Swarm Aggregation in Reciprocal Multi-Agent Systems with Fading Interaction Laws

Xudong Chen *

Abstract

In this paper, we will investigate a special class of multi-agent systems, as we call in this paper, the class of Reciprocal Multi-Agent systems. This class of multi-agent systems models a natural rule by which agents interact with its neighbors in a reciprocal way. By this rule, the dynamical system then evolves as a gradient flow on the configuration space. We develop, among other things, one basic property of the gradient flow associated with the model. We show that there is no collision of agents along the gradient flow, and moreover, there is no escape of agents to infinity so that the swarm aggregation is achieved. An upper/lower bound for the distance between any two agents in an equilibrium is also established. In the development of swarm aggregation, we establish a modern approach for proving the convergence of the gradient flow. In particular, we introduce a parametrized notion of clustering as a way of partitioning agents in a dilute configuration. This is a rich question relating to classic techniques such as $k$-mean algorithm and its variants, and is useful for studying multi-agent systems.

1 Introduction

Over the last two decades, multi-agent systems with reciprocal interaction laws, as we call in this paper reciprocal multi-agent systems (or in short RMA systems), have been investigated under various assumptions and from different perspectives. Questions about stability, robustness and etc. have all been treated to some degree (see, for example, [1–3,6–9,11–13,15]).

Each RMA system is defined by a connected undirected graph, together with a family of interaction functions. Let $\Gamma = (V,E)$ be the graph with $V := \{1,\cdots,N\}$ the set of vertices and $E$ the set of edges. Consider the motion of a set of $N$ agents in a purely kinematic model whereby agent $\vec{x}_i$ feels the presence of $\vec{x}_j$ if and only if they are adjacent, i.e, there is an edge between $i$ and $j$ in graph $\Gamma$, and they interact with each other reciprocally through an effect depending only on the pairwise distance between them. The equations of

*Xudong Chen is with CSL at UIUC, Champaign, IL 61820, USA xdchen@illinois.edu
motion of agents $\vec{x}_1, \cdots, \vec{x}_N \in \mathbb{R}^n$ with $N > n$ take the form

$$\dot{\vec{x}}_i = \sum_{j=1}^{N} g_{ij}(d_{ij})(\vec{x}_j - \vec{x}_i), \quad i = 1, \cdots, N \quad (1)$$

Each $g_{ij} : \mathbb{R}^+ \to \mathbb{R}$ is a continuous differentiable function depending only on the mutual distance between $\vec{x}_i$ and $\vec{x}_j$. We require $g_{ij} = g_{ji}$ for all $ij$, i.e., interactions between agents are reciprocal. We further require $g_{ij}$ be identically zero if $\vec{x}_i$ and $\vec{x}_j$ are not adjacent. In this paper, we will impose two conditions on each $g_{ij}$, $ij \in E$, and they are:

**Strong repulsion.** $\lim_{d \to 0} dg_{ij}(d) = -\infty$ and $\lim_{d \to 0} \int_{d}^{1} x g_{ij}(x) dx = -\infty$.

**Fading attraction.** $g_{ij}(d) > 0$ for sufficiently large $d$ and $\lim_{d \to \infty} dg_{ij}(d) = 0$.

Notice that it is $dg_{ij}(d)$ that represents the actual attraction/repulsion between $\vec{x}_i$ and $\vec{x}_j$, so for convenience, we let $\bar{g}_{ij}(d) := dg_{ij}(d)$.

As interactions between pairs of agents are reciprocal, consequently the centroid of the configuration is invariant along the evolution, so we may assume that the centroid of the configuration is located at the origin. We also observe that an interaction law $g_{ij}$ produces infinite repulsion at zero separation. Thus equation (1) is only well defined if we agree to limit our attention to configurations in which $\vec{x}_i$ doesn’t collide with $\vec{x}_j$ for each $ij \in E$. Thus our model should be thought of as being defined on an appropriate open subset of $\mathbb{R}^{n \times N}$.

So the configuration space is defined by

$$P := \{ (\vec{x}_1; \cdots; \vec{x}_N) \in \mathbb{R}^{n \times N} \mid \sum_{i=1}^{N} \vec{x}_i = 0 \text{ and } \vec{x}_i \neq \vec{x}_j, \forall ij \in E \} \quad (2)$$

An important observation is that the equations we adopt to describe the formation of the configuration are actually a gradient flow over $P$. The potential function associated is a symmetric function of the individual agents and is given by

$$\Psi(\vec{x}_1, \cdots, \vec{x}_N) := \sum_{ij \in E} \int_{1}^{d_{ij}} \bar{g}_{ij}(x) dx \quad (3)$$

In this paper, we will develop, among other things, one basic property of this class of RMA systems.

**Theorem 1.1.** The dynamics of the RMA system described by equation (1) is a gradient flow associated with the potential function $\Psi$ over the configuration space $P$. The set of equilibria associated with the gradient flow is compact, in particular, there exist two positive numbers $l_-$ and $l_+$ such that the distance between any two adjacent agents in an equilibrium configuration lies in the closed interval $[l_-, l_+]$. Moreover, for each initial condition, the gradient flow exists for all time and converges to the set of equilibria.

After this introduction, we proceed in steps. In section 2, we establish the potential function over the configuration space and develop a metric property.
of the equilibria associated with the gradient flow. In section 3, we show that
the resulting gradient flow of the RMA system is collision-free, so in particular,
the solution of the dynamical system exists for all time. The rest of this paper,
from section 4 to section 7, are about the analysis of swarm aggregation. In
section 4, we introduce a parametrized clustering of agents, and describe various
aspects associated with it. In section 5, we introduce the notion of dissipation
zone, and establish a lower bound for the loss of potential of a gradient flow
whenever it reaches a dissipation zone. In section 6, we introduce the notion of
semi-diverging gradient flow, and investigate its path-behavior by relating it to
clustering of agents. In section 7, we use the results established in the earlier
sections to prove the convergence of the gradient flow.

2 Existence of solution and bounded size of equilibria

For each initial condition \( p \), we let \( \varphi_t(p) \) be the solution of the gradient flow
at time \( t \). If the solution exists for all time \( t > 0 \), we then simply denote by
\( \varphi_{\geq 0}(p) \) the entire gradient flow. Let \( a, b \) be two numbers with \( 0 \leq a \leq b \leq \infty \),
we define a subset of \( P \) as

\[
P_{ab} := \{ p \in P | a \leq d_{ij} \leq b, ij \in E \}
\]

In this paper, we will simply write \( P^b \) if \( a = 0 \) and \( P^a \) if \( b = \infty \). Our goal in
this section is to develop theorem 2.1 and theorem 2.2.

**Theorem 2.1.** Consider the RMA system described by equation \( \text{(1)} \). For each
initial condition \( p \in P \), the solution of the gradient flow exists for all time. In
fact, there exists a number \( a > 0 \) associated with \( p \) such that
\( \varphi_{\geq 0}(p) \subset P^a \).

**Theorem 2.2.** Consider the RMA system described by equation \( \text{(1)} \). There
exist two positive numbers \( l_− \) and \( l_+ \) such that the distance between any two
adjacent agents in any equilibrium associated with the gradient flow lies in the
closed interval \([l_−, l_+]\), i.e, the set of equilibria is contained in \( P_{l_−}^{l_+} \).

We will first prove theorem 2.1, and then develop the metric property asso-
ciated with equilibria.

**Proof of theorem 2.1.** It suffices to show that there is a positive number \( a \)
associated with the initial condition \( p \) such that if \( \varphi_t(p) \) exists, then \( \varphi_t(p) \in P^a \).
Let

\[
\psi_0 := \inf \{ \int_1^d \bar{g}_{ij}(x) dx | d \in \mathbb{R}_+, ij \in E \}
\]

By conditions of strong repulsion and fading attraction, we know \( \psi_0 \) exists. On
the other hand, by condition of strong repulsion alone, we know exists a positive
number \( a' \) such that for any \( a' \in (0, a) \) and for any \( ij \in E \), we have

\[
\int_1^{a'} \bar{g}_{ij}(x) dx + (|E| - 1) \psi_0 > \Psi(p)
\]
where $|E|$ is the size of the graph $\Gamma$. The potential $\Psi(\varphi_t(p))$, as a function of $t$, is non-increasing along the evolution, so inequality (6) implies that at any time $t > 0$, the distance between any two adjacent agents in $\varphi_t(p)$ is bounded below by $a$. This then establishes the theorem.

The rest of this section is devoted to the proof of theorem 2.2. We will prove the existence of upper bound $l_+$ and the existence of lower bound $l_-$ in lemma 2.3 and corollary 2.5, respectively.

Before going on, we first define two positive numbers associated with the family of interaction laws. Let $\alpha$ and $\beta$ be defined so that

\[
g_{ij}(d) < 0, \quad \forall i,j \in E \quad \& \quad \forall d \in (0,\alpha)
\]

\[
g_{ij}(d) > 0, \quad \forall i,j \in E \quad \& \quad \forall d \in (\beta,\infty)
\]

These two numbers exist by conditions of \textit{strong repulsion} and \textit{fading attraction}. In fact, let $\alpha_{ij}$ and $\beta_{ij}$ be the first and last zero of $g_{ij}$, $ij \in E$, then we may just set

\[
\alpha = \min\{\alpha_{ij}| ij \in E]\]

\[
\beta = \max\{\beta_{ij}| ij \in E\}
\]

We now state lemma 2.3.

**Lemma 2.3.** Consider the RMA system described by equation (1). There is a positive number $l_+$ such that the distance between any two adjacent agents in an equilibrium is less than $l_+$.

**Proof.** The proof is done by contradiction. We assume that for any distance $d$, there is an equilibrium $p$ with a pair of adjacent agents, say $\vec{x}_1$ and $\vec{x}_N$, such that $d_{1N} \geq d$. We now choose $d := N\beta$.

Let $x_i^j$ be the $j$-th coordinate of agent $\vec{x}_i$. Rotate the configuration, if necessary, so that $\vec{x}_N - \vec{x}_1$ lies in the $x_1^1$-axis, and we assume $x_1^1 < x_1^N$. To keep $\vec{x}_1$ balanced, there is at least one agent, say $\vec{x}_2$, such that $x_2^1 \leq x_1^1 + \beta$ because otherwise, the interaction between $\vec{x}_1$ and $\vec{x}_i$, $1i \in E$, is an attraction, and the projection of each attraction on agent $\vec{x}_1$ is positive along the $x_1^1$-axis and hence, $\vec{x}_1$ can’t be balanced.

Consequently, we have $d_{2N} \geq (N-1)\beta > \beta$, so by assumption $g_{2N}(d_{2N}) \geq 0$. This, in particular, implies that

\[
g_{1N}(d_{1N})(x_N^1 - x_1^1) + g_{2N}(d_{2N})(x_N^1 - x_2^1) > 0
\]

In other words, agent $\vec{x}_N$ attracts $\vec{x}_1$ and $\vec{x}_2$, as a whole, along $x_1^1$-axis. So there must exist an agent among $\vec{x}_1$ and $\vec{x}_2$, say $\vec{x}_2$, such that

\[
g_{12}(d_{12})(x_1^1 - x_2^1) + g_{2N}(d_{2N})(x_N^1 - x_2^1) > 0
\]

Then to keep $\vec{x}_2$ balanced, there is at least one agent $\vec{x}_3$ such that $x_3^1 \leq x_2^1 + \beta \leq x_1^1 + 2\beta$. 

4
Since \( d_{3N} \geq (N - 2)\beta > \beta \), we have \( g_{3N}(d_{3N}) \geq 0 \). So \( \vec{x}_N \) attracts \( \vec{x}_1, \vec{x}_2 \) and \( \vec{x}_3 \), as a whole, along \( x^1 \)-axis. Consequently there is an agent among \( \vec{x}_1, \vec{x}_2 \) and \( \vec{x}_3 \), say \( \vec{x}_3 \), such that

\[
\sum_{i=1}^{2} g_{i3}(d_{i3})(x_i^1 - x_i^3) + g_{3N}(d_{3N})(x_N^1 - x_N^3) > 0
\]

Similarly, to keep \( \vec{x}_3 \) balanced, we locate agent \( \vec{x}_4 \) with \( x_4^1 \leq x_1^1 + 3\beta \).

Repeat the process, we then get \( x_k^1 \leq x_1^1 + (k - 1)\beta \) for each \( k = 1, \cdots, N - 1 \). Consequently \( d_{kN} > \beta \) for all \( k \neq N \). But then,

\[
\sum_{i=1}^{N-1} g_{iN}(d_{iN})(x_i^1 - x_N^3) < 0
\]

So, agent \( \vec{x}_N \) can’t be balanced. This contradicts to the assumption that \( p \) is an equilibrium. \( \square \)

The existence of upper bound \( l_+ \) is now clear. We will now prove the existence of lower bound \( l_- \). We start with a construction of a particular subset of the configuration space. Let \( \epsilon \) and \( r \) be two positive numbers, let \( \vec{x}_0 \) be the center of \( \vec{x}_1 \) and \( \vec{x}_2 \), i.e, \( \vec{x}_0 := \frac{1}{2}(\vec{x}_1 + \vec{x}_2) \), and let \( d_{i0} \) be the Euclidean distance between \( \vec{x}_i \) and \( \vec{x}_0 \). We then define a subset of \( P \) as

\[
Z(\epsilon, r, N) := \{ p \in P \mid d_{ij} \geq d_{12} = \epsilon, \forall i, j \in E \}
\]

So for each configuration \( p \) in \( Z(\epsilon, r, N) \), the distance \( d_{12} \) is minimal among all mutual distances between adjacent agents in \( p \), and all agents are contained in the closed ball of radius \( r \) centered at \( \vec{x}_0 \).

In this paper, we let \( f(p) \) be the gradient vector field associated with \( \Psi \), i.e,

\[
f(p) := -\nabla \Psi(p)
\]

and let \( f_i(p) \) be the restriction of \( f(p) \) to agent \( \vec{x}_i \), i.e,

\[
f_i(p) := \sum_{j=1}^{N} g_{ij}(d_{ij})(\vec{x}_j - \vec{x}_i)
\]

Let \( \tilde{p} \) be a sub-configuration of \( p \) formed by agents \( \vec{x}_{i_1}, \cdots, \vec{x}_{i_m} \) with \( i_1 < \cdots < i_m \). Similarly, we can define \( \tilde{f}(\tilde{p}) \) and \( \tilde{f}_{ij}(\tilde{p}) \) as if \( \tilde{p} \) is isolated. To be explicit, let \( \tilde{\Gamma} \) be a subgraph of \( \Gamma \) defined by restricting \( \Gamma \) to vertices \( i_1, \cdots, i_m \), and let

\[
\tilde{\Psi}(\vec{x}_{i_1}, \cdots, \vec{x}_{i_m}) := \sum_{i, j, k} \int_{1}^{d_{i_j, i_k}} \tilde{g}_{i_j, i_k}(x)dx
\]

be a sub-potential of \( \Psi \) by restricting \( \Psi \) to the subgraph \( \tilde{\Gamma} \). We then define

\[
\tilde{f}(\tilde{p}) := -\nabla \tilde{\Psi}(\tilde{p})
\]

and let \( \tilde{f}_{ij}(\tilde{p}) \) be the restriction of \( \tilde{f}(\tilde{p}) \) to agent \( \vec{x}_{i_j} \).
Lemma 2.4. Let $p$ be a configuration in $Z(\epsilon, r, N)$, and let $\vec{e}_{0i}$ be a unit vector defined by

$$
\vec{e}_{0i} := \frac{\vec{x}_i - \vec{x}_0}{|\vec{x}_i - \vec{x}_0|}
$$

so the projection of $f_i(p)$ along $\vec{e}_{0i}$ is $(f_i(p), \vec{e}_{0i})$. Let

$$
\xi(p) := \max\{\langle f_i(p), \vec{e}_{0i} \rangle | i = 1, \ldots, N \}
$$

and let

$$
\eta(\epsilon, r, N) := \inf\{\xi(p) | p \in Z(\epsilon, r, N)\}
$$

If the vertices 1 and 2 are adjacent in $\Gamma$, then $\lim_{\epsilon, r \to 0} \eta(\epsilon, r, N) = +\infty$. The result holds for all $N \geq 2$, and it doesn’t depend on how $\epsilon$ and $r$ go to zero.

Proof. There are choices of $(\epsilon, r, N)$ such that $Z(\epsilon, r, N)$ is the empty set, then $\eta(\epsilon, r, N) = +\infty$. So it suffices to consider the case $Z(\epsilon, r, N)$ is nonempty. The proof is done by induction on the number of agents.

Base case: consider the case where we have only two agents $\vec{x}_1$ and $\vec{x}_2$. By condition of strong repulsion on $g_{12}$, we have

$$
\lim_{\epsilon, r \to 0} \eta(\epsilon, r, 2) = \lim_{\epsilon \to 0} |\bar{g}_{12}(\epsilon)| = \infty
$$

this establishes the base case.

Inductive step: we assume that the lemma holds for $N \leq k - 1$ and we prove for the case $N = k$. The proof is done by contradiction, i.e, we assume for any $\epsilon > 0$ and $r > 0$, there exist $\epsilon'$ and $r'$ with $0 < \epsilon' \leq \epsilon$ and $0 < r' \leq r$, together with a configuration $p \in Z(\epsilon', r', k)$ such that $\xi(p)$ is bounded above by a fixed number $\xi_0 > 0$.

Since the graph $\Gamma$ is connected, there is a chain of subgraphs

$$
\emptyset = \Gamma_0 \subset \Gamma_1 \cdots \subset \Gamma_k = \Gamma
$$

each $\Gamma_i$ is a connected subgraph consisting exactly of $i$ vertices. Relabel the vertices, if necessary, so that the set of vertices associated with $\Gamma_i$ is $\{1, \ldots, i\}$ for each $i = 1, \ldots, k$.

Let $\hat{p}$ be a sub-configuration of $p$ formed by agents $\vec{x}_1, \ldots, \vec{x}_{k-1}$. By induction, $\xi(\hat{p})$ can be made arbitrarily large by shrinking $\epsilon$ and $r$. So choose a positive number $K$, sufficiently large, such that $\xi(\hat{p}) = K\xi_0$. We may as well assume that $\epsilon$ and $r$ are small enough so that all interactions among $p$ are repulsions.

Suppose $\xi(\hat{p}) = \langle f_j(\hat{p}), \vec{e}_{0j} \rangle$, and we consider the projection of $f_j(p)$ along $\vec{e}_{0j}$, we have

$$
\langle f_j(p), \vec{e}_{0j} \rangle = \xi(\hat{p}) + g_{jk}(d_{jk}) \langle \vec{x}_k - \vec{x}_j, \vec{e}_{0j} \rangle
$$

So to keep $\langle f_j(p), \vec{e}_{0j} \rangle \leq \xi(p) \leq \xi_0$, we must have

$$
g_{jk}(d_{jk})(\vec{x}_j - \vec{x}_k, \vec{e}_{0j}) \geq (K - 1)\xi_0
$$
This, in particular, implies \( \langle \vec{x}_k - \vec{x}_j, \vec{e}_{ij} \rangle > 0 \). An important observation is then \( d_{i0} > d_{j0} \) and \( \langle \vec{x}_k - \vec{x}_j, \vec{e}_{0j} \rangle < \langle \vec{x}_k - \vec{x}_j, \vec{e}_{0k} \rangle \). So if \( \vec{x}_j \) is the only agent that is adjacent to \( \vec{x}_k \), then

\[
\langle f_k(p), \vec{e}_{0k} \rangle = g_{jk}(d_{jk})\langle \vec{x}_j - \vec{x}_k, \vec{e}_{0k} \rangle > g_{jk}(d_{jk})\langle \vec{x}_j - \vec{x}_k, \vec{e}_{0j} \rangle > (K - 1)\xi_0 \tag{25}
\]

which is a contradiction because \( K \) can be made arbitrarily large.

So there is at least one agent, say \( \vec{x}_{i_1} \), other than \( \vec{x}_j \) such that the two vertices \( i_1 \) and \( k \) are adjacent. In addition, we have

\[
g_{1,k}(d_{1,k})\langle \vec{x}_k - \vec{x}_{i_1}, \vec{e}_{0k} \rangle > \frac{1}{k} [g_{jk}(d_{jk})\langle \vec{x}_j - \vec{x}_k, \vec{e}_{0k} \rangle - \xi_0] \geq \frac{K - 2}{k}\xi_0 \tag{26}
\]

because otherwise

\[
\langle f_k(p), \vec{e}_{0k} \rangle \geq g_{jk}(d_{jk})\langle \vec{x}_j - \vec{x}_k, \vec{e}_{0k} \rangle + \sum_{i \neq j, i \neq k} g_{ik}(d_{ik})\langle \vec{x}_i - \vec{x}_k, \vec{e}_{0k} \rangle > \xi_0 \tag{27}
\]

We assume \( K > 2 \), inequality \([26]\) then implies \( \langle \vec{x}_{i_1} - \vec{x}_k, \vec{e}_{0k} \rangle > 0 \), consequently we have \( d_{i_1,0} > d_{k0} \) and \( \langle \vec{x}_{i_1} - \vec{x}_k, \vec{e}_{0k} \rangle < \langle \vec{x}_{i_1} - \vec{x}_k, \vec{e}_{0i_1} \rangle \).

We now make a change of variable as \( K' := (K - 2)/k + 1 \), so then

\[
g_{i_1,k}(d_{i_1,k})\langle \vec{x}_k - \vec{x}_{i_1}, \vec{e}_{0i_1} \rangle > g_{i_1,k}(d_{i_1,k})\langle \vec{x}_k - \vec{x}_{i_1}, \vec{e}_{0k} \rangle > (K' - 1)\xi_0 \tag{28}
\]

By identifying inequality \([28]\) with inequality \([24]\), we then conclude that there is an agent \( \vec{x}_{i_2} \) adjacent to \( i_1 \) such that

\[
g_{i_2,i_1}(d_{i_2,i_1})\langle \vec{x}_{i_1} - \vec{x}_{i_2}, \vec{e}_{0i_1} \rangle > \frac{K' - 2}{k}\xi_0 \tag{29}
\]

So again \( \langle \vec{x}_{i_2} - \vec{x}_{i_1}, \vec{e}_{0i_1} \rangle > 0 \) and hence, \( d_{i_20} > d_{i_10} \).

Repeat the process, we then get a sequence of agents \( \vec{x}_{i_1}, \vec{x}_{i_2}, \cdots \) with \( d_{i_10} > d_{i_{k+1}0} \) for all \( k = 1, 2, \cdots \). In fact, this sequence must be infinite because \( K \) can be made arbitrarily large. This then contradicts to the fact that there are only finite number of agents. \( \square \)

The next corollary establishes the existence of lower bound \( l_- \).

**Corollary 2.5.** Consider the RMA system described by equation \([1]\). There is a positive number \( l_- \) such that the distance between any two adjacent agents in an equilibrium is greater than \( l_- \).

**Proof.** The proof is done by contradiction, i.e, we assume that for any \( \epsilon > 0 \), there is an equilibrium \( p \) with two adjacent agents, say \( \vec{x}_1 \) and \( \vec{x}_2 \), such that \( d_{12} \leq \epsilon \). We may as well assume that \( d_{12} \) is the minimum distance between adjacent agents.

Let \( \zeta_0 \) be a positive number, by condition of strong repulsion, we can choose \( \epsilon \) such that \( |\tilde{g}_{ij}(d)| > \zeta_0 \) for any \( d \in (0, \epsilon) \) and for any \( ij \in E \).

\[
p_1 := \sup\{d \in \mathbb{R}^+ | |\tilde{g}_{ij}(d)| = \zeta_0/N \text{ for some } ij \in E\} \tag{30}
\]
we assume $\zeta_0$ is large enough so that $\rho_1 < \alpha$. Then to keep $\vec{x}_1$ balanced, there is at least one agent, say $\vec{x}_3$, with $d_{13} \leq \rho_1$ because otherwise,

$$\sum_{i=3}^{N} g_{1i}(d_{i1})(\vec{x}_i - \vec{x}_1) < \sum_{i=3}^{N} |\bar{g}_{1i}(d_{i1})| < \sum_{i=3}^{N} |\bar{g}_{1i}(\rho_1)| < \zeta_0 < |\bar{g}_{12}(d_{12})|$$  \hspace{1cm} (31)

and hence, agent $\vec{x}_1$ can’t be balanced. Let $\vec{x}_0$ be the center of $\vec{x}_1$ and $\vec{x}_2$, then the three agents $\vec{x}_1$, $\vec{x}_2$ and $\vec{x}_3$ are contained in the open ball $B_{r_1}(\vec{x}_0)$ centered at $\vec{x}_0$ with radius $r_1 := \rho_1 + \epsilon$. Notice that both $\epsilon$ and $\rho_1$ go to zero as $\zeta_0$ goes to infinity, so $r_1$ can be made arbitrarily small by increasing $\zeta_0$.

Let $\hat{\rho}$ be the sub-configuration formed by agents $\vec{x}_1$, $\vec{x}_2$ and $\vec{x}_3$, and let $\xi(\hat{\rho})$ be defined by equation (19) as if $\hat{\rho}$ is isolated from other agents. Then by lemma 2.4, we can make $\xi(\hat{\rho})$ arbitrarily large by shrinking both $\epsilon$ and $r_1$. Let $\zeta_1 := \xi(\hat{\rho})$, and let

$$\rho_2 := \sup\{d \in \mathbb{R}^+ | |\bar{g}_{ij}(d)| = \zeta_1/N \text{ for some } ij \in E\}$$  \hspace{1cm} (32)

again, we assume $\zeta_1$ is sufficiently large so that $\rho_2 < \alpha$. Suppose $\zeta_1 = (f_3(\hat{\rho}), \bar{\epsilon}_0)$, then to keep $\vec{x}_3$ balanced, we have to place at least one agent, say $\vec{x}_4$, around $\vec{x}_3$ such that $d_{34} < \rho_2$. In other words, the four agents $\vec{x}_1, \vec{x}_2, \vec{x}_3$ and $\vec{x}_4$ are contained in the open ball $B_{r_2}(\vec{x}_0)$ with $r_2 := \epsilon + \rho_1 + \rho_2$, and $r_2$ can be made arbitrarily small by increasing $\zeta_0$.

Repeat the process, we then successively locate $\vec{x}_5, \cdots, \vec{x}_N$. All of the agents are contained in an open ball $B_{r}(\vec{x}_0)$. Moreover, the radius $r$ approaches to zero as $\zeta_0$ goes to infinity. If $r$ is sufficiently small, then $\xi(\rho) > 0$ by lemma 2.4. This contradicts to the assumption that $\rho$ is an equilibrium. \hfill $\square$

By combining lemma 2.3 and corollary 2.5, we then prove theorem 2.2. Notice that the set $P^+_l$ is compact, so the set of equilibria, as a closed subset in a compact set, is also compact. In the rest of this paper, we will show that each gradient flow converges to the set of equilibria. In other words, the set of equilibria is a global compact attractor of the gradient system.

### 3 Clustering agents in dilute configurations

A parametrized clustering $\sigma(l, \epsilon)$ with parameters $l > 0$ and $\epsilon > 0$ on a configuration $p$ is a partition of agents of $p$ into disjoint union of clusters $C_1, \cdots, C_M$. The partition has to satisfy three conditions:

1. Let $\Gamma_i$ be a subgraph of $\Gamma$ associated with $C_i$, i.e, if $C_i = \{\vec{x}_{i1}, \cdots, \vec{x}_{ik}\}$, then $\Gamma_i$ is the restriction of $\Gamma$ to vertices $i_1, \cdots, i_k$. We require each $\Gamma_i$ be connected.

2. Let $C_i$ and $C_j$ be two adjacent clusters, i.e, there is an agent $\vec{x}_i$ in $C_i$ and an agent $\vec{x}_j$ in $C_j$ such that the two agents are adjacent. Let $l_{ij}$ be an adjacent-cluster distance, i.e, $l_{ij}$ is the distance between centers of two adjacent clusters $C_i$ and $C_j$. We then require $l_{ij}$ be greater than $l$. 

8
3. Let $r_i$ be the radius of cluster $C_i$, then $r_i/l_{ij} < \epsilon$ for all $j$ with $C_j$ adjacent to $C_i$.

Any configuration admits the **trivial clustering**, namely the one with only one cluster containing all the agents. If a configuration admits a nontrivial clustering, then its size is bounded below. Conversely we show any configuration with sufficiently large radius will admit a nontrivial clustering.

**Lemma 3.1** (Nontrivial clustering). Given a pair of positive parameters $(l, \epsilon)$, there exists a number $r(l, \epsilon) > 0$ such that if the radius of a configuration is greater than $r(l, \epsilon)$, then the configuration admits a nontrivial clustering $\sigma(l, \epsilon)$.

**Proof.** Suppose not, then given any $r > 0$, there exists a configuration $p$ with its radius greater than $r$, but only admits the trivial clustering. For such a configuration, there must exist at least a pair of adjacent agents $\vec{x}_i$ and $\vec{x}_j$ such that $d_{ij} \leq l$, otherwise the agent-wise clustering will be admitted by $p$ and is nontrivial.

Divide $N$ agents in $p$ into disjoint groups by the following rule: two agents $\vec{x}_i$ and $\vec{x}_j$ are in the same group if and only if there is a chain $\vec{x}_{a_1}, \cdots, \vec{x}_{a_m}$ with $a_1 = i$, $a_m = j$ such that $\vec{x}_{a_q}$ and $\vec{x}_{a_{q+1}}$ are adjacent and $d_{a_qa_{q+1}} \leq l$, this holds for all $q = 1, \cdots, m - 1$. This rule uniquely determines the partition.

Let $G_1, \cdots, G_k$ be the groups associated with the partition. The radius of each group is less than $\frac{1}{2} N l$, so we may assume that there is more than one group because the radius of $p$ can be arbitrarily large. Then by the same reason, there exist at least a pair of adjacent groups $G_i$ and $G_j$ such that the distance between their centers is less than $N l / \epsilon$.

Integrate $G_1, \cdots, G_k$ by the following rule: two groups $G_i$ and $G_j$ will be integrated if and only if there is a chain $G_{a_1}, \cdots, G_{a_{m'}}$ with $a_1 = i$, $a_{m'} = j$ such that $G_{a_q}$ and $G_{a_{q+1}}$ are adjacent, and the distance between centers of $G_{a_q}$ and $G_{a_{q+1}}$ is less than $N l / \epsilon$, this holds for all $q = 1, \cdots, m' - 1$.

Let $G'_1, \cdots, G'_{k'}$, be groups of agents after integration. The radius of each $G'_k$ is less than $\frac{1}{2\epsilon} N^2 l$. The number of groups will strictly decrease after each step.
of integration, so \( k' < k \). If \( k' > 1 \), then we can repeat the process to integrate groups into even larger ones by applying the same rule. In finite steps, there remains only one group that contains all the agents, and its radius is bounded above by \( \frac{1}{d_N} N^N l \). This is a contradiction because we can choose radius of \( p \) arbitrarily large.

A configuration may admit multiple clusterings with respect to the same pair of parameters. We then ask whether there is a canonical one among all qualified clustering. A clustering \( \sigma(l, \epsilon) \) induces a partition on the set of vertices \( V \). Let \( (V_1, \cdots, V_k) \) be the family of disjoint, nonempty subsets with respect to \( \sigma(l, \epsilon) \), then each \( V_i \) collects indices of agents in cluster \( C_1 \). Then there is a partial order on clusterings describing the granularity of the partition. Suppose \( \sigma'(l, \epsilon) \) is another clustering and induces a different partition \( (V'_1, \cdots, V'_k) \) on \( V \). We denote by \( \sigma(l, \epsilon) \succ \sigma'(l, \epsilon) \) if \( k > k' \) and each \( V_i \) is a subset of \( V'_j \) for some \( j \).

**Lemma 3.2** (Linear order on clustering with fixed parameters). Given a pair of parameters \((l, \epsilon)\) with \( \epsilon < 1/4 \), any two different clustering \( \sigma(l, \epsilon) \) and \( \sigma'(l, \epsilon) \) is comparable and hence, all clusterings with fixed parameter \((l, \epsilon)\) form a linearly ordered set, i.e., \( \sigma_1(l, \epsilon) \succ \cdots \succ \sigma_n(l, \epsilon) \) with \( \sigma_n(l, \epsilon) \) the trivial clustering.

**Proof.** Suppose there are two non-comparable clusterings \( \sigma_1(l, \epsilon) = (C_1, \cdots, C_k) \), \( \sigma'(l, \epsilon) = (C'_1, \cdots, C'_{k'}) \) on \( p \). With out loss of generality, we assume

1. \( \vec{x}_1 \in C_1 \) and \( \vec{x}_2 \in C_2 \)
2. \( \vec{x}_1, \vec{x}_2 \in C'_1 \)
3. \( \vec{x}_3 \in C_1 \cap C'_2 \) and \( \vec{x}_3 \) is adjacent to \( \vec{x}_1 \)

Let \( r_1, r_2 \) and \( r'_1 \) be the radii of clusters \( C_1, C_2 \) and \( C'_1 \), respectively. Let \( r := \max\{r_1, r_2\} \), then \( r > 0 \) because \( C_1 \) contains at least two agents \( \vec{x}_1 \) and \( \vec{x}_3 \). Notice that \( d_{12} > (1/\epsilon - 2)r \) since \( \vec{x}_1 \) and \( \vec{x}_2 \) are in different clusters with respect to \( \sigma(l, \epsilon) \), while \( d_{12} < 2r'_1 \) as the two agents are both contained in \( C'_1 \). On the other hand, \( d_{13} > (1/\epsilon - 2)r'_1 \) because \( \vec{x}_1 \) and \( \vec{x}_3 \) are in different clusters with respect to \( \sigma'(l, \epsilon) \). So then \( \frac{1}{2}(\frac{1}{\epsilon} - 2)^2r \leq d_{13} \leq 2r \) which is a contradiction if \( \epsilon < 1/4 \).

**Remark.** In the rest of this section, we will implicitly assume \( \epsilon < 1/4 \). The clustering \( \sigma_1(l, \epsilon) \) is then indecomposable, and can be regarded as the canonical clustering with respect to \((l, \epsilon)\).

Let \( \sigma(l, \epsilon) \) and \( \sigma'(l', \epsilon') \) be two clusterings on \( p \) with respect to parameters \((l, \epsilon)\) and \((l', \epsilon')\) respectively, if the two clusterings induce the same partition on \( V \), we will then simply write \( \sigma(l, \epsilon) \simeq \sigma'(l', \epsilon') \).

**Lemma 3.3** (Variation on chains). Let \((l, \epsilon)\) and \((l', \epsilon')\) be two pairs of parameters with \( l' \geq l \) and \( \epsilon' \leq \epsilon \). Let \( \{\sigma_i(l, \epsilon)\}_{i=1}^n \) and \( \{\sigma'_j(l', \epsilon')\}_{j=1}^{n'} \) be the linearly ordered sets of clusterings with respect to \((l, \epsilon)\) and \((l', \epsilon')\) respectively, then for each \( k = 1, \cdots, n' \), we have \( \sigma'_k(l', \epsilon') \simeq \sigma_j(l, \epsilon) \) for some \( j = 1, \cdots, n \).
Proof. Suppose \((C_1, \cdots, C_k)\) is a partition with respect to \(\sigma'(l', \epsilon')\), so then each subgraph \(\Gamma_i\) associated with \(C_i\) is connected, and by assumption \(t_{ij} > l' \geq l\) and \(r_i/l_{ij} < \epsilon' \leq \epsilon\) if \(C_i\) and \(C_j\) are adjacent. So then the clustering \(\sigma(l, \epsilon) := (C_1, \cdots, C_k)\) satisfies all three defining conditions.

Clustering is useful for investigating gradient flows of dilute configurations, and we now state the main theorem of this section.

**Theorem 3.4** (Clustering on sequence of diverging configurations). If there were an initial condition \(p\) such that along the gradient flow \(\varphi_{\geq 0}(p)\) the radius of the configuration can’t be bounded above, then there would be a monotone sequence of times \(\{t_i\}_{i \in \mathbb{N}}\) approaching to infinity such that

1. each configuration in the sequence \(\{\varphi_{t_i}(p)\}_{i \in \mathbb{N}}\) admits a nontrivial clustering \(\sigma(l, \epsilon)\) that will induce the same partition on \(V\).

2. if we partition the agents in \(\varphi_{t_i}(p)\) into disjoint clusters with respect to \(\sigma(l, \epsilon)\), then the minimum adjacent-cluster distance approaches to infinity along the sequence.

3. there is a fixed number \(R > 0\) that bounds any radius of a cluster in a configuration in the sequence.

Proof. As the radius the of configuration along the flow curve isn’t bounded above, so there is a time sequence \(\{t_i\}_{i \in \mathbb{N}}\) approaching to infinity such that the radius of the configuration \(\varphi_{t_i}(p)\) monotonically increases and approaches to infinity along the sequence. For convenience, we let \(p_i\) denote \(\varphi_{t_i}(p)\).

Let \((l, \epsilon)\) be a pair of positive parameters, since the radius of the configuration diverges along the sequence, by lemma 3.1 there is an index \(n_{l, \epsilon}\) associated with \((l, \epsilon)\) such that for each configuration \(p_i\), \(i \geq n_{l, \epsilon}\), in the sequence, the canonical clustering \(\sigma(l, \epsilon)\) on \(p_i\) is nontrivial. Fix \(\epsilon\) and let \(\{t_i\}_{i \in \mathbb{N}}\) be a positive monotonically increasing sequence approaching to infinity, we then get the sequence of indices \(\{n_{t_i, \epsilon}\}_{i \in \mathbb{N}}\) correspondingly.

Pick a subsequence \(\{p_{j_i}\}_{i \in \mathbb{N}}\) out of \(\{p_i\}_{i \in \mathbb{N}}\), and the index \(j_i, i \in \mathbb{N},\) satisfies

\[
\begin{aligned}
    j_1 &> n_{t_1, \epsilon} \\
    j_i &> \max\{j_{i-1}, n_{t_i, \epsilon}\}, \quad i > 1
\end{aligned}
\]  

so then, the minimum adjacent-cluster distance in \(p_{j_i}\), associated with \(\sigma(l_{j_i}, \epsilon)\) will approach to infinity along the subsequence. For convenience, we may assume that the original sequence already satisfies this condition, and we will write \(\{p_i, \sigma(l_i, \epsilon)\}_{i \in \mathbb{N}}\) to emphasize the clustering on each \(p_i\).

There are only finitely many partitions on \(V\), so there is a subsequence \(\{p_{j_i}, \sigma(l_{j_i}, \epsilon)\}_{i \in \mathbb{N}}\) chosen in a way that each clustering \(\sigma(l_{j_i}, \epsilon)\) on \(p_{j_i}\) induces the same partition on \(V\). Again we assume that the original sequence has already satisfied this condition. So far we have constructed a sequence \(\{p_i\}_{i \in \mathbb{N}}\) that satisfies the first two conditions in the statement of the theorem.

Last we show that there is a subsequence out of \(\{p_i\}_{i \in \mathbb{N}}\) that will satisfy condition 3. The proof is done by induction on the number of agents.
**Base case:** The case $N = 2$ is trivially true.

**Induction step:** We assume that the theorem holds for $N < k$ and prove for the case $N = k$. Along the sequence $\{p_i, \sigma(l_i, \epsilon)\}_{i \in \mathbb{N}}$, we may assume that there is at least one cluster of agents, say $C_1$, with its radius approaching to infinity. For convenience, we denote by $\{\hat{p}_i\}_{i \in \mathbb{N}}$ the sequence of sub-configurations formed by agents in $C_1$. The number of agents in each $\hat{p}_i$ is fixed and less than $k$, so by induction there will be a subsequence $\{\hat{p}_{j_i}, \hat{\sigma}_{j_i}(\hat{l}_{j_i}, \epsilon)\}_{i \in \mathbb{N}}$, with each $\hat{\sigma}_{j_i}(\hat{l}_{j_i}, \epsilon)$ a clustering on $\hat{p}_{j_i}$, such that it satisfies all three conditions in the statement of the theorem. Since cluster $C_1$ is indecomposable with respect to $\sigma(l_{j_i}, \epsilon)$, it then implies that $\hat{l}_{j_i} < l_{j_i}$. So by lemma 3.3 and the illustration of figure 2, there is

![Figure 2](image)

Figure 2: In this figure, the clusters $C_1$, $C_2$ and $C_3$ are with respect to the canonical clustering $\sigma(l, \epsilon)$ on the whole configuration, while the clusters $C_{11}$, $C_{12}$ and $C_{13}$ are with respect to the canonical clustering $\hat{\sigma}(\hat{l}, \epsilon)$ on the sub-configuration formed by agents in $C_1$. There is a clustering $\sigma(\hat{l}, \epsilon)$ on $p$ that agrees with $\hat{\sigma}(\hat{l}, \epsilon)$ inside $C_1$, but agrees with $\sigma(l, \epsilon)$ outside. It suffices to show that the adjacent-cluster distance between $C_i, i \neq 1$, and $C_{1j}$ is greater than $\hat{l}$ if $C_i$ and $C_{1j}$ are adjacent. This holds because by assumption $C_i$ and $C_1$ are adjacent, so then $d(C_i, C_{1j}) > (l - r_1) > (1/\epsilon - 1)r_1 > (1/\epsilon - 1)\hat{l}/2 > 3/2\hat{l}$.

a clustering $\sigma(\hat{l}_{j_i}, \epsilon)$ on $p_{j_i}$, such that it agrees with $\hat{\sigma}(\hat{l}_{j_i}, \epsilon)$ inside $C_1$, but with $\sigma(l_{j_i}, \epsilon)$ outside. Pick the subsequence $\{p_{j_i}, \sigma(\hat{l}_{j_i}, \epsilon)\}_{i \in \mathbb{N}}$ out of $\{p_i, \sigma(l_i, \epsilon)\}_{i \in \mathbb{N}}$ with $\sigma(l_{j_i}, \epsilon)$ replaced with $\sigma(\hat{l}_{j_i}, \epsilon)$. If there is any other diverging cluster in the subsequence, then we repeat the process. The whole procedure terminates in finite steps of repetition, and we get a sequence that satisfies all three conditions in the statement of the theorem. \(\square\)
4 Dissipation zone

Let $d > 0$ be a distance, let $r$ be a radius and let $ij$ be an edge in $E$. We then define two subsets of the configuration space by

\begin{align*}
X_{ij}(d) &:= \{ p \in P | d_{ij} = d \} \\
Y(r) &:= \{ p \in P | r_p = r \}
\end{align*}

where $r_p$ is the radius of a configuration $p$. A distance $d$ (or a radius $r$) is said to be absent in a configuration $p$ if $d_{ij} \neq d$ for any $ij \in E$ (or $r(p) \neq r$), and $X_{ij}(d)$ (or $Y(r)$) is said to be a dissipation zone if the distance $d$ (or the radius $r$) is absent in any equilibrium associated with the gradient flow.

Our goal in this section is to develop theorem 4.1 and theorem 4.2; these two theorems establish positive lower bounds for the loss of potential of a gradient flow once it reaches a dissipation zone.

**Theorem 4.1.** Let $d$ be a distance absent in any equilibrium, and let

\[ \mu_{ij}(N,d) := \inf \{ |f(p)| | p \in X_{ij}(d) \} \]

then $\mu_{ij}(N,d) > 0$. If along a gradient flow $\varphi_{\geq 0}(p)$, there is a moment $t > 0$ at which $\varphi_t(p) \in X_{ij}(d)$, then there is a $\tau_{\mu} > 0$ such that during period $[t, t + \tau_{\mu}]$, the loss of potential along the gradient flow is at least $\mu_{ij}^2(N,d) \tau_{\mu}/4$.

There is a similar version of theorem 4.1 by replacing $X_{ij}(d)$ with $Y(r)$, as we state below.

**Theorem 4.2.** Let $r$ be a radius absent in any equilibrium, let

\[ \nu(N,r) := \inf \{ |f(p)| | r_p = r \} \]

then $\nu(N,r) > 0$. If along a gradient flow $\varphi_{\geq 0}(p)$, there is a moment $t > 0$ at which $\varphi_t(p) \in Y(r)$, then there is a $\tau_{\nu} > 0$ such that during period $[t, t + \tau_{\nu}]$, the loss of potential along the gradient flow is at least $\nu^2(N,d) \tau_{\nu}/4$.

In this section, we will mainly focus on the proof of theorem 4.1 but all the arguments along the development can be used to prove theorem 4.2. A complete proof will be given at the end of this section.

Let $I$ be a closed neighborhood of $d$, we then let $X_{ij}(I) := \{ X_{ij}(d) | d \in I \}$. Notice that

\[ \frac{d}{dt} \Psi(\varphi_t(p)) = -|f(\varphi_t(p))|^2 \]

So theorem 4.1 can be established if we can

1. find a closed neighborhood $I$ of $d$ such that $\inf \{ |f(p)| | p \in X_{ij}(I) \} > \frac{1}{2} \mu_{ij}(N,d)$.

2. compute a lower bound for the period that a gradient flow takes to escape out of $X_{ij}(I)$ from $X_{ij}(d)$.

The rest of this section is then organized by this order, we will first prove the existence the interval $I$, and then estimate the escaping time.
4.1 A lower bound for escaping velocity

Our goal here is to establish theorem 4.3.

**Theorem 4.3.** Let $d > 0$ be a distance absent in any equilibrium, and let $\mu_{ij}(N,d)$ be defined by equation (36), then $\mu_{ij}(N,d) > 0$. Moreover, as a function of $x$, $\mu_{ij}(N,x)$ is continuous at $d$. This, in particular, implies that there exists a closed neighborhood $I$ of $d$ with $\mu_{ij}(N,d') > \frac{1}{2}\mu_{ij}(N,d)$ for any $d' \in I$.

The proof of theorem 4.3 will be given after lemma 4.4. We start by proving the existence of a positive lower bound for $|f(p)|$ as $p$ varies over $X_{ij}(d)$.

**Lemma 4.4.** Suppose $d > 0$ is absent in any equilibrium, then $\mu_{ij}(N,d) > 0$.

**Proof.** As the edge $ij$ is fixed during the proof, so for simplicity, we will simply write $X(\cdot)$ and $\mu(\cdot, \cdot)$ by omitting their subindices. Let $a$, $b$ be two positive numbers with $0 \leq a \leq b \leq \infty$, and let $X^b_a(d) := X(d) \cap P^b_a$. Similarly, we will write $X^b$ if $a = 0$ and write $X_a$ if $b = +\infty$. Define

$$
\mu_1(N,d) := \inf \{|f(p)|| p \in X^b_a(d)\} \\
\mu_2(N,d) := \inf \{|f(p)|| p \in X(d) - X_a(d)\} \\
\mu_3(N,d) := \inf \{|f(p)|| p \in X(d) - X^b(d)\}
$$

then

$$
\mu(N,d) = \min\{\mu_1(N,d), \mu_2(N,d), \mu_3(N,d)\}
$$

so it suffices to find $a$ and $b$ such that $\mu_i(N,d) > 0$ for each $i = 1, 2, 3$.

Proof that $\mu_1(N,d) > 0$. This holds for any two positive numbers $a$ and $b$. Because the subset $X^b_a(d)$ is compact, and by assumption the gradient field doesn’t vanish over $X^b_a(d)$.

Proof that $\mu_2(N,d) > 0$. Let $\delta$ be a positive number, we show that there exists $a > 0$ such that $|f(p)| > \delta$ for any $p \in X(d) - X_a(d)$. Choose $a$ sufficiently small, we then apply the same arguments in the proof of theorem 2.2 to argue that if there are two adjacent agents in $p$ such that their distance is less than $a$, then to keep $|f(p)|$ less than $\delta$, all the other agents can be successively located in a small open ball. In particular, the radius of the ball approaches to zero as $a$ goes to zero. So then by lemma 2.4 $|f(p)|$ can be made arbitrary large by shrinking $a$.

Proof that $\mu_3(N,d) > 0$. The proof is done by induction on the number of agents.

Base case: In the case $N = 2$, the set $X(d)$ is a singleton and

$$
\mu(2,d) = \sqrt{2}|\hat{g}_{12}(d)| > 0
$$

Induction step: Assume that $\mu(N,d) > 0$ for $N \leq k - 1$ and we prove for the case $\mu_3(k,d) > 0$. Let

$$
\tilde{\mu}(k-1,d) := \min\{\mu(m,d)|2 \leq m \leq k - 1\}
$$
Let \((l, \epsilon)\) be a pair of positive parameters for clustering. By lemma \[3.1\] and lemma \[3.2\] the canonical clustering on \(p\) is nontrivial if the radius of \(p \in X(d)\) exceeds certain threshold. Let \(b\) be twice the threshold, we show that there is a positive lower bound for \(|f(p)|\) as \(p\) varies over \(X(d) - X^b(d)\).

Without loss of generality, we assume \(\vec{x}_1\) and \(\vec{x}_2\) are adjacent, and \(d_{12} = d\). As the distance \(d\) is fixed, if we choose \(l\) large enough with \(\epsilon\) fixed, then \(\vec{x}_1\) and \(\vec{x}_2\) are belong to the same cluster, say \(C_1\). Let \(\bar{p}\) be the sub-configuration formed by agents in \(C_1\). Since the clustering is nontrivial, the number of agents in \(C_1\) is less than \(k\). Then by induction, \(|\bar{f}(\bar{p})| \geq \mu(k - 1, d) > 0\). So there is at least one agent \(\vec{x}_i\) in \(\bar{p}\) such that \(|\bar{f}(\bar{p})| \geq \frac{1}{\sqrt{k-1}} \mu(k - 1, d)\).

Now take into account interactions between agents in adjacent clusters. By the condition of fading attraction, there exists \(d' > 0\) such that for any \(ij \in E\) and any \(d > d'\), we have

\[
|\bar{g}_{ij}(d)| < \frac{1}{2k^{\frac{1}{2}}} \mu(k - 1, d) \tag{43}
\]

Increase \(l\), if necessary, so that the resulting distance between agents in any two adjacent clusters is greater than \(d'\). So then \(|f_i(p)| > \frac{1}{2\sqrt{k}} \mu(k - 1, d)\) and hence,

\[
f(p) > \frac{1}{2\sqrt{k}} \mu(k - 1, d) > 0 \tag{44}
\]

for any \(p \in X(d) - X^b(d)\). Consequently \(\mu_3(k, d) \geq \frac{1}{2\sqrt{k}} \mu(k - 1, d) > 0\). \(\square\)

We now prove theorem \[4.3\]

**proof of theorem \[4.3\]**. If \(d\) is absent in any equilibrium, then there is a closed neighborhood \(I\) of \(d\) such that any distance in \(I\) is absent in any equilibrium. In the proof of lemma \[4.4\] we have showed that there is a positive number \(a\) such that if we let

\[
\mu'_{ij}(N, d) := \inf\{|f(p)| | p \in X_{ij}(d) \cap P_a\} \tag{45}
\]

then \(\mu'_{ij}(N, d) = \mu_{ij}(N, d)\), and this holds for all \(d \in I\). So it suffices to prove that \(\mu'_{ij}(N, d)\) is continuous at \(d\).

Given \(\epsilon > 0\), there exists \(l > 0\) such that if \(|d - d'| < l\), then for each \(p\) in \(X_{ij}(d)\), there is \(p'\) in \(X_{ij}(d')\) with \(|p - p'| < \epsilon|\). On the other hand, given \(\delta > 0\), there is \(\epsilon > 0\) such that if both \(p\) and \(p'\) are in \(P_a\), and if \(|p - p'| < \epsilon\), then \(|f(p) - f(p')| < \delta\). This is because each function \(\bar{g}_{ij}\) is bounded when restricted on the interval \([a, \infty)\), so given a positive number \(\delta\), there is \(\epsilon' > 0\) such that \(|\bar{g}_{ij}(d' + \epsilon'') - \bar{g}_{ij}(d'')| < \delta\) for any \(d' \geq a\), any \(\epsilon'' \in (0, \epsilon')\) and any \(ij \in E\).

Pick \(d'\) in \((d - l, d + l)\), and without loss of generality, we assume \(\mu'(N, d') \geq \mu'(N, d)\). Let \(\{p_k\}_{k \in \mathbb{N}}\) be a sequence of configurations in \(X_{ij}(d) \cap P_a\) with \(\lim_{k \to \infty} |f(p_k)| = \mu'(N, d)\). For each \(p_k\), we pick a configuration \(p'_k \in X_{ij}(d') \cap P_a\) with \(|p'_k - p_k| < \epsilon\). So then \(|f(p'_k) - f(p_k)| < \delta\) and hence,

\[
|\inf_{k \in \mathbb{N}} |f(p'_k)| - \mu'_{ij}(d, N)| < \delta \tag{46}
\]
On the other hand,

\[ \inf_{k \in \mathbb{N}} |f(p'_k)| \geq \mu'_{ij}(N, d') \geq \mu'_{ij}(N, d) \]  

which implies that \(|\mu'_{ij}(N, d') - \mu'_{ij}(N, d)| < \delta\). Since this holds for each \(d'\) in \((d - l, d + l)\), we then prove the continuity of \(\mu'_{ij}(N, d)\) at \(d\). \(\square\)

There is a similar version of theorem 4.3 for \(Y(r)\), and we state it below without a proof.

**Theorem 4.5.** Let \(r > 0\) be a radius absent in any equilibrium, then \(\nu(N, r) > 0\) and there exists a closed neighborhood \(J\) of \(r\) such that \(\nu(N, r') > \frac{1}{2}\nu(N, r)\) for any \(r' \in J\).

In the rest of this paper, we assume that the closed intervals \(I\) and \(J\) are chosen in a way so that they are symmetric around \(d\) and \(r\), respectively.

### 4.2 A lower bound for escaping time

Our goal here is to establish theorem 4.6 where we set up a lower bound for the period that a gradient flow takes to escape out of \(X_{ij}(I)\) from \(X_{ij}(d)\). In particular, we do this by establishing a lower bound for the width of \(X_{ij}(I)\) and an upper bound for the escaping velocity. These two bounds are established in lemma 4.7 and lemma 4.9, respectively.

**Theorem 4.6.** There is a positive number \(v_p\) associate with each gradient flow \(\varphi_{\geq 0}(p)\) such that \(v_p \geq |f(\varphi_t(p))|\) for all \(t \geq 0\). Let \(I = [d - l_{ij}, d + l_{ij}]\) be a closed interval in \(\mathbb{R}^+\). If there is a moment \(t > 0\) at which \(d_{ij} = d\), then it takes at least \(\tau_p = l_{ij}/\sqrt{2}v_p\) units of time for a gradient flow to escape out of \(X_{ij}(I)\) from \(X_{ij}(d)\)

We only consider the case \(p\) is not an equilibrium because otherwise, the gradient flow \(\varphi_{\geq 0}(p)\) stays at \(p\) for all time. The proof of theorem 4.6 will be given after lemma 4.7 and lemma 4.9.

**Lemma 4.7.** Let \(d\) and \(d'\) be two positive numbers with \(d > d'\), let \(D(X_{ij}(d), X_{ij}(d'))\) be the distance between \(X_{ij}(d)\) and \(X_{ij}(d')\), i.e.

\[ D(X_{ij}(d), X_{ij}(d')) := \inf\{|p - p'| | p \in X_{ij}(d), p' \in X_{ij}(d')\} \]  

then \(D(X_{ij}(d), X_{ij}(d')) = (d - d')/\sqrt{2}\).

**Proof.** It is clear that \(D(X_{ij}(d), X_{ij}(d')) \geq (d - d')/\sqrt{2}\) because

\[ D(X_{ij}(d), X_{ij}(d')) \geq \sqrt{\|x_i - \bar{x}_i\|^2 + \|x_j - \bar{x}_j\|^2} \geq (d - d')/\sqrt{2} \]  

On the other hand, for almost all configurations \(p\) in \(X_{ij}(d)\), there will be a configuration \(p'\) in \(X_{ij}(d')\) with \(|p - p'| = (d - d')/\sqrt{2}\). This is simply achieved by fixing all the agents in \(p\) but moving \(\bar{x}_i\) and \(\bar{x}_j\) to get \((d - d')\) closer to
each other along the line determined by themselves, then we will get a $p'$ in $X_{ij}(d')$ with $|p - p'| = (d - d')/\sqrt{2}$ as long as the new positions of $\vec{x}_i$ and $\vec{x}_j$ are not occupied by other agents. But this holds for almost all configurations in $X_{ij}(d)$.

There is a similar version for $Y(r)$.

Lemma 4.8. Let $r$, $r'$ be two radii with $r > r' > 0$, and let $D(Y(r), Y(r'))$ be the distance between $Y(r)$ and $Y(r')$, then $D(Y(r), Y(r')) > r - r'$.

Proof. The centroid of a configuration $p$ is at the origin. So if $p = (\vec{x}_1, \cdots, \vec{x}_N)$ is in $Y(r)$ while $p' = (\vec{x}'_1, \cdots, \vec{x}'_N)$ is in $Y(r')$, then there are at least two pairs of agents, say $(\vec{x}_1, \vec{x}'_1)$ and $(\vec{x}_2, \vec{x}'_2)$, such that $\vec{x}_i \neq \vec{x}'_i$, $i = 1, 2$. We may also assume that $\vec{x}_1$ and $\vec{x}'_1$ are the outermost agents in $p$ and $p'$, respectively. So then $|p - p'| \geq |\vec{x}_1 - \vec{x}'_1| + |\vec{x}_2 - \vec{x}'_2| > |\vec{x}_1 - \vec{x}'_1| \geq r - r'$.

Let $I = [d - l_\mu, d + l_\mu]$ be a closed neighborhood of $d$, then by lemma 4.7 we have $D(X_{ij}(d), X_{ij}(d \pm l_\mu)) = l_\mu/\sqrt{2}$. This then establishes a lower bound for the distance that a gradient flow has to travel to escape out of $X_{ij}(I)$ from $X_{ij}(d)$. We now show that for each gradient flow $\varphi_{\geq 0}(p)$, the velocity $|f(\varphi_t(p))|$ has a positive upper bound.

Lemma 4.9. Consider the RMA system described by equation (1). For each initial condition $p$, we let

$$ v_p := \sup \{ |f(\varphi_t(p))| | t \geq 0 \} \quad (50) $$

then $v_p$ exists.

Proof. By condition of fading attraction, the function $\tilde{g}_{ij}(d)$ is bounded above. On the other hand, by theorem 2.1 there exists $\alpha > 0$ such that $\varphi_{\geq 0}(p) \subset P_\alpha$. So the magnitude of the interaction between any two agents along the gradient flow is bounded above, and so is $|f(\varphi_t(p))|$ for any $t \geq 0$.

Theorem 4.6 is then proved by combining lemma 4.7 and lemma 4.9.

4.3 Proofs of theorem 4.1 and theorem 4.2

Proof of theorem 4.1. By theorem 4.3 there is a closed interval $I$ of $d$ such that $f(p) \geq \frac{1}{2} \mu_{ij}(N, d)$ for any $d \in I$. We may assume that $I = [d - l_\mu, d + l_\mu]$. By theorem 4.6 it takes at least $\tau_\mu = l_\mu/\sqrt{2} v_p$ units of time for the gradient flow to leave the dissipation zone $X_{ij}(I)$ from $X_{ij}(d)$, so during the period $[t, t + \tau_\mu]$, the loss of potential is given by

$$ \Psi_t(p) - \Psi_{t+\tau_\mu}(p) = \int_t^{t+\tau_\mu} |f(\varphi_s(p))|^2 ds \geq \frac{1}{4} \mu_{ij}(N, d)^2 \tau_\mu \quad (51) $$

This then completes the proof.
The arguments in the proof of theorem 4.1 can be directly used to prove theorem 4.2; we omit the details here.

We end this section with a little discussion on the dissipation zone. For each $k$-agent system, by theorem 2.2 there is a distance $d_k$ and a radius $r_k$ such that if $d \geq d_k$ (or $r \geq r_k$), then $d$ as a distance (or $r$ as a radius) is absent in any equilibrium. Let $D$ and $R$ be two positive numbers that satisfy

$$D \geq \max\{d_k | 2 \leq k \leq N\} \tag{52}$$

$$R \geq \max\{r_k | 2 \leq k \leq N\} \tag{53}$$

Here we abuse the notation $R$ as it is already defined in the statement of theorem 3.4, yet we may assume that $R$ is sufficiently large so that it can be applied for both cases.

Let $d$ and $r$ be chosen from $[D, \infty)$ and $[R, \infty)$ respectively. For each $k$-agent system, there correspond $\hat{\mu}_{ij}(d)$ and $\hat{\nu}(r)$, and we here define

$$\hat{\mu}_{ij}(d) := \min\{\mu_{ij}(k, d) | 2 \leq k \leq N\} \tag{54}$$

$$\hat{\nu}(r) := \min\{\nu(k, r) | 2 \leq k \leq N\} \tag{55}$$

The definitions of $D$, $R$, $\hat{\mu}_{ij}(d)$ and $\hat{\nu}(r)$ are useful for studying semi-diverging gradient flows as we will define in the next section.

5 Asymptotic path-behavior of a semi-diverging gradient flow

In this section, we will first define the notion of semi-diverging gradient flow, and investigate certain properties associated with it. Let $(l, \epsilon)$ be a fixed pair of positive parameters for clustering, and let $R$ be defined by equation (53). Fix a nontrivial clustering $\sigma(l, \epsilon)$, we say a gradient flow $\varphi \geq 0(p)$ is semi-diverging with respect to $\sigma(l, \epsilon)$ if for any $t > 0$, there is a moment $t_0 > t$ such that the configuration $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$, and the radius of each cluster in $\varphi_{t_0}(p)$ is bounded above by $R$. Our goal in this section is to develop theorem 5.1.

**Theorem 5.1.** Let $\sigma(l, \epsilon)$ be a nontrivial clustering, and let $\varphi_{\geq 0}(p)$ be a semi-diverging gradient flow with respect to $\sigma(l, \epsilon)$. There is a threshold $L$ for $l$ and a threshold $T$ for time such that if $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$ with $t_0 \geq T$ and $l \geq L$, then there is a moment $t \geq t_0$ such that $\varphi_{t_0}(p)$ will reach a dissipation zone $X_{ij}(d)$ with $d \in [L - 2R, L + 2R]$ for some $ij \in E$.

The theorem will be established after a sequence of lemmas and theorems, and the proof will be given at the end of this section.

5.1 Metric property of clustering along a semi-diverging gradient flow

Our goal here is establish theorem 5.2. This theorem describes a metric property of a semi-diverging gradient flow, in particular, it relates radii of clusters.
to adjacent-clusters distances in configurations along a semi-diverging gradient flow.

**Theorem 5.2.** Let $\sigma(l, \epsilon)$ be a nontrivial clustering, and let $\varphi_{\geq 0}(p)$ be a semi-diverging gradient flow with respect to $\sigma(l, \epsilon)$. There is a threshold $L_1$ for $l$ and a threshold $T$ for time such that if $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$ with $l \geq L_1$ and $t_0 \geq T$, then each radius of cluster will remain less than $R$ along the gradient flow $\varphi_{t_0}(p)$ as long as each adjacent-cluster distance keeps greater than $L_1$.

The proof of theorem 5.2 will be given after lemma 5.3.

**Lemma 5.3.** Let $\sigma(l, \epsilon)$ be a nontrivial clustering, and let $\varphi_{\geq 0}(p)$ be a semi-diverging gradient flow with respect to $\sigma(l, \epsilon)$. Suppose $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$, and suppose the maximum radius of cluster in $\varphi_{t_0}(p)$ is $R$. Let $\hat{\nu}(R)$ be defined by equation (55), then there is a threshold $L_1$ for $l$ and a fixed period $\tau_{\nu} > 0$ such that if $l \geq L_1$, then $|f(\varphi_t(p))| > \hat{\nu}(R)/4$ for any $t \in [t_0, t_0 + \tau_{\nu}]$.

Proof. We assume that the radius of $C_1$ in $\varphi_{t_0}(p)$ is $R$, and we denote by $\varphi_{t_0}(\hat{p})$ the sub-configuration formed by agents in $C_1$, so then $|f(\varphi_{t_0}(\hat{p}))| > \hat{\nu}(R)$. Let $J := [R - \delta, R + \delta]$ and we assume $\hat{\nu}(r) > \frac{1}{2}\hat{\nu}(R)$ for any $r \in J$.

Recall in lemma 4.8, we have proved that if $p_1 \in Y(R)$ and $p_2 \in Y(R(\pm \delta)$, then $|p_1 - p_2| > \delta$, in the lemma we implicitly assumed that both $p_1$ and $p_2$ are centered at the origin. We now consider one of its variations. Suppose $p_1$ is of radius $R$ centered at $\bar{c}_1$ while $p_2$ is of radius $R(\pm \delta$ centered at $\delta_2$, then

$$|p_1 - p_2| > |\bar{c}_1 - \bar{c}_2| + \delta > \delta$$

(56)

Let $v_p$ be defined by equation (50), then $v_p > 0$, and we let $\tau_{\nu} := \delta/v_p$. Inequality (56) then implies that the radius of $C_1$ will be within interval $J$ during the period $[t_0, t_0 + \tau_{\nu}]$.

By condition of fading attraction, there is a distance $\tilde{d} > 0$ such that for any $d > \tilde{d}$ and any $ij \in E$, we have

$$|\bar{g}_{ij}(d)| < \frac{1}{4N^{3/2}}\hat{\nu}(R)$$

(57)

Let $L_1 := \tilde{d} + 2v_p\tau_{\nu} + 2R$. Then if each adjacent-cluster distance in $\varphi_{t_0}(p)$ is greater than $L_1$, then the distance between any two agents of adjacent clusters will be greater than $\tilde{d} + 2v_p\tau_{\nu}$ at time $t_0$. Since $|f_i(\varphi_t(p))| < |f(\varphi_t(p))| \leq v_p$ for all $i = 1, \cdots, N$ and all $t > 0$, the distance between any two agents of adjacent clusters will be greater than $\tilde{d}$ during the period $[t_0, t_0 + \tau_{\nu}]$.

Let $f(\varphi_{t}(\hat{p}))$ be the restriction of $f(\varphi_{t}(p))$ to $\varphi_{t}(\hat{p})$, it is a sum of two parts: the intra-cluster part $f_{A}(\varphi_{t}(\hat{p}))$ and the inter-cluster part $f_{B}(\varphi_{t}(\hat{p}))$, i.e.,

$$f(\varphi_{t}(\hat{p})) = f_{A}(\varphi_{t}(\hat{p})) + f_{B}(\varphi_{t}(\hat{p}))$$

(58)

The intra-cluster part $f_{A}(\varphi_{t}(\hat{p}))$ is contributed by agents inside $C_1$, so by previous notation, it is just $f(\varphi_{t}(\hat{p}))$. We have chosen $\tau_{\nu}$ so that if $t \in [t_0, t_0 + \tau_{\nu}]$, then $|f_{A}(\varphi_{t}(\hat{p}))| > \hat{\nu}(R)/2$. The inter-cluster part $f_{B}(\varphi_{t}(\hat{p}))$ is contributed by
agents outside $C_1$. The threshold $L_1$ is chosen so that $|f_B(\varphi_t(\hat{p}))| < \hat{\nu}(R)/4$ for any $t \in [t_0, t_0 + \tau_\nu]$. So then, during the period $[t_0, t_0 + \tau_\nu]$, we have $|f(\varphi_t(p))| > |f(\varphi_t(\hat{p}))| \geq |f_A(\varphi_t(\hat{p}))| - |f_B(\varphi_t(\hat{p}))| > \hat{\nu}(R)/4$. 

We now prove theorem 5.2.

**Proof of theorem 5.2.** The theorem is trivially true if the clustering $\sigma(l, \epsilon)$ is agent-wise, so we assume otherwise.

The potential function $\Psi(p)$ is bounded below as $p$ varies over $P$ because each integral $\int d^d g_{ij}(x) dx$ is bounded below as $d$ varies over $\mathbb{R}^+$. This in particular implies that

$$\lim_{t \to \infty} \int_t^\infty |f(\varphi_s(p))|^2 ds = 0 \quad (59)$$

Let $\hat{\nu}(R) > 0$ be defined by equation (54), and let $\tau_\nu > 0$ be the period defined in the statement of lemma 5.3 then there is a moment $T > 0$ such that for any $t \geq T$, we have

$$\int_t^\infty |f(\varphi_s(p))|^2 ds < \frac{1}{16} \hat{\nu}^2(R) \tau_\nu \quad (60)$$

We now assume $t_0 > T$, $l > L_1$, and we assume that all adjacent-cluster distances keep greater than $L_1$ along the gradient flow $\varphi_{\geq t_0}(p)$. So then, if there is a moment $t \geq t_0$ at which the maximum radius of cluster reaches $R$, then by lemma 5.3, the loss of the potential of the gradient flow over the period $[t, t + \tau_\nu]$ will exceed $\hat{\nu}^2(R) \tau_\nu/16$ which is a contradiction.

5.2 Metric property of configurations along a semi-diverging gradient flow

Our goal here is to establish theorem 5.4. This theorem describes a metric property of a semi-diverging gradient flow, in particular, it relates the size of a configuration along a semi-diverging gradient flow to both radii of clusters and adjacent-cluster distances.

**Theorem 5.4.** Let $\sigma(l, \epsilon)$ be a nontrivial clustering, and let $\varphi_{\geq 0}(p)$ be a semi-diverging gradient flow with respect to $\sigma(l, \epsilon)$. There is a threshold $L_2$ for $l$ such that if $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$ with $l \geq L_2$, and further we assume that along the flow $\varphi_{\geq t_0}(p)$,

1. each adjacent-cluster distance remains greater than $L_2$.

2. each radius of cluster remains less than $R$.

then there exists a number $b > 0$ such that $\varphi_{\geq t_0}(p) \subset P_b$.

**Proof.** We may as well assume $t_0 = 0$. Choose $L_2$ large enough so that the interaction between any two agents of adjacent clusters is an attraction. Let $\hat{c}_i(t) \in \mathbb{R}^n$ be the center of cluster $C_i$ at time $t$, and let

$$\hat{\hat{c}}_i(t) := \begin{cases} \hat{c}_i(t)/|\hat{c}_i(t)| & \hat{c}_i(t) \neq 0 \\ 0 & \hat{c}_i(t) = 0 \end{cases} \quad (61)$$
We assume there are $M := S$ clusters, and each cluster $C_i$ contains $n_i$ agents. Let $S := \{1, \ldots, M\}$, then for each $i \in S$, and each subset $S' \subset S$, we define

$$\pi(i, S', t) := \frac{\langle \hat{c}_i(t), \sum_{j \in S'} n_j \hat{c}_j(t) \rangle}{\sum_{j \in S'} n_j}$$ (62)

it is understood as the projection of the center of agents in $\bigcup_{j \in S'} C_j$ along direction $\hat{c}_i(t)$. For each $m = 1, \ldots, M$, we then define

$$\pi_m(t) := \max\{\pi(i, S', t) | i \in S, |S'| = m\}$$ (63)

each $\pi_m(t)$, as a function of $t$, is continuous and piecewise differentiable. It suffices for us to show $\pi_1(t)$ is bounded above for all $t > 0$ because by definition, $\pi_1(t)$ is the radius of $\varphi_i(p)$. The proof is then done by contradiction, i.e., we assume for any distance $d$, there is a moment $t_1 > 0$ such that $\pi_1(t_1) = d$.

We may as well assume that $d > \pi_1(0)$ and $\pi_1(t) < d$ for $t < t_1$. Suppose $C_1$ is the outermost cluster at time $t_1$, then at the same moment, there is at least one cluster, say $C_2$, adjacent to $C_1$ such that

$$\langle \hat{c}_1(t_1), \hat{c}_2(t_1) \rangle > d - 2R$$ (64)

because otherwise, all clusters that are adjacent to $C_1$ will pull back $C_1$ along direction $\hat{c}_1(t_1)$. Then by time reversing, there is a number $\delta > 0$ such that $\hat{c}_1(t) > \hat{c}_1(t_1)$ for any $t \in (t_1 - \delta, t_1)$ which contradicts to our assumption on $t_1$. Inequality (64) then implies that

$$\pi_2(t_1) \geq \pi(1, \{1, 2\}, t_1) \geq d - 2R$$ (65)

We may assume $d$ is sufficiently large so that $d - 2R > \pi_2(0)$. Then during the period $[0, t_1]$, there exists a moment $t_2$ such that $\pi_2(t_2) = d - 2R$ and $\pi_2(t) < d - 2R$ for $t < t_2$. Without loss of generality, we assume

$$\pi_2(t_2) = \pi(1, \{1, 2\}, t_2) = d - 2R$$ (66)

and

$$\langle \hat{c}_1(t_2), \hat{c}_2(t_2) \rangle \leq \langle \hat{c}_1(t_2), \hat{c}_1(t_2) \rangle \leq d$$ (67)

so then

$$\langle \hat{c}_1(t_2), \hat{c}_2(t_2) \rangle \geq d - 2NR$$ (68)

Now apply the same arguments, we conclude that at the moment $t_2$, there is at least one cluster $C_3$ adjacent to either $C_1$ or $C_2$, or both, such that

$$\langle \hat{c}_1(t_2), \hat{c}_3(t_2) \rangle \geq d - 2(N + 1)R$$ (69)

this then, in turn, implies that

$$\pi_3(t_2) \geq \pi(1, \{1, 2, 3\}, t_2) \geq d - 2(N + 1)R$$ (70)
Repeat the process, we then find a sequence of decreasing moments $t_1 \geq \cdots \geq t_{M-1} > 0$ such that at each time $t_k$, there is an index $i \in S$ and a subset $S' \subset S$ with $|S'| = k + 1$ such that for any $j \in S'$

\[ \langle \hat{c}_i(t_k), \hat{c}_j(t_k) \rangle \geq d - 2 \sum_{l=0}^{k-1} N_l R > 0 \]  

(71)

Now consider the configuration $\varphi_{t_{M-1}}(p)$, inequality (71) then implies that the center of each cluster is on one side of the hyperplane that is perpendicular to $\hat{c}_i(t_{M-1})$. But this is a contradiction because $\varphi_{t_{M-1}}(p)$ is centered at the origin.

5.3 Proof of theorem 5.1

Proof of theorem 5.1. Let $D$ be defined by equation (52), let $L_1$, $T$ be defined in the statement of theorem 5.2, let $L_2$ be defined in the statement of theorem 5.4, and we define $L := \max\{L_1, L_2, D + 2R\}$.

The proof is done by contradiction. i.e, we assume $\varphi_{t_0}(p)$ admits $\sigma(l, \epsilon)$ with $l \geq L$ and $t_0 \geq T$, yet the flow $\varphi_{\geq t_0}(p)$ doesn’t intersect $X_{ij}(I)$ for any $ij \in E$. By theorem 5.2 each adjacent-cluster distance will keep greater than $L_1$, so the radius of each cluster is bounded above by $R$ along the gradient flow $\varphi_{\geq t_0}(p)$. Then by theorem 5.4 the gradient flow $\varphi_{\geq t_0}(p)$ is contained in $P^b$ for some $b > 0$.

Such a gradient flow will converge to the set of equilibria as we will see in the next section. However, the diameter of each equilibrium is bounded above by $D$, so there must exist a moment $t > t_0$ at which the gradient flow intersects $X_{ij}(D)$ for some $ij \in E$. Since $L \geq D + 2R$, so during the period $[t_0, t]$, the gradient flow intersects a dissipation zone $X_{ij}(d)$ for some $d \in [L - 2R, L + 2R]$ which is a contradiction. \qed

6 Convergence of the gradient flow

Our goal in this section is to prove that each gradient flow associated with the model described by equation (1) converges to the set of equilibria. This will be done after theorem 6.1.

Theorem 6.1 (Swarm aggregation). Consider the RMA system described by equation (1), for each configuration $p$, there exists a positive number $b > 0$ such that $\varphi_{\geq 0}(p)$ is contained in $P^b$.

We will first assume theorem 6.1 to establish the convergence of the gradient flow, and then give a proof of it. Recall the $\omega$-limit set of a configuration $p$ is defined to be

\[ \omega(p) := \bigcap_{t>0} \bigcup_{s \geq t} \varphi_s(p) \]  

(72)
By theorem 2.1 and theorem 6.1, for each initial condition \( p \), there are two positive numbers \( a \) and \( b \) associated with \( p \) such that the gradient flow \( \varphi_{\geq 0}(p) \) is contained in the compact set \( P_a^b \).

We now show that the \( \omega \)-limit set is nonempty and consists only of equilibria. Let \( \{ t_i \}_{i \in \mathbb{N}} \) be a sequence approaching to infinity, since \( \{ \varphi_{t_i}(p) \}_{i \in \mathbb{N}} \) is contained in a compact set \( P_a^b \), there must exist at least one accumulation point. On the other hand, all accumulation points must be equilibria. Because if not, say there exists a configuration \( p' \) with \( f(p') \neq 0 \) and \( p' \) is the limit of the sequence \( \{ \varphi_{t_i}(p) \}_{i \in \mathbb{N}} \). By continuity of the gradient field, there exists an open neighborhood \( U \) of \( p' \) and a time period \( \tau \) such that \( |f(\varphi_s(p''))| > \frac{1}{2}|f(p')| \) for any \( p'' \in U \) and any \( s \in [0, \tau] \). By passing to a subsequence, if necessary, we assume that the sequence \( \{ \varphi_{t_i}(p) \}_{i \in \mathbb{N}} \) is contained in \( U \) and \( t_{i+1} - t_i > \tau \). So then

\[
\int_0^\infty |f(\varphi_s(p))|^2 ds > \sum_{i \geq 1} \frac{1}{4}|f(p)|^2 \tau = \infty
\]  

(73)

which contradicts the fact that the potential function \( \Psi(p) \) is bounded below as \( p \) varies over \( P \). So at this moment, we have actually proved theorem 1.1, the main theorem of this paper. The remaining section is devoted to the proof of theorem 6.1.

**Proof of theorem 6.1.** The proof is done by contradiction, i.e., we assume that there exists a gradient flow \( \varphi_{\geq 0}(p) \) such that the radius of the configuration \( \varphi_t(p) \) can’t be bounded above as \( t \) approaches to infinity. Then by theorem 3.4, there is an infinite sequence \( \{ \varphi_{t_i}(p), \sigma(t_i, \epsilon) \}_{t_i \rightarrow \infty} \) such that

1. each clustering \( \sigma(t, \epsilon) \) on \( \varphi_{t_i}(p) \) induces the same partition on \( V \), the set of vertices in \( \Gamma \).

2. the sequence \( \{ t_i \}_{i \in \mathbb{N}} \) approaches to infinity.

3. the radius of each cluster in \( \varphi_{t_i}(p) \) is bounded above by \( R \) for all \( t_i \).

Let \( L \) and \( T \) be defined in the statement of theorem 5.1, and let \( I := [L - 2R, L + 2R] \), and we assume each distance \( d \in I \) is absent in any equilibrium because there is no harm to assume \( L \) is large. First we see that there is an index \( k_0 \) such that if \( k \geq k_0 \), then \( t_k \geq T \) and \( l_k \geq L \), we may assume that \( k_0 = 1 \) by passing to a subsequence. Then by theorem 5.1 for each \( t_k \), there is a moment \( t_k' > t_k \) such that the gradient flow falls into \( X_{ij}(I) \) for some \( ij \in E \) at time \( t_k' \). For each \( d \in I \), we have \( \tilde{\mu}_{ij}(d) > 0 \), and \( I \) is a closed interval, so

\[
\tilde{\mu}_{ij}(I) := \min \{ \tilde{\mu}_{ij}(d) | d \in I \} > 0
\]  

(74)

Let \( \tilde{I} \) be a thickened version of \( I \) defined by

\[
\tilde{I} := [L - 2R - \mu, L + 2R + \mu]
\]  

(75)

with \( \mu > 0 \) chosen so that \( \tilde{\mu}_{ij}(\tilde{I}) \geq \tilde{\mu}_{ij}(I)/2 \). Let \( v_p \) be defined by equation 50, and let \( \tau_\mu := \frac{l_\mu}{\sqrt{2} v_p} \). Then the loss of potential of the gradient flow exceeds \( \tilde{\mu}_{ij}^2(I) \tau_\mu/4 \) over the period \( [t_k', t_k' + \tau_\mu] \).
By passing to a subsequence, if necessary, we may assume that the two time sequences \( \{t_k\}_{i \in \mathbb{N}} \) and \( \{t'_k\}_{i \in \mathbb{N}} \) are interlacing, i.e.,

\[
t_k < t'_k < t_{k+1} < t'_{k+1}, \quad \forall k \in \mathbb{N}
\]

and let

\[
\Delta := \min\{\hat{\mu}_{ij}^2(i)\tau / 4 | ij \in E\} > 0
\]

then the loss of potential along the gradient flow \( \varphi_{\geq 0}(p) \) is

\[
\int_0^\infty |f(\varphi_t(p))|^2 dt > \sum_{k \in \mathbb{N}} \Delta = \infty
\]

This contradicts to the fact that the potential function \( \Psi(p) \) is bounded below over the configuration space.

\[\square\]

Acknowledgements

The author here is grateful for discussions with Prof. Roger W. Brockett.

References

[1] **Xudong Chen**, *Gradient flow for organizing multi-agent system*, accepted by 2014 American Control Conference.

[2] **Xudong Chen**, *Genericity of equivariant Morse functions in multi-agent systems with complete graph*, preprint

[3] **Xudong Chen** and **Roger W. Brockett** *Centralized and decentralized formation control with variable interaction laws*, submitted to 2014 53rd IEEE Conference on Decision and Control

[4] **Xudong Chen**, *Multi-agent systems with reciprocal interaction laws*, Thesis, Harvard University, School of Engineering and Applied Sciences, (to appear)

[5] Breder, C. M., *Equations descriptive of fish schools and other animal aggregations*, Ecology (1954): 361-370.

[6] Gazi, Veysel, and Kevin M. Passino, *Stability analysis of swarms.* IEEE Transactions on Automatic Control, 48.4 (2003): 692-697.

[7] Gazi, Veysel, and Kevin M. Passino, *A class of attractions/repulsion functions for stable swarm aggregations*, International Journal of Control 77.18 (2004): 1567-1579.

[8] Chu, Tianguang, Long Wang, and Tongwen Chen, *Self-organized motion in anisotropic swarms*, Journal of Control Theory and Applications 1.1 (2003): 77-81.
[9] Olfati-Saber, Reza, and Richard M. Murray, Consensus problems in networks of agents with switching topology and time-delays, Automatic Control, IEEE Transactions on 49.9 (2004): 1520-1533.

[10] Marshall, Joshua A., and Mireille E. Broucke, On invariance of cyclic group symmetries in multiagent formations, IEEE CONFERENCE ON DECISION AND CONTROL. Vol. 44. No. 1. IEEE; 1998, 2005.

[11] Baillieul, J., and A. Suri, Information patterns and hedging Brock-ett’s theorem in controlling vehicle formations, Decision and Control, 2003. Proceedings. 42nd IEEE Conference on. Vol. 1. IEEE, 2003.

[12] Anderson, Brian DO, Morse theory and formation control, Control & Automation (MED), 2011 19th Mediterranean Conference on. IEEE, 2011.

[13] Brian D.O. Anderson and U. Helmke, Equivariant Morse Theory and Formation Control, Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on. IEEE, 2013.

[14] Belabbas, M-A., Decentralized formation control part i: Geometric aspects, arXiv preprint arXiv:1101.2416 (2011).

[15] Krick, Laura, Mireille E. Broucke, and Bruce A. Francis, Stabilisation of infinitesimally rigid formations of multi-robot networks, International Journal of Control 82.3 (2009): 423-439.

[16] M. J. Field, Equivariant dynamical systems, Transactions of the American Mathematical Society 259.1 (1980): 185-205.