015), pp. 63–69. DOI: 10.1016/j.neucom.2012.10.044.

[121] H. Xia, H. Wang, and Z. Xuan. “Opinion Dynamics: A Multidisciplinary Review and Perspective on Future Research”. en. In: *International Journal of Knowledge and Systems Science (IJKSS)* 2.4 (Oct. 2011), pp. 72–91. DOI: 10.4018/jkss.2011100106.

[122] Zhilin Yang et al. “XLNet: Generalized Autoregressive Pretraining for Language Understanding”. In: *arXiv:1906.08237 [cs]* (Jan. 2020).

[123] W. Yu et al. “Second-Order Consensus for Multiagent Systems With Directed Topologies and Nonlinear Dynamics”. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 40.3 (June 2010), pp. 881–891. DOI: 10.1109/TSMCB.2009.2031624.

[124] Bernt Øksendal. “Stochastic Differential Equations”. en. In: *Stochastic Differential Equations: An Introduction with Applications*. Ed. by Bernt Øksendal. Universitext. Berlin, Heidelberg: Springer, 2003, pp. 65–84. DOI: 10.1007/978-3-642-14394-6_5.

APPENDIX A

APPENDIX

Proposition 10 *The matrix L is a diagonal dominant matrix, its eigenvalues satisfy* $Re(\lambda_i) \geq 0$ *and no eigenvalue* λ_i *is purely imaginary except zero.*

Proof
The diagonal entries of L satisfy $\sigma_i = \sum_{j,j\neq i} a_{ij}$ where $a_{ij} > 0$, thus summing over each row will give zero and we deduce that L is diagonal dominant. By the Gershgorin disc theorem the eigenvalues λ_i are contained in the closed discs $B(\sigma_i, \sigma_i)$ (see Fig. 3). Thus, the eigenvalues λ_i are either 0 or have a real part strictly positive. Moreover, 0 is always an eigenvalue of L as the constant vector $\mathbf{1} = (1, \ldots, 1)^T$ is always an eigenvector of L associated with the eigenvalue $\lambda = 0$. ■

Lemma 7 *If L is irreducible then the eigenvalue* 0 *is simple.*

Proof
By Proposition 1 we have that the real parts of the eigenvalues of L are all nonnegative and there are no purely imaginary eigenvalues. We also note that 0 is an eigenvalue of L of at least multiplicity one because it is associated to $\mathbf{1}$. Consider the matrix $D = L - cId$ where $c \in \mathbb{R}$ and notice that $\lambda = a + ib$ is an eigenvalue of L if and only if $\widehat{\lambda} = a - c + ib$ is an eigenvalue of D. Therefore, since 0 is an eigenvalue of L we must have that $-c$ is an eigenvalue of D. If we choose c such that:

$$c > \max_i \frac{b_i^2 + a_i^2}{2a_i},$$

we have that:

$$\sqrt{(a_i - c)^2 + b_i^2} < c \quad \text{for all } i.$$

Therefore $-c$ is the eigenvalue of largest modulus of D and since D is irreducible we have by the Perron-Frobenius theorem that $-c$ must be simple and since the eigenvalues of D are in a one to one correspondence with the eigenvalues of L we must have that 0 is a simple eigenvalue of L as desired. ■

Corollary 8 *Suppose L is irreducible and symmetric. Then the solution of the consensus model,* $\mathbf{s}(t)$*, satisfies:*

$$\mathbf{s}(t) \to \bar{s}(0)\,\mathbf{1} \quad as \quad t \to +\infty.$$

with $\bar{s}$ *the average opinion* (2.9)*. Moreover,* $|s_i(t) - \bar{s}(0)| \leq Ce^{-\lambda_2 t}$ *where C depends only on the initial condition and* λ_2 *is the second largest eigenvalue of L.*

Proof
The consensus model is a linear system, therefore its solution is given by:

$$\mathbf{s}(t) = e^{-tL}\mathbf{s_0}$$

where $\mathbf{s_0} = (s_1(0), ..., s_N(0))^T$. Note that since $\mathbf{1}$ is an eigenvector of L corresponding to the eigenvalue 0, it is also an eigenvector of e^{-tL} corresponding to the eigenvalue of 1. Therefore we may write:

$$\mathbf{s}(t) - \bar{s}(0)\mathbf{1} = e^{-tL}(\mathbf{s_0} - \bar{s}(0)\,\mathbf{1}). \tag{A.1}$$

Also note that since L is diagonal dominant and symmetric that there exists P such that

$$L = PDP^{-1} \tag{A.2}$$

where $D = \text{diag}(0, \lambda_2, \lambda_3, ...)$ and P is the matrix composed of the eigenvectors of L. Since L is symmetric these eigenvalues are real. By Proposition 1 and Lemma 1 there is exactly one zero eigenvalue and λ_i is strictly positive for $2 \leq i \leq n$. We now define:

$$\mathbf{y}(t) := P^{-1}(\mathbf{s}(t) - \bar{s}(0)\,\mathbf{1}),$$

this is the difference between the opinion at time t and the average opinion represented in the diagonal coordinate system. Notice that:

$$\mathbf{y}'(t) = \frac{d}{dt}[P^{-1}(\mathbf{s}(t) - \bar{s}(0)\,\mathbf{1})] = D\mathbf{y}(t).$$

Therefore, since this is an uncoupled system of linear equations we must have that:

$$y_i(t) = y_i(0)e^{-\lambda_i t}.$$

So for $i \geq 2$ we must have that $y_i(t) \to 0$ as $t \to +\infty$ exponentially with rate at least λ_2. To conclude, it remains to show that $y_1(t) = 0$. Using that the eigenvectors of L form an orthogonal basis, the entry $y_1(t)$ is given by:

$$y_1(t) = \langle \mathbf{s}(t) - \bar{s}(0)\,\mathbf{1}\,,\ \frac{\mathbf{1}}{\|\mathbf{1}\|} \rangle = \bar{s}(t) - \bar{s}(0) = 0$$

since the mean value $\bar{s}(t)$ is preserved over time. ■

A.2 Convergence of linear systems

Lemma 8 *Given a linear system defined by:*

$$\mathbf{x}' = A\mathbf{x}\,, \quad \mathbf{x}(0) = \mathbf{x}_0.$$

Assume A has d distinct eigenvalues and a zero eigenvalue of multiplicity m with m linearly independent associated eigenvectors. If for all $\lambda_i \in Sp(A)$ with $i > m$ we have that $Re(\lambda_i) < 0$ then

$$\lim_{t \to +\infty} \mathbf{x}(t) = \mathbf{u}$$

where $\mathbf{u}$ *is in the center subspace of A,* E^c.

Proof
For ease of notation we will write

$$\text{Spec}(A) = \{ \lambda_1, .., \lambda_d \}$$

where m_i is the algebraic multiplicity of λ_i. We will also write that $\lambda_1 = 0$. We know that given λ_i that we may find m_i linearly independent generalized eigenvectors of A, $\{ \mathbf{v}^1_{\lambda_i}, ..., \mathbf{v}^{m_i}_{\lambda_i} \}$. The generalized eigenspace corresponding to λ_i is given by $E_{\lambda_i} = \text{span}\{ \mathbf{v}^1_{\lambda_i}, ..., \mathbf{v}^{m_i}_{\lambda_i} \}$ and by the generalized eigenspace decomposition theorem we can find a basis of $\mathbb{R}^d$ consisting of generalized eigenvectors of A. Therefore, we may write:

$$\mathbb{R}^d = \bigoplus_{i=1}^{d} E_{\lambda_i}. \tag{A.3}$$

We know by the fundamental theorem of linear systems that the solution to the system is given by:

$$\mathbf{x}(t) = e^{At}\mathbf{x}_0.$$

By (A.3) we may write:

$$\mathbf{x}_0 = \mathbf{e}_1 + ... + \mathbf{e}_d$$

where for each $1 \leq i \leq d$ we have $\mathbf{e}_i \in E_{\lambda_i}$. Notice that since we are given that there are m_1 distinct eigenvectors associated with $\lambda_1 = 0$ that for each $w \in E_{\lambda_1}$ we have that $A\mathbf{w} = 0$. Consequently we have that $A_{|E_{\lambda_1}} = 0$ which implies that $A_{|E_{\lambda_1}} = 0$ for each $t \in \mathbb{R}$. Taking the matrix exponential of both sides yields

$$e^{tA_{|E_{\lambda_1}}} = \text{Id} \quad \text{for each } t \in \mathbb{R}.$$

Therefore we must have that

$$\begin{aligned} e^{At}\mathbf{x}_0 &= e^{At}\mathbf{e}_1 + ... + e^{At}\mathbf{e}_d \\ &= e^{tA_{|E_{\lambda_1}}}\mathbf{e}_1 + ... + e^{tA_{|E_{\lambda_d}}}\mathbf{e}_d \\ &= \mathbf{e}_1 + e^{tA_{|E_{\lambda_2}}}\mathbf{e}_2 + ... + e^{tA_{|E_{\lambda_d}}}\mathbf{e}_d. \end{aligned}$$

We now claim that for any $i \geq 2$:

$$\lim_{t \to +\infty} e^{tA_{|E_{\lambda_i}}} \mathbf{e}_i = 0.$$

Since we are writing $\mathbf{x_0}$ with respect to the basis of generalized eigenvectors of A we have that

$$A_{|E_{\lambda_i}} = \lambda_i Id + N_i$$

where N_I is nilpotent. Consequently

$$\begin{aligned} e^{tA_{|E_{\lambda_i}}} e_i &= e^{t(\lambda_i \mathrm{Id} + N_i)} \mathbf{e}_i \\ &= e^{t\mathrm{Re}(\lambda_i)} e^{t\mathrm{Im}(\lambda_i)i} (Id + tN_i + ... + \frac{1}{k!} t^k N_i^k) \mathbf{e}_i \\ &\leq C_1 e^{t\mathrm{Re}(\lambda_i)} (Id + tN_i + ... + \frac{1}{k!} t^k N_i^k) \mathbf{e}_i \end{aligned}$$

with $C_1 > 0$. Therefore, since $\mathrm{Re}(\lambda_i) < 0$ and every coordinate of $(Id + tN_i + ... + \frac{1}{k!} t^k N_i^k)\mathbf{e}_i$ is polynomial in t we must have that there exist $C > 0$ and $0 < \epsilon < -\mathrm{Re}(\lambda_i)$ such that

$$\left\| e^{tA_{|E_{\lambda_i}}} \mathbf{e}_i \right\| \leq C e^{(\mathrm{Re}(\lambda_i) + \epsilon)t} \to 0 \quad \text{as} \quad t \to +\infty. \tag{A.4}$$

We deduce that for $i \geq 2$:

$$\lim_{t \to +\infty} e^{tA_{|E_{\lambda_i}}} \mathbf{e}_i = 0$$

which implies

$$\lim_{t \to +\infty} e^{At} \mathbf{x_0} \;=\; \lim_{t \to +\infty} \mathbf{e}_1 + e^{tA_{|E_{\lambda_2}}} \mathbf{e}_2 + ... + e^{tA_{|E_{\lambda_d}}} \mathbf{e}_d \;=\; \mathbf{e}_1.$$

Since $\mathbf{e}_1 \in E^c$ we have that $\lim_{t \to +\infty} \mathbf{x}(t) \in E^c$ as desired. ■

A.3 Decomposition into strongly connected components

Lemma 9 *If L is the Laplacian of a directed graph $G = (V, E)$ then by relabeling vertices L can be represented:*

$$L = \begin{bmatrix} B_1 & & & \\ & B_2 & & 0 \\ & & B_3 & \\ & * & & \ddots \\ & & & & B_k \end{bmatrix}$$

Proof
We may partition the vertex set of G, V, into its strongly connected components $\{ U_1, \ldots, U_k \}$, so that:

$$V = \uplus_i^k U_i.$$

If we consider the set $\{ U_1, \ldots, U_k \}$ as the vertex set of a new graph G^* with edge set E^* given by:

$$(U_m, U_n) \in E^* \text{ if there exists } u \in U_m \text{ and } v \in U_n \quad \text{with} \quad (u, v) \in E. \tag{A.5}$$

Then the graph G^* is a directed acyclic graph as G has been partitioned into strongly connected components; if a cycle existed all vertices included in the cycle would represent the same strongly connected component of G by (A.5) contradicting the partition of V into strongly connected components. Therefore there exists a topological ordering $\leq^*$ on the vertex set of G^*. This ordering is given by:

$$U_m \leq^* U_n \quad \text{if} \quad (U_m, U_n) \in E^*.$$

We only need to label the vertices of V in such a way that they respect the topological ordering. Then, this labeling of V produces L in the desired form. ■

A.4 Projection lemma

Lemma 10 *Let* $\{ u_i \}_i \subseteq \mathbb{R}^d$ *and consider* $C = \{ v \mid \langle v, u_i \rangle \geq 0 \text{ for all } i \}$. *If* u *is in* C *then* $\langle P_C(x), u \rangle \geq \langle x, u \rangle$ *for all* x *in* $\mathbb{R}^d$.

Proof
$P_C(x)$ is by definition the solution to the minimization problem:

$$\begin{aligned} &\text{minimize } f(y) = \frac{1}{2}\|y - x\|^2 \\ &\text{subject to } g_i(y) = \langle y, u_i \rangle \geq 0 \end{aligned}$$

Since C is closed and convex there exists a unique minimizer y^* and since the constraints are linear the Karush-Kuhn-Tucker conditions imply that:

$$\nabla f(y^*) = \sum_i \lambda_i \nabla g_i(y^*)$$

where $\lambda_i \geq 0$. Therefore:

$$y^* = x + \sum_i \lambda_i u_i$$

which implies by definition of C that:

$$\langle P_C(x), u \rangle = \langle x + \sum_i \lambda_i u_i, u \rangle = \langle x, u \rangle + \sum_i \lambda_i \langle u_i, u \rangle \geq \langle x, u \rangle$$

as desired. ■

A.5 1-D decay inequality

Lemma 11 *For any* $0 < r_* < 1$, $\eta > 0$ *and* $N > 0$, *there exist* $T > 0$ *and* $\delta > 0$ *such that:*

$$\begin{aligned} \delta + 2T &\leq \min(r_*, 1 - r_*) && \text{(A.6)} \\ \frac{T^2}{2N^2}\eta\Big(\eta(1 - \delta - 2T) - 2N(\delta + 2T)\Big) &\geq \delta. && \text{(A.7)} \end{aligned}$$

Proof
We let $\delta = T^3$ and show that for $T > 0$ sufficiently small both equations are satisfied. Indeed, the substitution leads to:

$$\begin{aligned} T^3 + 2T &\leq \min(r_*, 1 - r_*) && \text{(A.8)} \\ \frac{1}{2N^2}\eta\Big(\eta(1 - T^3 - 2T) - 2N(T^3 + 2T)\Big) &\geq T && \text{(A.9)} \end{aligned}$$

Since $\min(r_*, 1 - r_*) > 0$ and $T^3 + 2T \overset{T\to 0}{\longrightarrow} 0$, there exists $T_1 > 0$ such that (A.8) is satisfied for $0 < T \leq T_1$. Similarly, for the equation (A.9), we notice that

$$\frac{1}{2N^2}\eta\Big(\eta(1 - T^3 - 2T) - 2N(T^3 + 2T)\Big) \overset{T\to 0}{\longrightarrow} \frac{\eta^2}{2N^2} > 0.$$

Therefore, there exists $T_2 > 0$ such that (A.9) is satisfied for $0 < T \leq T_2$. Taking $T = \min(T_1, T_2)$ and $\delta = T^3$, we deduce a solution to (3.37)-(3.38). ■