BOUNDED CONFIDENCE DYNAMICS AND GRAPH CONTROL: ENFORCING CONSENSUS

GuanLin Li

Georgia Institute of Technology
Program in Quantitative Biosciences, Georgia Institute of Technology School of Physics
Atlanta, GA 30332, USA

Sebastien Motsch and Dylan Weber

Arizona State University
School of Mathematical and Statistical Sciences
Tempe, AZ 85257-1804, USA

Abstract. A generic feature of bounded confidence type models is the formation of clusters of agents. 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 converge to a consensus as the control still maintains connectivity of the interaction network.

1. **Introduction.** Mathematical models of opinion formation have long been objects of theoretical interest. Models in this context are often posed in an agent based framework where the potential for agents to interact is encoded in a network [6,12,17,20,26,28]. The rise of social networks as some of the main forums for the exchange of ideas clearly motivates the need to continue the study of these models. Analysis of the metadata associated with social networks shows that an emergent feature is the formation of polarized communities or "echo chambers" [9–11,23]. In this paper we study a class of models that exhibit this phenomenon. We are especially interested in the emergence of a *consensus* - when the opinions of all agents agree.

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 [19, 24, 27]. 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 [19, 27]. 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 [20, 24, 25, 29]. 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 [19].

We study a class of models inspired by the Hegselmann-Krause bounded confidence model [12] 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. In fact, it is relatively easy to show that a consensus can only occur if the initial opinions of all the agents are within the interaction range of each other. For this reason much of the study of this class of models has focused on characterizing the clustering behavior [2, 3, 7, 14, 16, 18]. We take a different viewpoint in this manuscript and instead investigate controls on collections of agents [5, 21] 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 whose opinions 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 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.

2. **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:

$$\dot{\mathbf{x}}_i = \sum_{j=1}^{N} a_{ij}(\mathbf{x}_j - \mathbf{x}_i), \quad a_{ij} = \frac{\Phi_{ij}}{\sum_{k=1}^{N} \Phi_{ik}} \quad \text{with} \ \Phi_{ij} = \Phi(|\mathbf{x}_j - \mathbf{x}_i|). \qquad (2.1)$$

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:

$$m = \min_{r \in [0,1]} \Phi(r) \quad , \quad M = \max_{r \in [0,1]} \Phi(r). \qquad (2.2)$$

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 [12]. 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 1. Notice that since the interaction coefficients a_{ij} satisfy $\sum_{j=1}^{N} a_{ij} = 1$ we can rewrite the dynamics (2.1) as:

$$\dot{\mathbf{x}}_i = \overline{\mathbf{x}}_i - \mathbf{x}_i, \quad \overline{\mathbf{x}}_i = \sum_{j=1}^{N} a_{ij}\mathbf{x}_j, \qquad (2.3)$$

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 1).
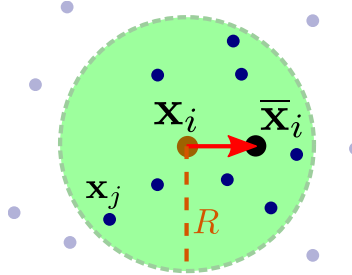

Figure 1. The movement of an agent according to the bounded confidence dynamics (2.3).

In this manuscript we will be concerned with conditions that cause the opinions of all agents to converge to a consensus.

Definition 1. We say that the dynamics (2.1) converge to a **consensus** if there exists $\mathbf{x}_*$ such that:

$$\lim_{t \to +\infty} \mathbf{x}_i(t) = \mathbf{x}_* \quad \text{for any } i. \tag{2.4}$$

We will see that due to the local consensus assumption, the notion of *connectivity* is crucial to the formation of a consensus.

Definition 2. 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:

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

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 3-left); the dynamics (2.1) must be modified in some manner for connectedness of the initial condition to be sufficient for consensus.

3. **No one left behind - enforcing consensus.**

3.1. **Critical region.** We now modify the dynamics (2.1) 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 2).

Definition 3. 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:

$$\mathcal{B}_i = \{\mathbf{x} \in \mathbb{R}^d \mid 1 - r_* \leq |\mathbf{x} - \mathbf{x}_i| \leq 1 \text{ and } \langle \overline{\mathbf{x}}_i - \mathbf{x}_i, \mathbf{x} - \mathbf{x}_i \rangle \leq 0\}. \tag{3.1}$$

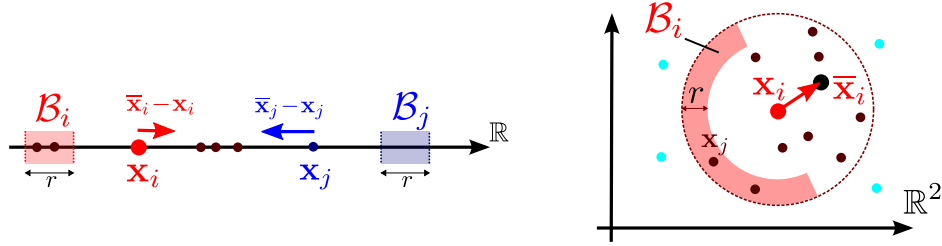where $\overline{\mathbf{x}}_i$ is given by (2.3).

Figure 2. Illustration of the critical regions (3.1) 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 $\overline{\mathbf{x}}_i$ and 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. **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, (2.3); 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 (2.1). 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:

$$\dot{x}_i = \mu_i(\overline{x}_i - x_i), \quad \mu_i = \begin{cases} 0 & \text{if there exists } x_j \in \mathcal{B}_i \\ 1 & \text{otherwise} \end{cases} \tag{3.2}$$

where $\overline{x}_i$ is the local average defined in (2.3).

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 (2.1) 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 2. 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:

$$(i, j) \in E^{\mathcal{B}} \text{ if } j \in \mathcal{B}_i. \tag{3.3}$$

In this notation we could write (3.2) as:

$$\dot{x}_i = \mu_i(\overline{x}_i - x_i), \quad \mu_i = \begin{cases} 0 & \text{if there exists } j \text{ such that } (i, j) \in E_{\mathcal{B}} \\ 1 & \text{otherwise} \end{cases}$$
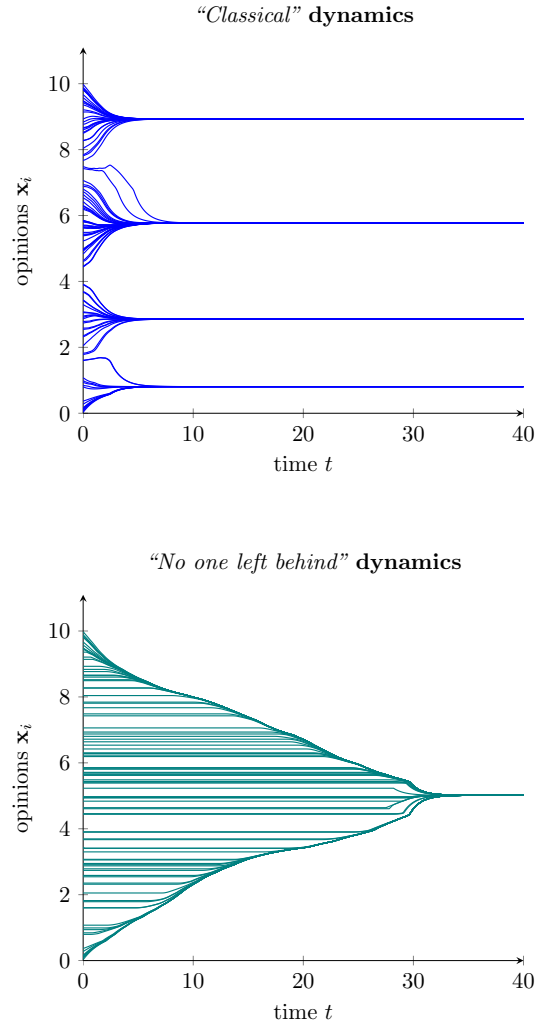
"Classical" **dynamics**

"No one left behind" **dynamics**

Figure 3. Simulation of the opinion dynamics without and with control (resp. left and right figure), e.g. solving resp. (2.3) and Model 1 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 4.

We examine the effect that the augmented dynamics have on the long time behavior of the opinions in Figure 3 and note that for the same initial condition that normal bounded confidence dynamics do not result in a consensus but four "clusters"
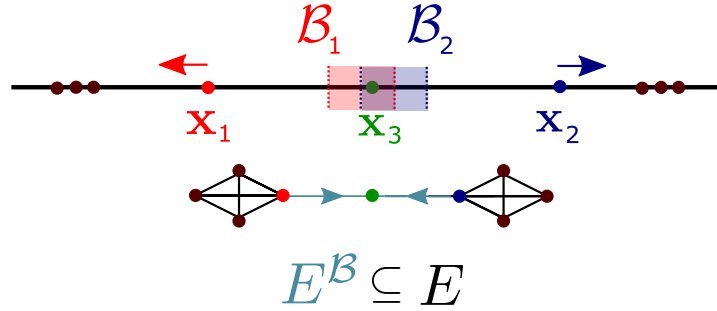
Figure 4. A configuration of agents (top) and the resulting inter-action graph (edge set E, black) and behind graph (edge set $E^{\mathcal{B}}$), light blue). Note that the behind graph is a directed subgraph of the interaction graph.

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.

3.3. **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 4 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. 5):

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

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 (3.2), 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.
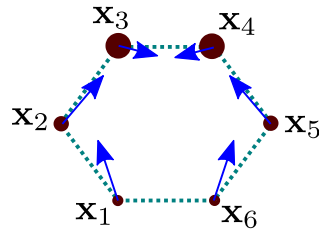

Figure 5. Counter-example in multi-dimension. Blue arrow is the velocity of each cluster. In this setting, every agent has someone in its critical region $\mathcal{B}_i$. Thus, the *naive* control in Model 1 would prevent anyone from moving.

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 4. 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:

$$\mathcal{C}_i = \{\mathbf{v} \in \mathbb{R}^d \,|\, \langle \mathbf{v}, \mathbf{x}_j - \mathbf{x}_i \rangle \geq 0 \quad \text{for all } \mathbf{x}_j \in \mathcal{B}_i\}. \tag{3.4}$$

where $\mathcal{B}_i$ is the critical region (3.1) associated to agent i. If the critical region $\mathcal{B}_i$ is empty, then $\mathcal{C}_i = \mathbb{R}^d$.

Remark 3. We note that we can define the cone of admissible velocities in terms of the behind graph introduced in Remark 2. We just have to replace $\mathbf{x}_j \in \mathcal{B}_i$ by $(i, j) \in E^{\mathcal{B}}$ in the definition of $\mathcal{C}_i$.
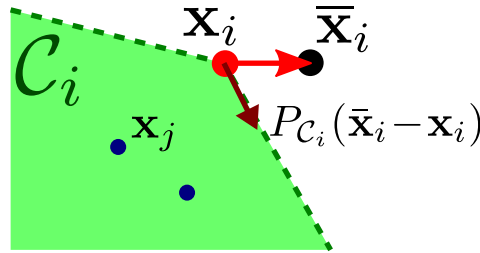

Figure 6. The velocity of agent i is the projection of the *desired* velocity $\overline{\mathbf{x}}_i - \mathbf{x}_i$ onto the cone of admissible velocity $\mathcal{C}_i$.

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 [13]. 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:

$$\mathbf{x}_i' = P_{\mathcal{C}_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i) \tag{3.5}$$

where $\overline{\mathbf{x}}_i$ is the average velocity defined in (2.3) and $P_{\mathcal{C}_i} : \mathbb{R}^d \to \mathcal{C}_i$ is the projection operator associated to the cone of admissible velocities $\mathcal{C}_i$ (3.4).

Remark 4. We note that the 1-D NOLB dynamics introduced in Model 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

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

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

$$\mathbf{x}_i' = P_{\mathcal{C}_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i) = \overline{\mathbf{x}}_i - \mathbf{x}_i.$$

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:

$$\mathcal{C}_i = \{\, \mathbf{v} \in \mathbb{R} \mid \mathbf{v} \leq 0 \,\}.$$

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

$$\mathbf{x}'_i = P_{\mathcal{C}_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i) = 0.$$

We illustrate the dynamics (3.5) and its long term behavior in Figure 7. As in the 1D case, a consensus is reached after some time whereas the classical dynamics generate multiple clusters.
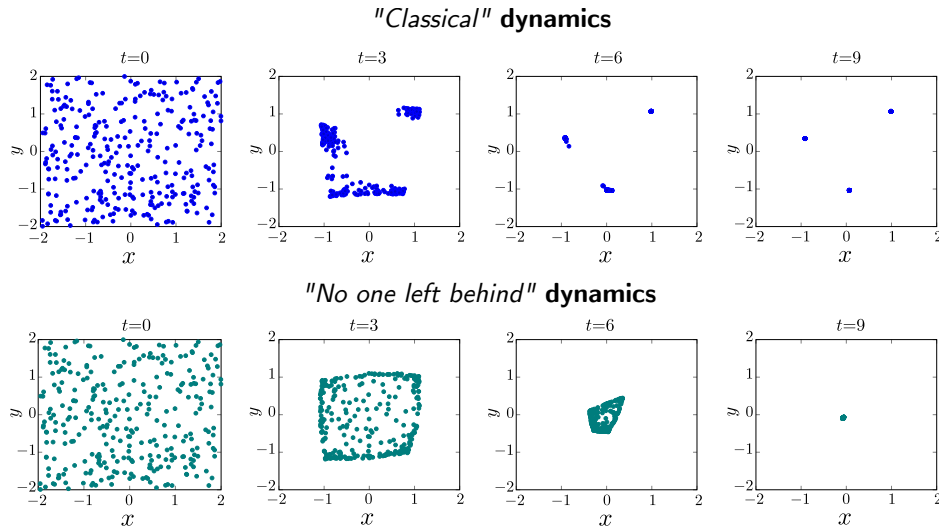

Figure 7. 2D simulation of opinion dynamics without and with control (resp, top and bottom figure), e.g. solving resp. (1) and (3.5) with $r_* = \frac{1}{2}$. With the control (bottom), the dynamics converge to a consensus.

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.

4. Convergence to a consensus.

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

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

The agents contract in the following sense (see [14] for a proof).

Proposition 1. *If $\{\, x_i \,\}_i$ evolves according to the bounded confidence dynamics or the NOLB dynamics then the convex hull Ω satisfies:*

$$\Omega(t_1) \subset \Omega(t_0) \quad \text{for any} \quad t_1 \geq t_0 \tag{4.1}$$

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 1. *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 2. *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:*

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

So in both the bounded confidence model and the NOLB model, agents remain close to each other in the sense of Proposition 1. 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 (3.5) 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 3 and 7. The mechanism through which the local control accomplishes this can be seen in the following result.

Proposition 2. *Fix $0 \leq r_* \leq 1$. Suppose that at time t the opinions of agents i and j evolve according to (3.5) 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:*

$$\frac{d}{dt}|\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 0. \tag{4.2}$$

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:

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

Denote $\mathbf{v}_i = P_{\mathcal{C}_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i)$. 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 $\langle P_{\mathcal{C}_i}(\overline{\mathbf{x}}_i - \mathbf{x}_i), \mathbf{x}_j - \mathbf{x}_i\rangle \geq 0$. We can prove similarly that $\langle P_{\mathcal{C}_j}(\overline{\mathbf{x}}_j - \mathbf{x}_j), \mathbf{x}_i - \mathbf{x}_j\rangle \geq 0$ and therefore,

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

$\square$

The critical region acts as a "trap". If the distance between two agents 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 3. *Suppose $|\mathbf{x}_i(t_0) - \mathbf{x}_j(t_0)| \leq 1$. Then $|\mathbf{x}_i(t) - \mathbf{x}_j(t)| \leq 1$ for all $t \geq t_0$.*

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 solution to the dynamics is 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 2 $|\mathbf{x}_i(t) - \mathbf{x}_j(t)|$ is decaying on this time interval, a contradiction as $|\mathbf{x}_i(t_1) - \mathbf{x}_j(t_1)| = 1$. $\square$

So, adding the control indeed prevents situations like the ones presented in Figures 3 and 7 and preserves connectivity.

Corollary 4. *The dynamics* (3.5) *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 (3.5) leads to a consensus.

4.2. **Emergence of a consensus.** We will now see that under the NOLB dynamics defined in (3.5), 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 vector field 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.

4.2.1. *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 8. Additionally we note that this example provides an illustration that the vector field associated to the NOLB dynamics is *not* continuous as the dynamics converge to a state that is not an equilibrium.

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 1, $\mathbf{X}^\infty$, 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.
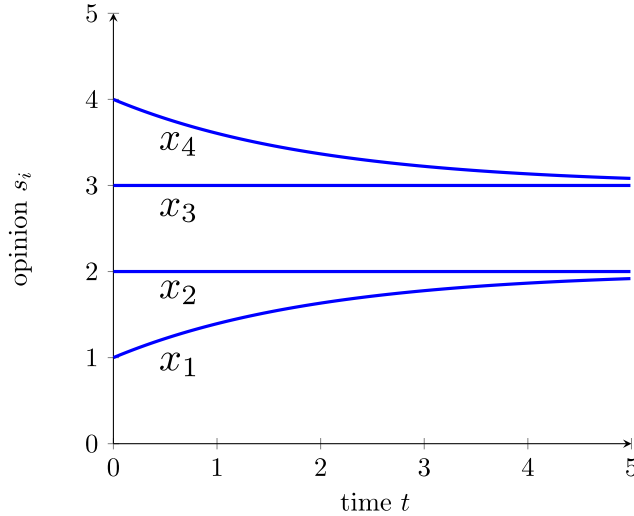
Figure 8. Preserving connectivity does not imply the convergence to a consensus. Here, when $r_* = 1$, the extreme points x_1 and x_4 will converge towards x_2 and x_3 respectively. However, x_2 and x_3 cannot move since x_1 and x_4 are always in their respective critical regions.

We will 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 1 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 1. We now prove the main result.

Theorem 1. *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* (3.5) *it will converge to a consensus.*

Proof. If $\{\mathbf{x}_i\}_i$ evolves according to the NOLB dynamics then by Corollary 1 (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 3 (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 $\mathcal{E}_p^\infty$:

$$\mathcal{N}_p^\infty = \{j \mid \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| \le 1\} \tag{4.3}$$
$$\mathcal{E}_p^\infty = \{j \mid \|\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\| = 1\}. \tag{4.4}$$

We are going to investigate 3 cases detailed in figure 9.

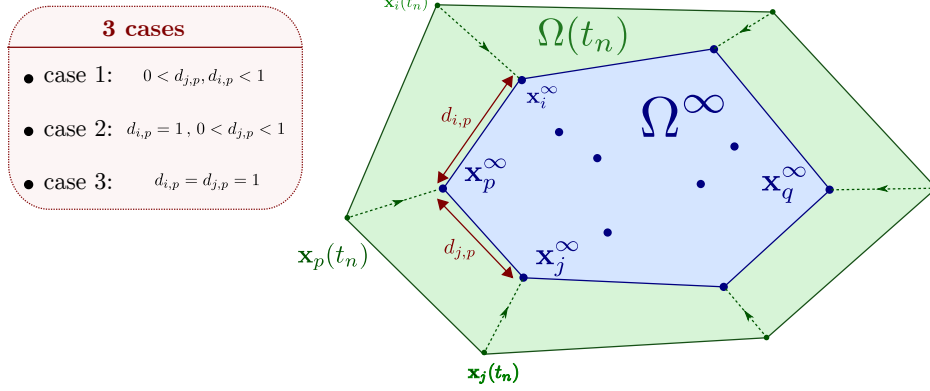• **Case 1. no extreme neighbors $\mathcal{E}_p^\infty = \emptyset$.**

Figure 9. 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 on 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 1), 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 10). 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 $\varphi_p(\overline{\mathbf{x}}_p^\infty) > 0$.

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:

$$\frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] = \langle \mathbf{u}_p, P_{\mathcal{C}_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)) \rangle. \tag{4.5}$$

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

Lemma 1. *Let $\{\mathbf{x}_i\}_i \subseteq \mathbb{R}^d$ and consider $C = \{\mathbf{v} \mid \langle \mathbf{v}, \mathbf{x}_i \rangle \geq 0 \text{ for all } i\}$. If $\mathbf{u}$ is in C then:*

$$\langle P_C(\mathbf{x}), \mathbf{u} \rangle \geq \langle \mathbf{x}, \mathbf{u} \rangle \quad \text{for all } \mathbf{x} \in \mathbb{R}^d. \tag{4.6}$$
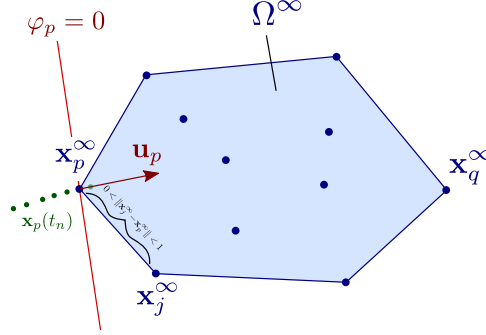
Figure 10. If the limit configuration $\{\mathbf{x}_k^\infty\}_k$ is not a consensus, the extreme point $\mathbf{x}_p(t_n)$ will eventually get inside the convex hull Ω^∞ which gives a contradiction.

It remains to show that $\mathbf{u}_p \in \mathcal{C}_p(t_n)$. Notice that $\langle \mathbf{x} - \mathbf{x}_p^\infty, \mathbf{u}_p \rangle \geq 0$ for all $\mathbf{x}$ in Ω^∞ and therefore $\mathbf{u}_p \in \mathcal{C}_{x_p^\infty}$. Eventually this is true for the approximating sequence $\mathbf{x}_p(t_n)$ as well: $\mathbf{u}_p \in \mathcal{C}_{x_{p^*}(t_n)}$ 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

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

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. $\mathbf{u}_p \in \mathcal{C}_p(t_n)$.

We can now continue our computation (4.5):

$$\frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] = \langle \mathbf{u}_p, P_{\mathcal{C}_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)) \rangle$$
$$\geq \langle \mathbf{u}_p, \overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n) \rangle \qquad (4.7)$$
$$= \varphi_p(\overline{\mathbf{x}}_p(t_n)) - \varphi_p(\mathbf{x}_p(t_n)) \quad \to \quad \varphi_p(\overline{\mathbf{x}}_p^\infty) - \varphi_p(\mathbf{x}_p^\infty) > 0$$

by continuity of φ_p. Indeed, $\overline{\mathbf{x}}_p(t_n) \overset{n \to +\infty}{\longrightarrow} \overline{\mathbf{x}}_p^\infty$ 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:

$$i) \qquad \varphi_p(\mathbf{x}_p(t_n)) \to \varphi_p(\mathbf{x}_p^\infty) = 0 \quad \text{since } \mathbf{x}_p(t_n) \to \mathbf{x}_p^\infty, \qquad (4.8)$$

$$ii) \qquad \frac{d}{dt}[\varphi_p(\mathbf{x}_p(t_n))] \geq \varphi_p(\overline{\mathbf{x}}_p^\infty) > 0 \text{ for large } t_n. \qquad (4.9)$$

● **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 connect 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,

$$\liminf_{t_n\to+\infty} \langle \mathbf{u}_p, \overline{\mathbf{x}}_p(t_n)\rangle = \liminf_{t_n\to+\infty} \left(\sum_{k=1..N,\, k\neq j} a_{pk}\langle \mathbf{u}_p, \mathbf{x}_k(t_n)\rangle + a_{pj}\langle \mathbf{u}_p, \mathbf{x}_j(t_n)\rangle \right)$$

$$\geq 0 + \frac{m}{N \cdot M}\langle \mathbf{u}_p, \mathbf{x}_j^\infty\rangle > 0.$$

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 finite-time, non-extreme 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 scenarios:

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 2). 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 8: $\|\mathbf{x}_3(t) - \mathbf{x}_2(t)\| = 1$ 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:

$$\langle P_{\mathcal{C}_p(t_n)}(\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n)), \mathbf{x}_j(t_n) - \mathbf{x}_p(t_n)\rangle = 0. \tag{4.10}$$

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,

$$\overline{\mathbf{x}}_p(t_n) - \mathbf{x}_p(t_n) \overset{n \to +\infty}{\longrightarrow} \sum_{j \in \mathcal{E}_p} a_{pj}^\infty \left(\mathbf{x}_j^\infty - \mathbf{x}_p^\infty\right). \tag{4.11}$$

Moreover, $a_{pj}^\infty = c > 0$ 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 (4.10):

$$\langle P_{C_p^\infty} \big(\sum_{k \in \mathcal{E}_p} c(\mathbf{x}_k^\infty - \mathbf{x}_p^\infty)\big), \mathbf{x}_j^\infty - \mathbf{x}_p^\infty \rangle = 0, \tag{4.12}$$

where C_p^∞ is the critical region defined by:

$$\mathcal{C}_p^\infty = \{\mathbf{v} \in \mathbb{R}^d \,|\, (\mathbf{v}, \mathbf{x}_j^\infty - \mathbf{x}_p^\infty) \geq 0 \quad \forall j \in \mathcal{E}_p\}. \tag{4.13}$$

Notice that the definition of $\mathcal{C}_p^\infty$ 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:

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{v} \rangle = 0 \qquad \text{with} \quad \mathbf{v} = \sum_{j \in \mathcal{E}_p} (\mathbf{x}_j^\infty - \mathbf{x}_p^\infty). \tag{4.14}$$

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

$$\langle P_{C_p^\infty}(\mathbf{v}), P_{C_p^\infty}(\mathbf{v}) \rangle = \langle P_{C_p^\infty}(\mathbf{v}), \mathbf{v} \rangle, \tag{4.15}$$

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 10), we have:

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{u}_p \rangle \geq \langle \mathbf{v}, \mathbf{u}_p \rangle \tag{4.16}$$

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

$$\langle P_{C_p^\infty}(\mathbf{v}), \mathbf{u}_p \rangle \geq \sum_{j \in \mathcal{E}_p} \langle \mathbf{x}_j^\infty, \mathbf{u}_p \rangle > 0, \tag{4.17}$$

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

4.2.2. *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 4 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 3. *Suppose $d(0) \leq 1$. Then:*

$$d(t) \leq d(0)e^{-\frac{m}{M} \cdot t} \tag{4.18}$$

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:

$$\begin{aligned}
\frac{d}{dt}[d^2(t)] = 2(\dot{x}_p - \dot{x}_q, x_p - x_q) &= 2\Big((\overline{x}_p - \overline{x}_q)(x_p - x_q) - (x_p - x_q)^2\Big) \\
&\leq 2\Big(|\overline{x}_p - \overline{x}_q||x_p - x_q| - |x_p - x_q|^2\Big).
\end{aligned} \tag{4.19}$$

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 (2.2). Notice that:

$$\begin{aligned}
|\overline{x}_p - \overline{x}_q| &= \left| \sum_{i=1}^{N} a_{pi}x_i - \sum_{i=1}^{N} a_{qi}x_i \right| = \left| \sum_{i=1}^{N}(a_{pi} - \eta_i)x_i - \sum_{i=1}^{N}(a_{qi} - \eta_i)x_i \right| \\
&=: \left| \sum_{i=1}^{N} \widetilde{a}_{pi}x_i - \sum_{i=1}^{N} \widetilde{a}_{qi}x_i \right|.
\end{aligned} \tag{4.20}$$

Let $\boldsymbol{\eta} = \sum_{i=1}^{N} \eta_i$. We have by (4.20) that:

$$|\overline{x}_p - \overline{x}_q| = (1 - \boldsymbol{\eta})\left| \sum_{i=1}^{N} \frac{\widetilde{a}_{pi}x_i}{1 - \boldsymbol{\eta}} - \sum_{i=1}^{N} \frac{\widetilde{a}_{qi}x_i}{1 - \boldsymbol{\eta}} \right| := (1 - \boldsymbol{\eta})\,|\widetilde{x}_p - \widetilde{x}_q| \tag{4.21}$$

Notice that

$$\sum_{i=1}^{N} \frac{\widetilde{a}_{qi}}{1 - \boldsymbol{\eta}} = \sum_{i=1}^{N} \frac{\widetilde{a}_{pi}}{1 - \boldsymbol{\eta}} = 1, \tag{4.22}$$

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 (4.21) we have that $|\overline{x}_p - \overline{x}_q| \leq (1 - \boldsymbol{\eta})|x_p - x_q|$ which by (4.19) implies:

$$\begin{aligned}
\frac{d}{dt}[d^2(t)] &\leq 2\Big((1 - \boldsymbol{\eta})|x_p - x_q||x_p - x_q| - |x_p - x_q|^2\Big) \\
&= -2\boldsymbol{\eta}|x_p - x_q|^2.
\end{aligned} \tag{4.23}$$

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 (4.23) that:

$$\frac{d}{dt}[d^2(t)] = 2d(t)d'(t) \leq -\frac{m}{M}d(t). \tag{4.24}$$

An application of Gronwall's lemma provides the final result. $\qquad \square$

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 2. *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:*

$$d(t) \leq d(0) + \eta\delta(N \cdot T - t). \tag{4.25}$$

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 3. 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:

$$\delta + 2T \quad \leq \quad \min(r_*, 1 - r_*) \tag{4.26}$$

$$\frac{T^2}{2N^2}\eta\big(\eta(1 - \delta - 2T) - 2N(\delta + 2T)\big) \quad \geq \quad \delta. \tag{4.27}$$

It is always possible to find such constant (see lemma 2).

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 $|x_p(t_* + T) - x_{p_2}(t_* + T)| \leq 1$.
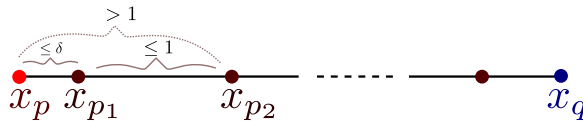

Figure 11. Situation in the case 2. The extreme point x_p needs x_{p_2} the neighbor of its neighbor x_{p_1} to be pushed further to the right.

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 (4.26). 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

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

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:

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

Coming back to x_p, we obtain:

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

using (4.27). 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 (4.25) which concludes the proof. $\qquad\square$
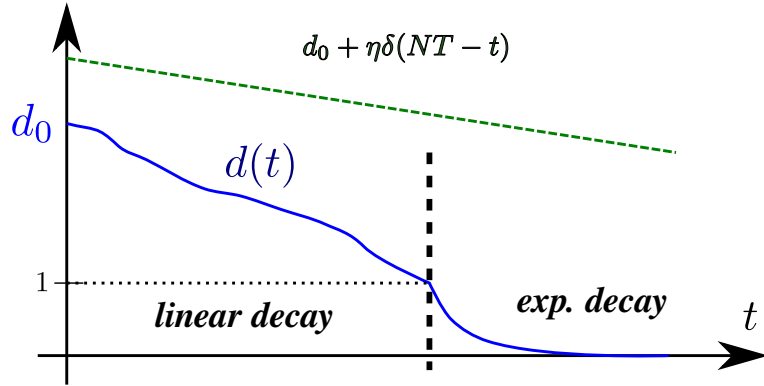
Figure 12. The decay of the diameter $d(t)$ is first linear and then exponential after the diameter $d(t)$ becomes less than 1.

Remark 5. 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 no one 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.

4.3. **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 2 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 13-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:

$$\tau = \min_{t \geq 0}\{d(t) \leq 1\}, \tag{4.28}$$

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 (4.28). The results are plotted in Figure 13-right. We find that indeed τ is lower for higher values of r_*. Interestingly, this effect seems to become less prominent for $r_* \geq 0.2$.
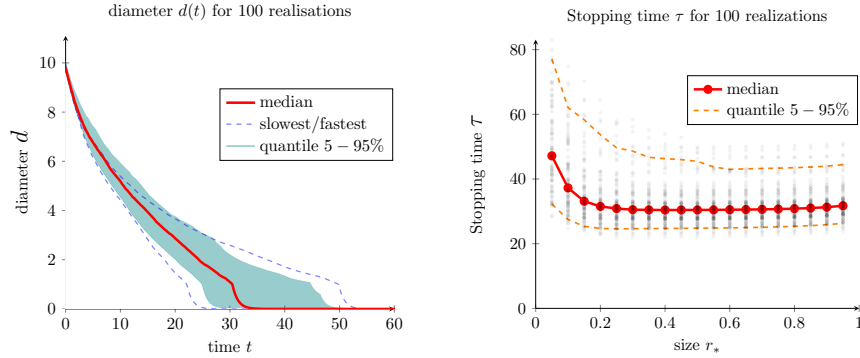

Figure 13. Left: diameter $d(t)$ over time for 100 realizations (quantile representation). **Right:** stopping time τ (4.28) depending on the size of the critical region r_*.

5. **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 (3.5) preserve connectivity between individual *agents* by Proposition 2 - 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.

5.1. **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 2). 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 14); 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 5. 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 $(j, k) \notin \widetilde{E}^B$.
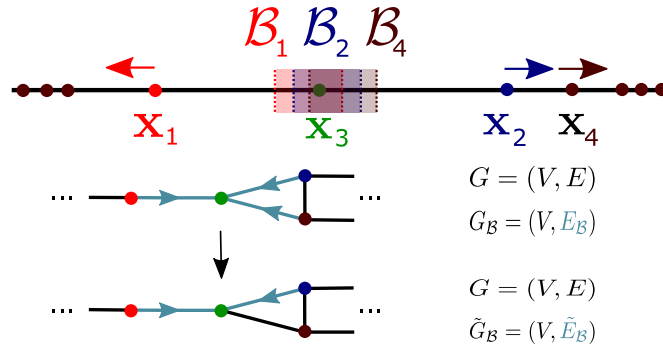

Figure 14. An example of how the behind graph can be relaxed while still ensuring that the interaction graph remains connected. The interaction graph is represented by undirected and directed edges, the behind graph is represented by only the blue directed edges. Agent 3 is in the behind region of both 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:

$$\mathbf{x}'_i = P_{\widetilde{\mathcal{C}}_i}\big(\overline{\mathbf{x}}_i - \mathbf{x}_i\big) \tag{5.1}$$

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

$$\widetilde{\mathcal{C}}_i = \{\mathbf{v} \in \mathbb{R}^d \mid \langle \mathbf{v}, \mathbf{x}_j - \mathbf{x}_i \rangle \geq 0 \quad \forall j \text{ such that } (i,j) \in \widetilde{E}^{\mathcal{B}}\}. \tag{5.2}$$

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 15 we demonstrate that under these dynamics that individual agents are indeed allowed to disconnect however the connection of the whole configuration is maintained.

Algorithm 1 Compute relaxed behind graph

1: **procedure** COMPUTE RELAXED BEHIND GRAPH
2:     Choose an order $\sigma \sim \text{Unif}(\text{permutations of } \{1, \ldots, N\})$
3:     **for** $i \in \{1, \ldots, N\}$ **do**
4:         **for** $j \in \{1, \ldots, N\}$ **do**
5:             **if** $(i,j) \in E^{\mathcal{B}}$ and there is no k such that $(i,k) \in E$ and $(k,j) \in \widetilde{E}^{\mathcal{B}}$ **then**
6:                 Add (i,j) to $\widetilde{E}^{\mathcal{B}}$
7:             **end if**
8:         **end for**
9:     **end for**
10: **end procedure**

Notice that in the example of the NOLB dynamics in Figure 15 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 4. *The RNOLB dynamics maintain connectivity of the whole configuration of agents.*

5.2. **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 (2.1) 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
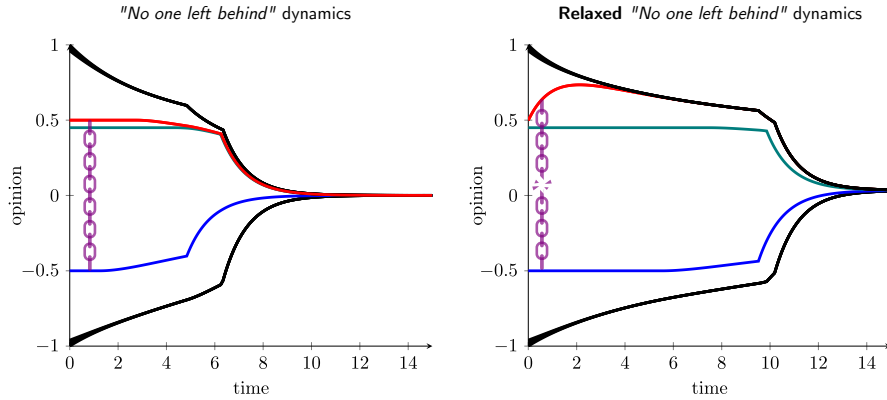
Figure 15. The NOLB dynamics do not allow the red agent to disconnect from the blue agent (illustrated with a purple chain). The RNOLB dynamics allow this disconnection to occur but maintain connectivity of the whole configuration.

if there is another agent in their critical region prevents the formation of clusters (see for example, Figures 3 and 7). 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 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 16 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.

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 16. Notice that in the initial period of cluster formation (roughly before $t = 10$)
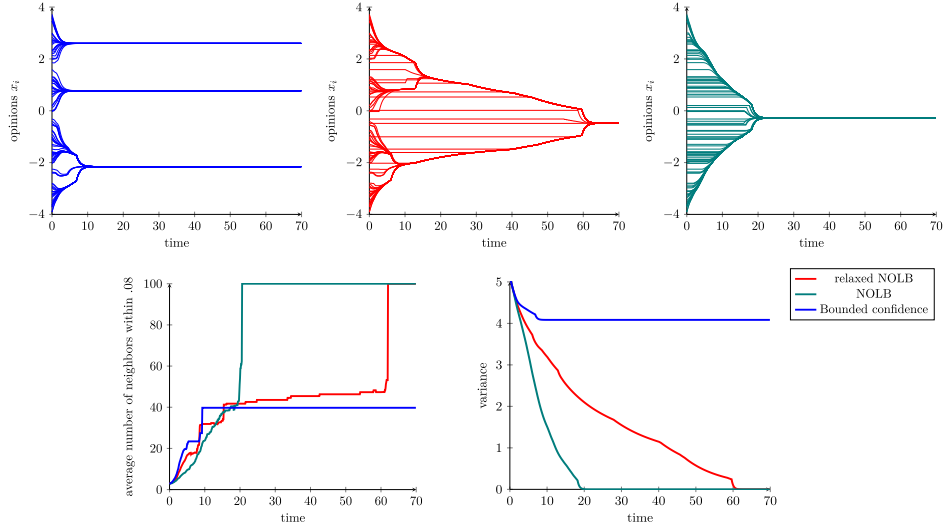
Figure 16. The RNOLB dynamics can be seen as an interpolation between NOLB and bounded confidence.

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 16. 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 16 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.

5.3. **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 - see Figure 17 for the results.

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
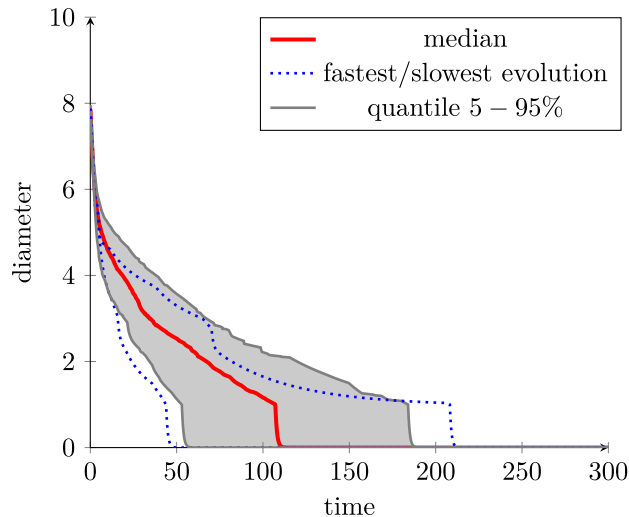case of the RNOLB dynamics as well.

Figure 17. Diameter, $d(t)$ over time for 100 realizations of the
RNOLB dynamics (quantile representation).

6. **Conclusion and future work.** In this paper we studied variants of the Hegsel-
mann -Krause bounded confidence dynamics introduced in [12]. 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 dy-
namics we introduced a variant dubbed *No one left behind* (NOLB) that maintains
connectivity between agents. We rigorously demonstrated that this control is suf-
ficient 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.

To our knowledge, models with a bounded confidence type interaction (one that depends on an interaction radius) have only been studied in a Euclidean setting. Motivated by this observation we hope to study the behavior of models with a bounded confidence style interaction in topology other than Euclidean space, eg. the circle. It would be interesting to extended this study and investigate whether the corresponding NOLB and RNOLB dynamics result in consensus in such spaces. Empirical studies of models of decentralized collective behavior outside of statistical physics have been challenging in the past due to lack of data that both describes their motion and their pairwise interactions. However, there has been some progress in recent years in biological studies of swarming animals such as birds and ants [1, 4, 22]. We believe the advent of social media provides a data source through which an empirical study of consensus in the context of opinion formation could be possible [8, 15]. A possible goal of such a study would be to confirm the phenomenon found in this study; connectivity in the interaction network of agents is a main driver of consensus.

Appendix A. Appendix.

Lemma 1. *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:

$$\text{minimize } f(y) = \frac{1}{2}\|y - x\|^2$$
$$\text{subject to } g_i(y) = \langle y, u_i \rangle \geq 0$$

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. $\qquad\square$

Lemma 2. *For any $0 < r_* < 1$, $\eta > 0$ and $N > 0$, there exist $T > 0$ and $\delta > 0$ such that:*

$$\delta + 2T \quad \leq \quad \min(r_*, 1 - r_*) \qquad \text{(A.1)}$$

$$\frac{T^2}{2N^2}\eta\big(\eta(1 - \delta - 2T) - 2N(\delta + 2T)\big) \quad \geq \quad \delta. \qquad \text{(A.2)}$$

Proof. We let $\delta = T^3$ and show that for $T > 0$ sufficiently small both equations are satisfied. Indeed, the substitution leads to:

$$T^3 + 2T \quad \leq \quad \min(r_*, 1 - r_*) \qquad \text{(A.3)}$$

$$\frac{1}{2N^2}\eta\big(\eta(1 - T^3 - 2T) - 2N(T^3 + 2T)\big) \quad \geq \quad T \qquad \text{(A.4)}$$

Since $\min(r_*, 1 - r_*) > 0$ and $T^3 + 2T \xrightarrow{T \to 0} 0$, there exists $T_1 > 0$ such that (A.3) is satisfied for $0 < T \leq T_1$. Similarly, for the equation (A.4), we notice that

$$\frac{1}{2N^2}\eta\big(\eta(1 - T^3 - 2T) - 2N(T^3 + 2T)\big) \xrightarrow{T \to 0} \frac{\eta^2}{2N^2} > 0.$$

Therefore, there exists $T_2 > 0$ such that (A.4) is satisfied for $0 < T \leq T_2$. Taking $T = \min(T_1, T_2)$ and $\delta = T^3$, we deduce a solution to (4.26)-(4.27). $\qquad\square$

REFERENCES

[1] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna and E. Cisbani, et al., Empirical investigation of starling flocks: A benchmark study in collective animal behaviour, *Animal Behaviour*, **76** (2008), 201–215.

[2] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, On Krause's multi-agent consensus model with state-dependent connectivity, *IEEE Trans. Automat. Control*, **54** (2009), 2586–2597.

[3] V. D. Blondel, J. M. Hendrickx and J. N. Tsitsiklis, On the 2R conjecture for multi-agent systems, *2007 European Control Conference (ECC)*, Kos, Greece, 2007.

[4] J. Buhl, D. J. T. Sumpter, I. D. Couzin, J. J. Hale and E. Despland, et al., From disorder to order in marching locusts, *Science*, **312** (2006), 1402–1406.

[5] M. Caponigro, M. Fornasier, B. Piccoli and E. Trélat, Sparse stabilization and optimal control of the Cucker-Smale model, *Math. Control Relat. Fields*, **3** (2013), 447–466.

[6] C. Castellano, S. Fortunato and V. Loreto, Statistical physics of social dynamics, *Rev. Mod. Phys.*, **81** (2009), 591–646.

[7] G. Deffuant, D. Neau, F. Amblard and G. Weisbuch, Mixing beliefs among interacting agents, *Adv. Complex Syst.*, **3** (2000), 87–98.

[8] E. Estrada, E. Vargas-Estrada and H. Ando, Communicability angles reveal critical edges for network consensus dynamics, *Phys. Rev. E (3)*, **92** (2015), 10pp.

[9] K. Garimella, G. De Francisci Morales, A. Gionis and M. Mathioudakis, Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship, *Proc. 2018 World Wide Web Conference*, 2018, 913–922.

[10] E. Gilbert, T. Bergstrom and K. Karahalios, Blogs are echo chambers: Blogs are echo chambers, 2009 42nd Hawaii International Conference on System Sciences, Big Island, HI, 2009.

[11] D. Goldie, M. Linick, H. Jabbar and C. Lubienski, Using Bibliometric and social media analyses to explore the "echo chamber" hypothesis, *Educational Policy*, **28** (2014).

[12] R. Hegselmann and U. Krause, Opinion dynamics and bounded confidence: models, analysis and simulation, *J. Artificial Societies Social Simulation*, **5** (2002).

[13] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*, Grundlehren Text Editions, Springer-Verlag, Berlin, 2001.

[14] P.-E. Jabin and S. Motsch, Clustering and asymptotic behavior in opinion formation, J. Differential Equations, **257** (2014), 4165–4187.

[15] D. Kempe, J. Kleinberg and E. Tardos, Maximizing the spread of influence through a social network, Proc. Ninth ACM SIGKDD Internat. Conference Knowledge Discovery Data Mining, 2003, 137–146.

[16] U. Krause, A discrete nonlinear and non-autonomous model of consensus formation, Communications in Difference Equations, Gordon and Breach, Amsterdam, 2000, 227–236.

[17] J. Lorenz, Continuous opinion dynamics under bounded confidence: A survey, Internat. J. Modern Phys. C, **18** (2007), 1819–1838.

[18] J. Lorenz, Consensus strikes back in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence, J. Artificial Societies Social Simulation, (2006).

[19] S. Motsch and E. Tadmor, Heterophilious dynamics enhances consensus, SIAM Rev., **56** (2014), 577–621.

[20] R. Olfati-Saber, J. A. Fax and R. M. Murray, Consensus and cooperation in networked multi-agent systems, Proc. IEEE, **95** (2007), 215–233.

[21] B. Piccoli, N. Pouradier Duteil and E. Trélat, Sparse control of Hegselmann–Krause models: Black hole and declustering, SIAM J. Control Optim., **57** (2019), 2628–2659.

[22] L.-A. Poissonnier, S. Motsch, J. Gautrais, J. Buhl and A. Dussutour, Experimental investigation of ant traffic under crowded conditions, eLife, **8** (2019).

[23] W. Quattrociocchi, A. Scala and C. R. Sunstein, Echo Cchambers on Facebook, SSRN, in progress.

[24] R. O. Saber and R. M. Murray, Consensus protocols for networks of dynamic agents, Proc. 2003 American Control Conference, Denver, CO, 2003.

[25] D. Spanos, R. Olfati-Saber and R. Murray, Dynamic consensus on mobile networks, IFAC World Congress, Citeseer, 2005, 1–6.

[26] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen and O. Shochet, Novel type of phase transition in a system of self-driven particles, Phys. Rev. Lett., **75** (1995), 1226–1229.

[27] D. Weber, R. Theisen and S. Motsch, Deterministic versus stochastic consensus dynamics on graphs, J. Stat. Phys., **176** (2019), 40–68.

[28] H. Xia, H. Wang and Z. Xuan, Opinion dynamics: A multidisciplinary review and perspective on future research, Internat. J. Knowledge Syst. Sci. (IJKSS), **2** (2011), 72–91.

[29] W. Yu, G. Chen, M. Cao and J. Kurths, Second-order consensus for multiagent systems with directed topologies and nonlinear dynamics, IEEE Trans. Syst. Man Cybernetics, Part B, **40** (2010), 881–891.

E-mail address: guanlinl@gatech.edu
E-mail address: smotsch@asu.edu
E-mail address: djweber3@asu.edu