View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Illinois Digital Environment for Access to Learning and Scholarship Repository

© 2012 Seyed Rasoul Etesami

ANALYZING THE OPINION DYNAMICS MODELS DISCRETE & CONTINUOUS

BY

SEYED RASOUL ETESAMI

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Adviser:

Assistant Professor Angelia Nedić

ABSTRACT

In this thesis we analyze some of the opinion dynamics in both discrete and continuous cases. In the discrete case, we will find some criteria under which we can say more about the behavior of the dynamics such as convergence of the agents to the same opinion, or consensus. For this purpose, we first consider the agent-based bounded confidence model of the Hegselmann-Krause where multiple agents want to agree on a common scalar, or they can be divided in several subgroups, with each subgroup having its own agreement value. In this model, we restrict ourselves to the case when all the agents have the same bound of confidence, often referred to as homogeneous case. We are interested to study the number of iterations which is enough for the termination of the Hegselmann-Krause algorithm. In other words, we want to give an upper bound on the number of iterations which guarantees the termination of the algorithm independently of reaching a consensus or not. Assuming the consensus is achieved in the Hegselmann-Krause model, we first give an upper bound on the number of iterations and then we provide another upper bound without any assumption.

In chapter 3 we use some analysis based on Lyapunov function theory to improve our upper bound substantially. In our analysis we use two differnt type of Lyapunov functions which each of them gives us a polynomial upper bound for the termination time. In chapter 4 we consider the Hegselmann-Krause model in higher dimensions. We will see that in higher dimensions we don't have lots of nice properties which exist in the scalar case. Then, we will find some upper bounds for the termination time. Also, at the end we will consider an extension of the Hegselmann-Krause model to continuous case such that the time is discrete but the density of the agents is continuous over the real line. In chapter 5 we use the matrix representation for the discrete dynamics and we provide some conditions on a chain of stochastic matrices based on their decomposition by permutation matrices such

that it can guarantee the convergence of the chain to a consensus matrix. Also, we provide some examples and one necessary condition for finite time convergence of an especial case of averaging gossip algorithms.

To my parents, for their love and support.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In this chapter we bring some of the notation that we will use through the thesis. After that, we discuss some of the known results, as the background material, and we use them later to prove our results.

1.1 Notation and Terminology

We consider all the vectors as column vectors. For a vector x we write x_i to denote its ith entry. We use I for the identity matrix. Also for a matrix A, we write A_i and A^j to denote its ith row and jth column, respectively. We use $R^{m\times m}$ to denote the set of $m \times m$ matrices with real entries. Also we denote the real numbers and integers by R and Z, respectively. For a set S we write $diam(S) = \max\{|a-b| : a, b \in S\}$. For a matrix $A = (a_{ij})$ with non-negative entries we define $\min^+(A)$ to be $\min\{a_{ij}|a_{ij} > 0\}$. We use e_i to denote a vector which all of its entries equal to zero except for the ith entry which is equal to 1. Also, we use e and J to denote a vector and a matrix with all of their entries equal to 1, respectively. We use x' for the transpose of vector x. Also, for two vectors x and y we use $x'y$ for the inner product of these two vectors. For an arbitrary vector x, We use $\|x\|$ to denote the euclidian norm of x, i.e. $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$. A matrix is said stochastic if all of its entries are non-negative and the sum of entries in each row is equal 1. A scrambling matrix is a stochastic matrix such that the inner product of each pair of its rows is positive. A matrix A is doubly stochastic if A and A' are both stochastic. Suppose that we have a chain of stochastic matrices $A(1), A(2), \ldots$. Then, with left product of this chain we mean $\ldots A(2)A(1)$. We usually refer to the left product of this chain from time t_0 to time $t_1 > t_0$ by $A(t_0, t_1) = A(t_1)\ldots A(t_0)$. A permutation matrix is a doubly stochastic matrix which has exactly one 1 in each row and column. We say that a column of a matrix is positive if all entries in that column are positive. We denote the set of all the $n \times n$ permutation matrices with $\mathcal{G}_{n\times n}$. Since we usually consider the dimension of the matrices to be n, for simplicity we often use $\mathcal{G}$ instead of $\mathcal{G}_{n\times n}$. The set of permutation matrices with ordinary product of matrices is an algebraic group which identity element being I. Since the number of permutation matrices of size n is $n!$, thus the cardinality of $\mathcal{G}$ is $n!$. Let $f : (R^m \times Z^+) \to R$ be a function and assume that $x(t_0) \in R^m$ for an arbitrary integer $t_0 \in Z^+$. We say that $\{x(k)\}$, defined by:

$$x(k+1) = f(x(k), k) \quad \forall k \geq t_0 \tag{1.1}$$

is a dynamic at starting time t_0. Also, we say that a point $x \in R^m$ is an equilibrium poind for the dynamic (1.1), if $x = f(x, k)$ for any $k \geq t_0$.

1.1.1 Graph theory and matrix representation

A simple graph $T(n, \mathcal{E})$ is a set of n nodes with indices from $\{1, 2, \ldots, n\}$ such that the node i is connected to node j if and only if $\{i, j\} \in \mathcal{E}$. We say $\mathcal{E}$ is the set of edges of the graph. We say that the vertex i is neighbor with j if and only if $\{i, j\} \in \mathcal{E}$. We denote the set of all the neighbors of i with $N_i(T)$. A directed graph is a graph that the set of its edges is composed of ordered pairs. In other words, in a directed graph $H(n, \mathcal{E})$ there is an edge from i to j if and only if $(i, j) \in \mathcal{E}$. A path in a graph is an ordered sequence of vertices such that there is an edge between every two consecutive vertices. We say that a simple graph is connected if there is a path between every two vertices of the graph, and, similarly we say a directed graph is strongly connected if there exists a directed path between every two nodes in the graph. For every $n \times n$ symmetric matrix A with non-negative entries we can consider a graph with n nodes such that node i is connected to node j if and only if the $A_{ij} = A_{ji} > 0$. Similarly for a $n \times n$ matrix B with non-negative entries we can assign a directed graph such that the link (i, j) is in the graph if and only if $B_{ij} > 0$.

1.1.2 Group theory

A finite group $(G, \times)$ of order m is a set of m elements with a product induced on it such that has three properties:

1. G is closed the product $\times$.
2. There exists an identity element e_0 in G such that the product of e_0 with every element from left and right is the same element.
3. Every element $x \in G$ has its inverse, it means that the multiplication of every element with its inverse is identity (e_0).

The order of an element $a \in G$ is the smallest positive integer k such that the multiplication of a with itself k times is identity, i.e. $a^k = e$. We usually denote this number by $O(a)$. A subgroup is a proper subset of a group which satisfies all the conditions of a group, (it means that a subgroup is a group itself). A cyclic group is a group which is generated by different powers of an element. Therefor, for an arbitrary element a in a group G, the cardinality of the cyclic group which is generated by a is $O(a)$. We usually refer to this cyclic group by $<a>= \{e_0, a, a^2, \ldots, a^{O(a)-1}\}$. For a positive integer k the number of positive integers which are less than k and are prime with respect to k is denoted by $\varphi(k)$.

1.1.3 Real Analysis

A scalar function f is convex if it has the following property:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y),\ \forall x, y \in D_f, \forall \lambda \in [0,1],$$

where D_f is the domain of f. If in the above relation the inequality is strict whenever $x \neq y$ and $\lambda \in (0,1)$, we say that the function is strictly convex. Moreover, we say that a function is strongly convex with parameter $\mu > 0$ if:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\mu}{2}\lambda(1-\lambda)\|x-y\|^2,\ \forall x, y \in D_f.$$

We say that a sequence of functions $f_n(x) : R \to R, n = 1, 2, \ldots$ converges uniformly to a function $f(x)$ if:

$$\forall \epsilon > 0 \; \exists N : \forall n \geq N, \forall x \in R \; |f_n(x) - f(x)| < \epsilon.$$

Also, we say a function $f : D_f \subset R \to R$ is uniformly continuous on D_f if:

$$\forall \epsilon > 0 \; \exists \delta > 0 : |x - y| < \delta \Rightarrow |f_n(x) - f(x)| < \epsilon.$$

1.2 Basic Results

In this part we present some basic facts which we use in the subsequent chapters to prove our main results.

1.2.1 Matrix analysis

Lemma 1.1. *(Birkhof-von Neumann, [1], p. 185) Every doubly stochastic matrix can be represented by a convex combination of permutation matrices.*

Suppose that we show all the $n \times n$ permutation matrices by $P_1, P_2, \ldots, P_{n!}$. The above lemma states that for a $n \times n$ doubly stochastic matrix A, there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_{n!}$ in [0,1] such that:

$$A = \sum_{k=1}^{n!} \lambda_k P_k.$$

Theorem 1.1. *([2]) Suppose $C \in R^{n \times n}$ is a stochastic matrix and x and y be two vectors such that $y = Cx$. Then,*

$$diam(y) \leq (1 - \mu(C))diam(x).$$

Where $\mu(C) = \min_{i \neq j}(\sum_{k=1}^{n} \min(c_{ik}, c_{jk}))$. *Furthermore, when C is a scrambling matrix* *with* $\min^{+} C \geq \delta$*, then we have* $\mu(C) \geq \delta$ *and, thus,*

$$diam(y) \leq (1-\delta)diam(x).$$

A short proof of the above theorem can be found in [2]. Stochastic and scrambling matrices are so applicable and you can find some more properties of them in [3], [4], [5] and [6].

Lemma 1.2. *([7], p. 112) Suppose* $\{A(t)\}$ *is a sequence of stochastic matrices such that in each* $A(t)$ *has at least one positive column. Furthermore, suppose there exists a positive* δ *such that* $\min^{+}(A(t)) > \delta$ *for all* t*. Then, the product of this sequence will converge to a consensus matrix.*

Lemma 1.3. *Suppose* A *and* B *are two matrices with non-negative entries such that their diagonal entries are all positive. If the inner product of two rows of either* A *or* B *is positive, then the inner product of the corresponding rows in both* AB *and* BA *are also positive.*

Proof. Let us assume that the inner product of ith and jth rows of A is positive, i.e. $A_i'.A_j > 0$. Then,

$$(AB)_i' = (A_i'B^1, A_i'B^2, \ldots, A_i'B^n)$$
$$(AB)_j' = (A_j'B^1, A_j'B^2, \ldots, A_j'B^n).$$

Therefore, the inner product of ith and jth rows in AB is equal to:

$$R = \sum_{k=1}^{n}(A_i'B^k)(A_j'B^k).$$

For all $k = 1, \ldots, n$ we have:

$$(A_i'B^k).(A_j'B^k) = (\sum_{t=1}^{n} a_{it}b_{tk})(\sum_{t=1}^{n} a_{jt}b_{tk}) \geq \sum_{t=1}^{n} a_{it}a_{jt}b_{tk}^2.$$

Therefore, we can write:

$$R \geq \sum_{k=1}^{n} \sum_{t=1}^{n} a_{it} a_{jt} b_{tk}^2 = \sum_{t=1}^{n} a_{it} a_{jt} \sum_{k=1}^{n} b_{tk}^2. \tag{1.2}$$

Now if we let $c_B = \min_{k=1,\dots,n}\{b_{kk}\}$ (and similarly $c_A = \min_{k=1,\dots,n}\{a_{kk}\}$), because of positivity of the diagonal entries in B and A we have $c_B, c_A > 0$, implying $\sum_{k=1}^{n} b_{tk}^2 \geq c_B^2$. Thus, from (1.2) we obtain:

$$R \geq c_B^2 \sum_{t=1}^{n} a_{it} a_{jt} = c_B^2 (A_i' A_j) > 0.$$

Therefore the inner product of the $(AB)_i$ and $(AB)_j$ is also positive. For the other statement we want to show that the inner product of $(BA)_i$ and $(BA)_j$ is also positive. For this, it suffices to $(AB)_i'(AB)_j > 0$ with the assumption that $B_i' B_j > 0$. therefore we can again, work with AB instead of BA but this time we suppose $B_i' B_j > 0$ and we don't have any assumption on positivity of $A_i' A_j > 0$. Therefore the inner product of these two rows is equal to:

$$\begin{aligned}
L &= \sum_{k=1}^{n} (A_i' B^k)(A_j' B^k) \\
&= \sum_{k=1}^{n} (\sum_{\ell=1}^{n} a_{i\ell} b_{\ell k})(\sum_{t=1}^{n} a_{jt} b_{tk}) \\
&\geq (\sum_{k=1}^{n} a_{ii} b_{ik} . a_{jj} b_{jk} = a_{ii} a_{jj} \sum_{k=1}^{n} b_{ik} b_{jk} \\
&= a_{ii} a_{jj} (B_i' B_j) \geq c_A^2 (B_i' B_j).
\end{aligned}$$

Again since the diagonal elements of A are all positive and since $B_i' B_j > 0$, we have $L > 0$. This completes the proof. **Q.E.D.**

1.2.2 Group theory

Lemma 1.4. *(Zorn Lemma, [8]) Suppose that $(P, \leq)$ is a partially ordered set such that every chain in P has a maximal element. Then, there exists an element in P which is maximal.*

Lemma 1.5. *([9], p. 123) In a cyclic group $< a >$ with cardinality n the number of different generators is $\varphi(n)$.*

Theorem 1.2. *([9], p. 89-96) The cardinality of each subgroup of a finite group divides the cardinality of the group.*

Since the cardinality of a cyclic group which is generated by an particular element is equal to the order of that element, from above theorem we have the following corollary.

Corollary 1.1. *The order of every element in an arbitrary finite group divides the cardinality of the group.*

1.2.3 Convex functions and measure theory

Lemma 1.6. *[10] Consider a sequence of nested sets $A_1 \subset A_2 \subset \dots$ which are all measurable with respect to a given measure μ. Then, we have:*

$$\lim_{k\to\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k).$$

Furthermore, if we have $\dots \subset A_2 \subset A_1$ and $\mu(A_1) < \infty$, then:

$$\lim_{k\to\infty} \mu(A_k) = \mu(\bigcap_{k=1}^{\infty} A_k).$$

Lemma 1.7. *[11]. Every convex function is continuous in relative interior of its domain. Therefore, a convex function which is defined over the real line is continuous.*

Lemma 1.8. *Suppose $\{f_k(x)\}_{k=1}^{\infty}$ is a sequence of convex functions over R which converges to $f(x)$ point-wise. Then, $f(x)$ is also convex.*

Proof. Choose x and y arbitrary and fix $\lambda \in [0,1]$. We have:

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \lim_{k\to\infty} f_k(\lambda x + (1-\lambda)y) \\ &\leq \lim_{k\to\infty} (\lambda f_k(x) + (1-\lambda) f_k(y)) \\ &= \lambda f(x) + (1-\lambda) f(y). \end{aligned}$$

Q.E.D.

Lemma 1.9. *(Jensen Inequality, [11]) Suppose that* $f(\cdot)$ *is a convex function and* X *is a random variable, then* $f(\mathbb{E}\ (X)) \leq \mathbb{E}\ (f(X))$. *In particular, if* $a = (a_1, a_2, \ldots, a_n)$ *is a stochastic vector and* $x_j \geq 0,\ \forall j = 1, 2, \ldots n$, *then,*

$$\sum_{j=1}^{n} a_j x_j \geq \Pi_{j=1}^{n} x_j^{a_j}.$$

CHAPTER 2

UPPER BOUND ON THE TERMINATION TIME OF THE HEGSELMANN-KRAUSE MODEL

2.1 General Concepts

One of the important questions in the network problems is the concept of consensus or agreement among agents. In these kind of problems we have a set of agents which want to share their individual information so that they agree on a same opinion. There have been developed many of local algorithms such that guarantee convergence to consensus under certain conditions [7], [12], [5], [13]. In some of these iterative algorithms the opinion of the agents in the next step depends on the opinion of the agents at the past stages. In other words, not only their opinion profiles depend on the time steps, but also the dynamic of the algorithm changes at every stage based on the past opinion profiles. In this section we describe an opinion dynamics model which was introduced earlier by Hegselmann and Krause and provide some basic concepts for the model. Let us assume we have n agents and we want to survey the interactions among their opinions. Specially we consider the following matrix representation form for the dynamic:

$$x(t+1) = A(t, x(t), \epsilon(t))x(t), \tag{2.1}$$

which A is a n by n matrix. The entries of A are function of time step t, current profile $x(t)$ and communication regime $\epsilon(t)$. This dynamics model is general and therefore we confine ourselves to a typical class of these dynamics which was introduced earlier as an agent-based bounded confidence [14]. This dynamics usually comes to the account when the agents are more conservative about their opinions. In this case, every agent will interact with agents which have opinions closer to its own opinion. We will continue with the definition of homogenous Hegselmann-Krause model which can be considered under the framework of (2.1). Suppose we have $n \in N$ agents and a static bound of confidence $\epsilon > 0$. Given an initial profile $x(0)$ define the matrix $A(t, x(t), \epsilon(t))$ in (2.1) in the following way:

$$x_i(t+1) = |N_i(x(t))|^{-1} \sum_{j \in N_i(x(t))} x_j(t)$$

$$A_{ij}(t) = \begin{cases} \frac{1}{|N_i(x(t))|}, & \text{if } j \in N_i(x(t)) \\ 0, & \text{else} \end{cases} \tag{2.2}$$

where $N_i(x(t)) = \{1 \leq j \leq n : \|x_i(t) - x_j(t)\| \leq \epsilon\}$. Also, we use $|N_i(x(t))|$ to show the number of elements in $N_i(x(t))$. Therefore we notice that according to the above definition we can write $x(t+1) = A(t)x(t)$, where $A(t)$ is a stochastic matrix and the positive entries in each row are the same. Furthermore, we can say that the positive entry in jth row is equal to $\frac{1}{|N_j(x(t))|}$. It has proven before that for any initial profile and any positive bound of confidence the Hegselmann-Krause model will terminate after finitely many steps [14]. However, depending on the initial profile and the bound of confidence, the final state could be consensus or not. It has been shown that the final state is a function of initial profile $x(0)$ and bound of confidence ϵ [15]. More insight about the behavior of the Hegselmann-Krause model can be found in [7]. Also, note that in this chapter our focus is only on the scalar case of the Hegselmann-Krause model where each agent has a real value as her opinion and we will postpone the higher dimensions to the next chapter. We next list some of the properties of Hegselmann-Krause model in one dimension which we use frequently in the rest of this chapter. We refer to the set of opinions of agents in time step t as the tth profile. Given an initial profile, without loss of generality we can relabel our agents such that the initial profile values are non decreasing. In other words: $x_1(0) \leq x_2(0) \leq \ldots \leq x_n(0)$. We note that Hegselmann-Krause model has these two following properties (as shown in [14]).

- The Hegselmann-Krause model preserves the order of the agents, it means that if in the initial profile we rearrange the agents opinion as a non decreasing sequence, then this order will be preserved in the next profiles as well.

- If the difference between opinions of two subsequent agents in a profile is greater than ϵ, then this difference will remain more than ϵ in all the next profiles. We usually refer to this fact as a "break" between agents. As a result, a necessary condition for consensus is that there is no break in the initial profile. Furthermore, if we assume that consensus is reachable, then we have break in non of the profiles.

2.2 A Loose Upper Bound

In this section we will develop an upper bound on the number of iterations in Hegselmann-Krause model needed to reach an equilibrium point. For this purpose, we first assume that consensus is reachable and provide an upper bound under this assumption. Later we will find an upper bound in the general sense without assuming that consensus is reached.

2.2.1 Upper Bound Under Consensus Assumption

First, notice that we can formulate the Hegselmann-Krause model as $x(m+1) = A(m)x(m)$ where A is a stochastic matrix with the same positive entries in each row. The positive entries in jth row are equal to $\frac{1}{|Nj(x(m))|}$ at instant m. Next lemma shows that in the Hegselmann-Krause model the information can be transferred from each agent to another agent after not more than $\frac{n+1}{2}$ time steps. Here, We assume $n \geq 2$ (otherwise the case is not interesting).

We consider the matrix representation of the Hegselmann-Krause model and interpret the entry changes when this algorithm is applying on the agents. In every stage we can consider the Hegselmann-Krause matrix composed of multiplication of two matrices; One is a diagonal matrix which carries the weights and the second is an incident 0-1 matrix which describes the neighborhood connections and the existence of link among agents. In other words, $A(m) = D(m)C(m)$, such that:

$$D_{ij}(m) = \begin{cases} 0, & i \neq j \\ \frac{1}{|N_i(x(m))|}, & i = j \end{cases} \qquad C_{ij}(m) = \begin{cases} 1, & j \in N_i(x(m)) \\ 0, & j \notin N_i(x(m)) \end{cases}$$

Lemma 2.1. *Suppose that Q is a $n \times n$ which is defined as following:*

$$Q_{ij} = \begin{cases} 1, & i = j, i = j+1, j = i-1 \\ 0, & else \end{cases}$$

. Then, $Q^{[\frac{n+1}{2}]}$ has at least one positive column, where $[\frac{n+1}{2}]$ denotes the integer part of $\frac{n+1}{2}$.

Proof. It can be seen easily that Q is the adjacency matrix of an undirected path of length $n-1$ such that each node has self loop. On the other side, we know that for every $k \in Z^+$, Q^k_{ij} shows the number of tours of length k between nodes i and j in the corresponding graph. Now by focusing on the middle node in the path, it can bee seen that there is always a tour of length $[\frac{n+1}{2}]$ between this middle node and every other node in the path. (note that the graph has self loops on each vertex). This means that each entry in the middle column of $Q^{[\frac{n+1}{2}]}$ ($[\frac{n+1}{2}]$th column) is at least one and hence is positive. **Q.E.D.**

Lemma 2.2. *Suppose that we consensus is reached in the Hegselmann-Krause model. Then, the consecutive left product of at most every $[\frac{n+1}{2}]$ of the matrices $A(m)$ is a scrambling matrix.*

Proof. We show that for every $k \in Z^+$, $A(k, k+[\frac{n+1}{2}])$ has at least one positive column where, by the definition we can write:

$$\begin{aligned} A(k, k+[\frac{n+1}{2}]) &= A(k+[\frac{n+1}{2}]) \dots A(k+1)A(k) \\ &= D(k+[\frac{n+1}{2}])C(k+[\frac{n+1}{2}]) \dots D(k+1)C(k+1)D(k)C(k). \end{aligned}$$

Also, we note that the existence of diagonal matrices $D(m), m = k \dots k+[\frac{n+1}{2}]$ doesn't affect the positivity of the entries $A(k, k+[\frac{n+1}{2}])$. However, they affect the amount of the entries. Since we are interested to the scrambling property of $A(k, k+[\frac{n+1}{2}])$, thus, simply we can remove the diagonal matrices $D(m), m = k \dots k+[\frac{n+1}{2}]$ from the above product and just survey the scrambling property of the $C(k+[\frac{n+1}{2}]) \dots C(k+1)C(k)$. Now, since we assumed that consensus is reached, therefore in each time step, every middle agent has at least one neighbor before and one neighbor after herself. Otherwise, based on the facts that we mentioned earlier there would be a break in one of the profiles and this is contradiction with this fact that consensus is reached. In the other words, every $C(t), t = k, \ldots, k + [\frac{n+1}{2}]$ has the following configuration:

$$C_{n\times n}(t) = \begin{pmatrix} 1 & 1 & * & * & * & \cdots & * & * \\ 1 & 1 & 1 & * & * & \cdots & * & * \\ * & 1 & 1 & 1 & * & \cdots & * & * \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ * & * & * & * & * & \cdots & 1 & 1 \end{pmatrix}. \tag{2.3}$$

Note that in the above configuration we assumed that the initial profile $x(0)$ has been ordered. Also, each "*" shows that the corresponding entry can admit 0 or 1 depends on its situation. Therefore, for every $t = k, \ldots, k + [\frac{n+1}{2}]$, we can write $C(t) = Q + R(t)$, where Q is the same matrix introduced in lemma (2.1). Also, $R(t)$ is a 0-1 matrix which captures the entries corresponding to *'s in (2.3). Now, let us show the set of positive entries of a matrix A by $P(A)$. Then, we have:

$$P(A(k, k + [\frac{n+1}{2}])) = P(\prod_{j=0}^{[\frac{n+1}{2}]} C(k+j)) = P(\prod_{j=0}^{[\frac{n+1}{2}]} (Q + R(k+j)). \tag{2.4}$$

Therefore, we can write $\prod_{j=0}^{[\frac{n+1}{2}]}(Q + R(k+j) = Q^{[\frac{n+1}{2}]} + T$ which T is some matrix with non-negative entries. Thus,

$$P(\prod_{j=0}^{[\frac{n+1}{2}]} (Q + R(k+j))) = P(Q^{[\frac{n+1}{2}]} + T) \supseteq P(Q^{[\frac{n+1}{2}]}). \tag{2.5}$$

By comparing (2.4) and (2.5) and using lemma (2.1) we get that $A(k, k + [\frac{n+1}{2}])$ has at least one positive column and thus is a scrambling matrix. **Q.E.D.**

Lemma (2.2) guarantees that in the worst case the product of at most $[\frac{n+1}{2}]$ consequent matrices of Hegselmann-Krause model is a scrambling matrix. However, it may happen that product of less than $[\frac{n+1}{2}]$ of the consecutive matrices has also the scrambling property. Now, let us assume $d \leq [\frac{n+1}{2}]$ is the number of steps such that $A(k, k+d)$ is a scrambling matrix. Therefore, since $\min^+(A(t)) \geq \frac{1}{n}, \forall t = 0, 1, \ldots,$ then we have:

$$\overset{+}{\min} A(k, k+d) \geq (\frac{1}{n})^d. \tag{2.6}$$

Theorem 2.1. *Suppose in the Hegselmann-Krause model consensus is reached, then an upper bound for the number of iterations is:*

$$(\frac{\ln(\frac{\epsilon}{D})}{\ln(1 - (\frac{1}{n})^{[\frac{n+1}{2}]})})\frac{(n+1)}{2}$$

which can be estimated by:

$$\ln(\frac{D}{\epsilon})\sqrt{n}^{n+3},$$

where $D = diam(x(0))$.

Proof. Let us assume T is a particular time step. Then, we can write:

$$A\Big(0,\ [\frac{n+1}{2}][\frac{T}{\frac{n+1}{2}}]-1\Big) = \prod_{i=1}^{[\frac{T}{\frac{n+1}{2}}]} A\Big((i-1)[\frac{n+1}{2}],\ i[\frac{n+1}{2}] - 1\Big).$$

By using (2.6) we have:

$$\overset{+}{\min} A\Big((i-1)[\frac{n+1}{2}],\ i[\frac{n+1}{2}] - 1\Big) \geq (\frac{1}{n})^{[\frac{n+1}{2}]}, \quad \forall i = 1, 2, \ldots, [\frac{T}{\frac{n+1}{2}}].$$

Furthermore, according to lemma (2.2), each of the $A\Big((i-1)[\frac{n+1}{2}], i[\frac{n+1}{2}]-1\Big)$ is a scrambling matrix. On the other side one can write:

$$x([\frac{n+1}{2}][\frac{T}{\frac{n+1}{2}}]-1) = A\Big(0, [\frac{n+1}{2}][\frac{T}{\frac{n+1}{2}}] - 1\Big)x(0), \quad [\frac{n+1}{2}][\frac{T}{\frac{n+1}{2}}]-1 \leq T-1.$$

Therefore, by applying theorem (1.1) $[\frac{T}{\frac{n+1}{2}}]$ times we get:

$$diam(x(T-1)) \leq diam\Big(x([\frac{n+1}{2}][\frac{T}{\frac{n+1}{2}}]-1)\Big) \leq \Big(1-(\frac{1}{n})^{[\frac{n+1}{2}]}\Big)^{[\frac{T}{\frac{n+1}{2}}]} diam(x(0)). \qquad (2.7)$$

Therefore, if we let T to be $[\frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{n})^{[\frac{n+1}{2}]})}\frac{(n+1)}{2}]$, by using the first and last term in (2.7) we get:

$$diam(x(T-1)) \leq \frac{\epsilon}{D} diam(x(0)) = \epsilon$$

This shows that after running $T = [\frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{n})^{[\frac{n+1}{2}]})}\frac{(n+1)}{2}]$ steps all the agents would be in the ϵ-neighborhood of each other and thus, in the next iteration the dynamic will reaches to consensus. Therefore, $[\frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{n})^{[\frac{n+1}{2}]})}\frac{(n+1)}{2}]$ is an upper bound for the termination time. Finally, if we estimate $\ln(1-(\frac{1}{n})^{[\frac{n+1}{2}]})$ by a larger amount $-(\frac{1}{n})^{\frac{n+1}{2}}$, then, we get the following upper bound:

$$\ln(\frac{D}{\epsilon})\sqrt{n}^{n+3}.$$

Q.E.D.

2.2.2 An Upper Bound in General Case

In this section with a complementary discussion we give an upper bound on the termination time of the Hegselmann-Krause model independent of any knowledge about the existence of consensus. This upper bound is almost the same as the previous upper bound with an additional term which is negligible in compare to the dominated term. First, we start with two definitions.

Definition 1. *For a given $r \in \{1, 2, \ldots, n-1\}$, we say that rth break happens in a particular profile for the first time if the number of breaks is r in that particular profile but it is less than r for all the previous profiles.*

Definition 2. *Suppose we have a profile that rth breaks happens in it for the first time.*

These breaks divide that profile to $r+1$ sub profiles and therefore, every sub profile has its own diameter. We define D_r to be The maximum value among all of these $r+1$ diameters.

Theorem 2.2. *Let T_n be the termination time for the Hegselmann-Krause model in one dimension. Then,*

$$T_n \leq \ln(\frac{D}{\epsilon})\sqrt{n}^{n+3} + \frac{n(n+1)}{2}$$

where, $D = diam(x(0))$.

Proof. Let $T = (\frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{n})^{[\frac{n+1}{2}]})} + n)\frac{(n+1)}{2}$ and consider $A(0,[T])$. If there is not any break in all of the profiles, then we will reach to consensus and according to theorem (2.1), T steps is enough for the termination of the dynamics. Otherwise, assume that the first break occurs in the tth profile such that $t = i[\frac{n+1}{2}] + j$ for some $i \geq 0$ and $0 \leq j < [\frac{n+1}{2}]$. Therefore, according to lemma (2.2) we can write $A(0, i[\frac{n+1}{2}])$ as the product of i scrambling matrices such that each of them includes $[\frac{n+1}{2}]$ matrices. Since before the first break there exist at least $i[\frac{n+1}{2}]$ time steps which we can look at them as the product of i scrambling matrices, therefore, by using theorem (1.1) i times we can see easily that D_1 satisfies to the following inequality:

$$D_1 \leq \left(1 - (\frac{1}{n})^{[\frac{n+1}{2}]}\right)^i D \tag{2.8}$$

In fact, the above inequality says that the product of these scrambling matrices decreases the diameter of the initial profile by a factor of $\left(1 - (\frac{1}{n})^{[\frac{n+1}{2}]}\right)^i$. Now, assume that we are running the algorithm T times, then after $t = i[\frac{n+1}{2}] + j$ steps, we still have

$$S = \frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{\frac{n+1}{2}})^{\frac{n+1}{2}})}\frac{n+1}{2} + \frac{n(n+1)}{2} - (i\frac{n+1}{2} + j) \tag{2.9}$$

more steps. By combining (2.8) and (2.9) we get:

$$\begin{aligned} S &= \frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D})}{\ln(1-(\frac{1}{n})^{\lceil\frac{n+1}{2}\rceil})} - i\Big) + \frac{n(n+1)}{2} - j \\ &\geq \frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D_1})}{\ln(1-(\frac{1}{n})^{\lceil\frac{n+1}{2}\rceil})}\Big) + \frac{n(n+1)}{2} - j \\ &\geq \frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D_1})}{\ln(1-(\frac{1}{n})^{\lceil\frac{n+1}{2}\rceil})}\Big) + \frac{(n-1)(n+1)}{2}. \end{aligned} \tag{2.10}$$

Now we can imagine that we just start the Hegselmann-Krause algorithm with this difference that we apply this algorithm on more than one profile (sub profiles) simultaneously and such that the diameter of each sub profile is less than or equal to D_1.

To complete the proof, suppose that we have totally ℓ breaks in the steady state or equivalently $\ell + 1$ subgroups. It is clear that $\ell \leq n - 1$. With repeating the same argument we can see that after ℓth break the diameter each subprofile is less than D_ℓ. But the remaining steps is more than or equal to:

$$\begin{aligned} &\frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D_\ell})}{\ln(1-(\frac{1}{n})^{\lceil\frac{n+1}{2}\rceil})}\Big) + \frac{(n-\ell)(n+1)}{2} \\ &\geq \frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D_\ell})}{\ln(1-(\frac{1}{n})^{\lceil\frac{n+1}{2}\rceil})}\Big). \end{aligned}$$

On the other side, we know that after ℓth break there is no any other break and all the separated sub profiles have diameter less than or equal to D_ℓ. Therefore, according to theorem (2.1) the remaining steps $\frac{n+1}{2}\Big(\frac{\ln(\frac{\epsilon}{D_l})}{\ln(1-(\frac{1}{\frac{n+1}{2}})^{\frac{n+1}{2}})}\Big)$ is enough for termination of the dynamics in each of the sub profiles. Therefore, T is an upper bound for the termination time. Finally, by applying the same approximation in theorem (2.1) we get the proper result. **Q.E.D.**

In this chapter, we studied the behavior of the Hegselmann-Krause model in scalar case and we gave a loose upper bound on the number of iterations. In this work we considered the homogenous case of the dynamic, but most of the ideas can be applied for inhomogeneous case as well. We first came up with an upper bound for the termination time as a function of number of the agents n, bound of confidence ϵ and diameter of the initial profile D, conditioned that consensus is reached, then we generalized it for the arbitrary case. Furthermore, one can get an upper bound explicitly as a function of n by using $D \leq n\epsilon$.

CHAPTER 3

POLYNOMIAL UPPER BOUND FOR THE TERMINATION TIME OF THE HEGSELMANN-KRAUSE MODEL

In this section, we again consider the scalar case of the Hegselmann-Krause model on the real line. The same as before, let us assume n to be the number of agents. Once again, we assume that the agents are sorted in an increasing order. Also, without loss of generality we assume that the initial profile is positive, i.e. $x(0) = (x_1(0), x_2(0), \ldots, x_n(0))$. Otherwise, we can translate all the agents by a positive constant to move them to the positive side of the real line. It is not hard to see that this translation does not change the dynamic. Moreover, let us show the dynamic by in its close form the same as following:

$$x(t+1) = A(t)x(t), \ \forall t \geq 0, \tag{3.1}$$

where,

$$A_{ij}(t) = \begin{cases} \frac{1}{|N_i(x(t))|}, & \text{if } j \in N_i(x(t)) \\ 0, & \text{else} \end{cases}.$$

We next discuss the adjoint dynamics for the Hegselmann- Krause dynamics. It has been shown in [16] and [17] that, for the Hegselmann-Krause dynamics, there exists a sequence of stochastic vectors $\{\pi(t)\}, t \geq 0$, backward in time, such that:

$$\pi'(t+1)A(t) = \pi'(t). \tag{3.2}$$

This dynamics is the adjoint for the original dynamics (3.1). Furthermore, It can be seen easily from (3.1) and (3.2) that:

$$\pi'(t+1)x(t+1) = \pi'(t)x(t), \forall t \geq 0. \tag{3.3}$$

Our analysis in this section uses a Lyapunov comparison function that can be constructed by using the adjoint dynamics [16]. First, we use the Entropy function to construct such a comparison, and in the next section we improve the results slightly by using a quadratic comparison function.

3.1 Entropy Comparison Function

Based on the adjoint dynamics which was introduced in (3.2), one can define:

$$H_{\pi(t)}(x(t)) = -\sum_{i=1}^{n} \pi_i(t) \ln(x_i(t)), \; t \geq 0. \tag{3.4}$$

Note that since we already assumed that the agents are sorted in the increasing order, thus $0 < x_1(0) \leq x_1(1) \leq x_1(2) \leq \ldots$. Therefore, all the profiles at each time step will remain positive and the above function is well defined at each time step.

Lemma 3.1. *Suppose that $\{\pi(t)\}_{t\geq 0}$ is a sequence of stochastic vectors which satisfies in the adjoint dynamics (3.2), then,*

$$H_{\pi(t+1)}(x(t+1)) \leq H_{\pi(t)}(x(t)).$$

Proof.

$$\begin{aligned} H_{\pi(t+1)}(x(t+1)) - H_{\pi(t)}(x(t)) &= \ln[\frac{\Pi_{i=1}^{n} x_i(t)^{\pi_i(t)}}{\Pi_{i=1}^{n} x_i(t+1)^{\pi_i(t+1)}}] \\ &= \ln[\frac{\Pi_{i=1}^{n} x_i(t)^{\pi_i(t)}}{\Pi_{i=1}^{n} (A(t)x(t))_i^{\pi_i(t+1)}}]. \end{aligned} \tag{3.5}$$

On the over side, by using Jensen inequality (1.9), we have:

$$
\begin{aligned}
(A(t)x(t))_i^{\pi_i(t+1)} = \Big(\sum_{j=1}^{n} A_{ij}(t)x_j(t)\Big)^{\pi_i(t+1)} &\geq \Big(\Pi_{j=1}^{n} x_j(t)^{A_{ij}(t)}\Big)^{\pi_i(t+1)} \\
\Rightarrow \Pi_{i=1}^{n}(A(t)x(t))_i^{\pi_i(t+1)} &\geq \Pi_{i=1}^{n}\Big(\Pi_{j=1}^{n} x_j(t)^{A_{ij}(t)}\Big)^{\pi_i(t+1)} \\
&= \Pi_{i=1}^{n}\Big(\Pi_{j=1}^{n} x_j(t)^{A_{ij}(t)\pi_i(t+1)}\Big). \qquad (3.6)
\end{aligned}
$$

Now, by using (3.2) in (3.6) we get:

$$
\begin{aligned}
(A(t)x(t))_i^{\pi_i(t+1)} &\geq \Pi_{i=1}^{n}\Big(\Pi_{j=1}^{n} x_j(t)^{A_{ij}(t)\pi_i(t+1)}\Big) \\
&= \Pi_{j=1}^{n} x_j(t)^{\sum_{i=1}^{n} A_{ij}(t)\pi_i(t+1)} = \Pi_{i=1}^{n} x_i(t)^{\pi_i(t)}. \qquad (3.7)
\end{aligned}
$$

Finally, applying (3.7) in (3.5) gives us the proper result. **Q.E.D.**

The above lemma tells us that the amount of Entropy function is non-increasing by passing the time and thus, a natural question is to figure out how large the amount of decrease would be. Therefore, In the next, our goal is to find a lower bound on the amount of decrease of the Entropy Function. We begin our analysis by the following definition.

Definition 3. *We say that a merging time happens at the time instant t if there are at least two agents i and j such that $x_i(t-1) \neq x_j(t-1)$ and $x_i(t) = x_j(t)$.*

The following lemma tells us that if there is no merging time at some time instant, then, the disjoint agents can not be so close to each other. In other words, if the agents are sufficiently close to each other, then we will have a merging time in the next iteration.

Theorem 3.1. *Consider the Hegselmann-Krause model with n agents and the bound of confidence $\epsilon > 0$. Also, let us show the t-th ordered profile by $x(t)$. Then, we have either t is a merging time or every two separated agents in $x(t)$ have a distance at least $\frac{\epsilon}{n^2}$.*

Proof. Let us assume that t is not a merging time. Choose an arbitrary agent i and assume that $x_{i+1}(t) \neq x_i(t)$ to be one of her closest neighbors. Since t is not a merging time, therefore, $N_i(t) \neq N_{i+1}(t)$. Otherwise, in the next iteration agents i and $i+1$ will merge.

Also, assume that $|N_{i+1}(t) \setminus N_i(t)| = x$, $|N_i(t) \setminus N_{i+1}(t)| = y$ and $|N_i(t) \cap N_{i+1}(t)| = m$. Furthermore, let $A = \sum_{j \in N_{i+1}(k) \cap N_i(k)} x_j(k)$ and $x_{i+1}(k) - x_i(k) = d < \epsilon$. Therefore, we notice that:

- $x_i(t) + \epsilon \leq x_j(t) < x_{i+1}(t) + \epsilon, \ \forall j \in N_{i+1}(t) \setminus N_i(t)$
- $x_i(t) - \epsilon < x_j(t) \leq x_{i+1}(t) - \epsilon, \ \forall j \in N_i(t) \setminus N_{i+1}(t)$.

Then, by the definition of Hegselmann-Krause model and using two above inequalities, one can write:

$$\begin{aligned} \frac{A + x(x_{i+1}(t) - \epsilon)}{m + x} \leq x_i(t+1) \leq \frac{A + x(x_{i+1}(t) - \epsilon)}{m + x} \\ \frac{A + y(x_i(t) - \epsilon)}{m + y} \leq x_{i+1}(t+1) \leq \frac{A + y(x_{i+1}(t) + \epsilon)}{m + y}. \end{aligned} \tag{3.8}$$

Moreover, we can get the following range of changes for A.

$$\begin{aligned} A &\leq (m-2)(x_i(t) + \epsilon) + x_i(t) + x_{i+1}(t) \\ &= (m-2)(x_i(t) + \epsilon) + x_i(t) + x_i(t) + d \\ &= m x_i(t) + d + (m-2)\epsilon. \end{aligned}$$

$$\begin{aligned} A &\geq (m-2)(x_{i+1}(t) - \epsilon) + x_i(t) + x_{i+1}(t) \\ &= (m-2)(x_i(t) + d - \epsilon) + x_i(t) + x_i(t) + d \\ &= m x_i(t) + d + (m-2)(d - \epsilon), \end{aligned}$$

or, simply we can write:

$$A - m x_i(k) \in [d + (m-2)(d - \epsilon), \ d + (m-2)\epsilon]. \tag{3.9}$$

Now we show that $x_{i+1}(t) - x_i(t) \geq \frac{\epsilon}{n^2}$. For this purpose, let us define:

$$\begin{aligned} f(x,y) &= \frac{A + y(x_i(t) - \epsilon)}{m+y} - \frac{A + x(x_{i+1}(t) - \epsilon)}{m+x} \\ &= \frac{(x-y)(A - mx_i(t)) + m(x+y)\epsilon - mdx + xy(2\epsilon - d)}{(m+x)(m+y)}, \end{aligned}$$

where, in the last equality we have used $x_{i+1}(t) = x_i(t) + d$. By using (3.8), it is clear that $x_{i+1}(t) - x_i(t) \geq f(x,y)$. We consider two different cases:

- Case 1: $x > y$ and $(A - mx_i(t)) \leq 0$.

Since $(A - mx_i(t)) \leq 0$ by using (3.9) we get $d + (m-2)(d-\epsilon) \leq (A - mx_i(t)) \leq 0$ and therefore, $(x-y)[d + (m-2)(d-\epsilon)] \leq (x-y)(A - mx_i(t))$. Thus, we can write:

$$\begin{aligned} f(x,y) &= \frac{(x-y)(A - mx_i(t)) + m(x+y)\epsilon - mdx + xy(2\epsilon - d)}{(m+x)(m+y)} \\ &\geq \frac{(x-y)[d + (m-2)(d-\epsilon)] + m(x+y)\epsilon - mdx + xy(2\epsilon - d)}{(m+x)(m+y)} \\ &= \frac{2x\epsilon + 2(m-1)y\epsilon - dx + 2xy\epsilon - (m-1)yd - dxy}{(m+x)(m+y)} \\ &= \frac{(2\epsilon - d)(x + xy + (m-1)y)}{(m+x)(m+y)} \geq \frac{(2\epsilon - d)(x-y)}{(m+x)(m+y)} \\ &\geq \epsilon(\frac{1}{(m+y)} - \frac{1}{(m+x)}) > \frac{\epsilon}{n^2}. \end{aligned}$$

Where in the last inequality we have used this fact that $m + x + y \leq n$.

- Case 2: $y \geq x$ and $(A - mx_i(t)) \geq 0$

Again, by using (3.9) we have $d + (m-2)\epsilon \geq (A - mx_i(t)) \geq 0$ and therefore, $(x-y)[d +$

$(m-2)\epsilon] \leq (x-y)(A-mx_i(t))$. Thus, we can write:

$$\begin{aligned} f(x,y) &= \frac{(x-y)(A-mx_i(t))+m(x+y)\epsilon-mdx+xy(2\epsilon-d)}{(m+x)(m+y)} \\ &\geq \frac{(x-y)[d+(m-2)\epsilon]+m(x+y)\epsilon-mdx+xy(2\epsilon-d)}{(m+x)(m+y)} \\ &= \frac{2y\epsilon+2(m-1)x\epsilon-dx+2xy\epsilon-(m-1)xd-dxy}{(m+x)(m+y)} \\ &= \frac{(2\epsilon-d)(y+xy+(m-1)x)}{(m+x)(m+y)} \geq \frac{(2\epsilon-d)(y-x)}{(m+x)(m+y)} \\ &\geq \epsilon(\frac{1}{(m+x)}-\frac{1}{(m+y)}) > \frac{\epsilon}{n^2}. \end{aligned}$$

Note that if non of the above cases happens, then, $f(x,y)$ is even larger than the previous cases, because in the nominator of $f(x,y)$ we are adding one extra positive term and therefore the results still hold. This completes the proof. **Q.E.D.**

Lemma 3.2. *Suppose that $f(\cdot)$ is a strongly convex function with parameter μ and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a stochastic vector, then:*

$$f(\sum_{i=1}^{n}\alpha_i x_i) \leq \sum_{j=1}^{n}\alpha_i f(x_i) - \frac{\mu}{2}\sum_{j=0}^{n-2}\frac{\alpha_{n-j}}{(\sum_{\ell=1}^{n-j}\alpha_\ell)(\sum_{\ell=1}^{n-j-1}\alpha_\ell)}\Big\|\sum_{i=1}^{n-j-1}\alpha_i\Big(x_{n-j}-x_i\Big)\Big\|^2.$$

Proof. We start from the left hand side of the above inequality and with an inductively argument extract the terms of the right hand side. By applying the strong convexity of f

we have:

$$
\begin{aligned}
f(\sum_{i=1}^{n}\alpha_i x_i) &= f(\alpha_n x_n + (1-\alpha_n)(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i)) \\
&\leq \alpha_n f(x_n) + (1-\alpha_n) f(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i) \\
&- \frac{\mu}{2}\alpha_n(1-\alpha_n)\|x_n - \frac{\sum_{i=1}^{n-1}\alpha_i x_i}{1-\alpha_n}\|^2 \\
&= \alpha_n f(x_n) + (1-\alpha_n) f(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i) \\
&- \frac{\mu}{2}\frac{\alpha_n}{1-\alpha_n}\|\sum_{i=1}^{n-1}\alpha_i(x_n - x_i)\|^2. \qquad (3.10)
\end{aligned}
$$

Since $\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n} = 1$, therefore, we can apply (3.10) on $f(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i)$ and we get:

$$
\begin{aligned}
f(\sum_{i=1}^{n-1}\frac{\alpha_i}{1-\alpha_n}x_i) &\leq \frac{\alpha_{n-1}}{1-\alpha_n}f(x_{n-1}) + (1-\frac{\alpha_{n-1}}{1-\alpha_n})f(\sum_{i=1}^{n-2}\frac{\alpha_i}{1-\alpha_n-\alpha_{n-1}}x_i) \\
&- \frac{\mu}{2}\frac{\frac{\alpha_{n-1}}{1-\alpha_n}}{1-\frac{\alpha_{n-1}}{1-\alpha_n}}\|\sum_{i=1}^{n-2}\frac{\alpha_i}{1-\alpha_n}(x_{n-1}-x_i)\|^2 = \frac{\alpha_{n-1}}{1-\alpha_n}f(x_{n-1}) \\
&+ \frac{(1-\alpha_n-\alpha_{n-1})}{1-\alpha_n}f(\sum_{i=1}^{n-2}\frac{\alpha_i}{1-\alpha_n-\alpha_{n-1}}x_i) \\
&- \frac{\alpha_{n-1}}{(1-\alpha_n-\alpha_{n-1})}\|\sum_{i=1}^{n-2}\frac{\alpha_i}{1-\alpha_n}(x_{n-1}-x_i)\|^2. \qquad (3.11)
\end{aligned}
$$

By replacing (3.11) in (3.10) and repeating this process we get:

$$
\begin{aligned}
f(\sum_{i=1}^{n}\alpha_i x_i) \leq \sum_{i=1}^{n}\alpha_i f(x_i) &- \frac{\mu}{2}\Big[\frac{\alpha_n}{1-\alpha_n}\|\sum_{i=1}^{n-1}\alpha_i(x_n-x_i)\|^2 \\
&+ \frac{\alpha_{n-1}(1-\alpha_n)}{(1-\alpha_n-\alpha_{n-1})}\|\sum_{i=1}^{n-2}\frac{\alpha_i}{1-\alpha_n}(x_{n-1}-x_i)\|^2 \\
&+ \frac{\alpha_{n-2}(1-\alpha_n-\alpha_{n-1})}{(1-\alpha_n-\alpha_{n-1}-\alpha_{n-2})}\|\sum_{i=1}^{n-3}\frac{\alpha_i}{1-\alpha_n-\alpha_{n-1}}(x_{n-2}-x_i)\|^2 + \dots\Big],
\end{aligned}
$$

which can be written in the form of lemma (3.2). **Q.E.D.**

Now, we have enough tools to find a lower bound on the decreasing amount of the Entropy function.

Theorem 3.2. *In the Hegselmann-Krause model we have:*

$$H_{\pi(t)}(x(t)) - H_{\pi(t+1)}(x(t+1)) \geq \frac{\mu}{2|N_k(t)|^2} \sum_{i=1}^{n} \pi_i(t+1) d_i(t)^2,$$

where, $d_i(t) = \max\{\|x_r(t) - x_s(t)\| : r, s \in N_i(t)\}$ *and* μ *is a positive constant.*

Proof. First of all, note that we can rescale the initial profile and assume that all the agents are in $[\theta, 1]$ for some $\theta > 0$. Also, we know that $-\ln(\cdot)$ is a strongly convex function on $[\theta, 1]$. Once again, let us show the parameter of the strong convexity by μ. According to (3.1), for an arbitrary $k \in \{1, \dots n\}$ we can write $x_k(t+1) = A'_k(t)x(t)$. Since $A_k(t)$ is a stochastic vector, hence, we can apply lemma (3.2) for $A(t)'_k x(t)$. Let us define $\delta_k(t)$ to be:

$$\delta_k(t) = \frac{\mu}{2} \sum_{j=0}^{n-2} \frac{A(t)_{k(n-j)}}{(\sum_{\ell=1}^{n-j} A(t)_{k\ell})(\sum_{\ell=1}^{n-j-1} A(t)_{k\ell})} \Big\| \sum_{i=1}^{n-j-1} A(t)_{ki} \Big(x_{n-j}(t) - x_i(t) \Big) \Big\|^2. \tag{3.12}$$

Let us show the smallest and the largest indices in $N_j(k)$ by r and s, respectively. Therefore, by the definition of $A(t)$ we have:

$$A_{k\ell}(t) = \begin{cases} \frac{1}{|N_k(t)|}, & \text{if } \ell \in \{r, r+1, \dots, s\} \\ 0, & \text{else} \end{cases}. \tag{3.13}$$

By using (3.13) in (3.12) and simplifying we get:

$$\begin{aligned} \delta_k(t) &= \frac{\mu}{2|N_k(t)|} \sum_{j=n-s}^{n-r-1} \frac{1}{(n-j-r)(n-j-r+1)} \Big\| \sum_{i=r}^{n-j-1} (x_{n-j}(t) - x_i(t)) \Big\|^2 \\ &= \frac{\mu}{2|N_k(t)|} \sum_{m=r+1}^{S} \frac{1}{(m-r)(m-r+1)} \Big\| \sum_{i=r}^{m-1} (x_m(t) - x_i(t)) \Big\|^2 \\ &\geq \frac{\mu}{2|N_k(t)|^3} \sum_{m=r+1}^{S} \Big\| \sum_{i=r}^{m-1} (x_m(t) - x_i(t)) \Big\|^2 \geq \frac{\mu}{2|N_i(k)|^3} \Big(|N_k(t)| d_k^2(t) \Big) \\ &= \frac{\mu}{2|N_i(k)|^2} d_k^2(t). \end{aligned}$$

Therefore,

$$\delta_k(t) \geq \frac{\mu}{2|N_i(k)|^2} d_k^2(t). \tag{3.14}$$

On the other side:

$$\begin{aligned}
H_{\pi(t)}(x(t)) - H_{\pi(t+1)}(x(t+1)) &= \sum_{k=1}^{n} \pi_k(t+1) \ln(x_k(t+1)) - \sum_{k=1}^{n} \pi_k(t) \ln(x_k(t)) \\
&= \sum_{k=1}^{n} \pi_k(t+1) \ln(A'_k(t)x(t)) - \sum_{k=1}^{n} \pi_k(t) \ln(x_k(t)) \\
&\geq \sum_{k=1}^{n} \Big(\pi_k(t+1) [\sum_{j=1}^{n} A_{kj}(t) \ln(x_j(t)) + \delta_k(t)] \Big) - \sum_{k=1}^{n} \pi_k(t) \ln(x_k(t)) \\
&= \sum_{j=1}^{n} \Big(\sum_{k=1}^{n} \pi_k(t+1) A_{kj}(t) \Big) \ln(x_j(t)) + \\
&+ \sum_{k=1}^{n} \pi_k(t+1) \delta_k(t) - \sum_{k=1}^{n} \pi_k(t) \ln(x_k(t)) \\
&= \sum_{j=1}^{n} \pi_j(t) \ln(x_j(t))) + \sum_{k=1}^{n} \pi_k(t+1) \delta_k(t) - \sum_{k=1}^{n} \pi_k(t) \ln(x_k(t)) \\
&= \sum_{k=1}^{n} \pi_k(t+1) \delta_k(t), \qquad (3.15)
\end{aligned}$$

where, in the second last equality we have used (3.2). Finally by using (3.14) in (3.15) we get:

$$H_{\pi(t)}(x(t)) - H_{\pi(t+1)}(x(t+1)) \geq \sum_{k=1}^{n} \pi_k(t+1) \delta_k(t) \geq \frac{\mu}{2|N_i(k)|^2} \sum_{k=1}^{n} \pi_k(t+1) d_k^2(t).$$

Q.E.D.

3.2 Polynomial Upper Bound by Using Entropy Function

Theorem 3.3. *The Hegselmann-Krause dynamics in one dimension reaches to its steady state no more than*

$$-\frac{2n^6}{\mu\epsilon^2}\ln(x_1(0)) + n$$

steps, where, x(0) denotes the initial profile and μ is a positive constant.

Proof. By taking summation on the relation given in theorem (3.2), we obtain:

$$\begin{aligned} H_{\pi(0)}(x(0)) - H_{\pi(T)}(x(T)) &\geq \sum_{t=0}^{T-1} \frac{\mu}{2|N_k(t)|^2} \sum_{i=1}^{n} \pi_i(t+1) d_i(t)^2 \\ &\geq \frac{\mu}{2n^2} \sum_{t=0}^{T-1} \sum_{i=1}^{n} \pi_i(t+1) d_i(t)^2. \end{aligned}$$

Now, by using theorem (3.1) we know that if at time instant t we don't have any merging, then, $d_i(t) \geq \frac{\epsilon}{n^2}$. Hence if we don't have any merging time in the time interval $[0, T]$, then,

$$\begin{aligned} H_{\pi(0)}(x(0)) - H_{\pi(T)}(x(T)) \geq \frac{\mu\epsilon^2}{2n^6} \sum_{t=0}^{T-1} \sum_{i=1}^{n} \pi_i(t+1) = \frac{\mu\epsilon^2}{2n^6} T \\ \Rightarrow H_{\pi(0)}(x(0)) - \frac{\mu\epsilon^2}{2n^6} T \geq H_{\pi(T)}(x(T)) \geq 0 \\ \Rightarrow T \leq \frac{2n^6}{\mu\epsilon^2} H_{\pi(0)}(x(0)) \leq -\frac{2n^6}{\mu\epsilon^2} \ln(x_1(0)). \end{aligned}$$

Furthermore, since the total number of merging times can not exceed the number of agents (n), hence, $-\frac{2n^6}{\mu\epsilon^2}\ln(x_1(0)) + n$ is an upper bound for the termination time. **Q.E.D.**

3.3 Quadratic Comparison Function

In this part, we use another type of comparison functions in our analysis to come up with an improvement upper bound on the termination time of the Hegselmann-Krause model. This comparison function $V(\cdot)$ which is constructed by using the adjoint dynamics in (3.2) is defined by:

$$V(t) = \sum_{i=1}^{n} \pi_i(t)(x_i(t) - \pi'(t)x(t))^2.$$

This function was introduced earlier in [17] and [16]. In particular, the decrease of this comparison function plays a critical role in our analysis. The following result which was shown in [16] gives us an exact and also a lower bound on the amount of decrease at each time step.

Theorem 3.4. *For any $t \geq 0$, we have*

$$V(t) - V(t+1) = \frac{1}{2}\sum_{i,j=1}^{n} H_{ij}(t)(x_i(t) - x_j(t))^2 \geq \frac{1}{2n}\sum_{i=1}^{n} \frac{\pi_i(t+1)d_i^2(t)}{|N_i(t)|},$$

where $H(t) = A'(t)diag(\pi(t+1))A(t)$ and $d_i(t) = \max\{|x_p(t) - x_q(t)| : p, q \in N_i(t)\}$.

To start our analysis, we state some lemmas which we are going to use them through our main theorem.

Lemma 3.3. *Suppose that t is not a merging time. Then, for every arbitrary agent i such that $|N_i(t)| > 1$ (it means that agent i is not singleton at time t), we have:*

$$\pi_i(t+1) \leq \sum_{\ell \in N_i(t) \setminus \{i\}} \pi_\ell(t+1).$$

Proof. let $x_r(t)$ and $x_s(t)$ to be the neighbors right before and after $x_i(t)$ which are different from it, i.e. $x_r(t) \neq x_i(t)$ and $x_r(t) \neq x_i(t)$. It may happen that one of them doesn't exist, but since we assumed that t is not a merging time, therefore, the other one has to exist. Also, since the order of agents will be preserved during the dynamics and since t is not a merging time, therefore, in the next time step $t+1$, these agents will still remain as the neighbors right before and after $x_i(t+1)$. This shows that $N_i(t+1) \subseteq N_r(t+1) \cup N_s(t+1) \subseteq$

$\cup_{\ell \in N_i(t)\setminus\{i\}} N_\ell(t+1)$. Now by using the definition of adjoint dynamics (3.2) we can write:

$$\begin{aligned}
\pi_i(t+1) = \pi'(t+2)A^i(t+1) &= \sum_{j \in N_i(t+1)} \frac{\pi_j(t+2)}{|N_j(t+1)|} \\
&\leq \sum_{j \in N_r(t+1) \cup N_s(t+1)} \frac{\pi_j(t+2)}{|N_j(t+1)|} \\
&\leq \sum_{j \in N_r(t+1)} \frac{\pi_j(t+2)}{|N_j(t+1)|} + \sum_{j \in N_s(t+1)} \frac{\pi_j(t+2)}{|N_j(t+1)|} \\
&= \pi_r(t+1) + \pi_s(t+1) \leq \sum_{\ell \in N_i(t)\setminus\{i\}} \pi_\ell(t+1).
\end{aligned}$$

Q.E.D.

Lemma 3.4. *Suppose that t is not a merging time. Moreover, assume that for two different p and q we have $d_p(t) < \frac{\epsilon}{2}$ and $d_q(t) < \frac{\epsilon}{2}$, then, $N_p(t) \cap N_q(t) = \emptyset$.*

Proof. Suppose that the lemma is not true. Therefore, there exists at least a $r \in N_p(t) \cap N_q(t)$. But in this case:

$$|x_p(t) - x_q(t)| \leq |x_p(t) - x_r(t)| + |x_r(t) - x_q(t)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This mean that $q \in N_p(t)$. Since we know that $d_p(t) < \frac{\epsilon}{2}$, therefore, $|x_p(t) - x_q(t)| \leq \frac{\epsilon}{2}$. Now, for every agent $j \in N_q(t)$ we have:

$$|x_p(t) - x_j(t)| \leq |x_p(t) - x_q(t)| + |x_q(t) - x_j(t)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This shows that $j \in N_p(t)$ and hence, $N_q(t) \subseteq N_p(t)$. With a similar argument $N_p(t) \subseteq N_q(t)$. Therefore, $N_p(t) = N_q(t)$. This shows that we will have a merging time at time t which is a contradiction. **Q.E.D.**

3.4 Polynomial Upper Bound by Using Quadratic Function

Theorem 3.5. *Let us assume that t_i denotes the number of time steps between the $(i-1)$th and ith merging times. Then,*

$$V(\sum_{i=1}^{n-1} t_i) \leq V(0) - \frac{\epsilon^2}{8n} \sum_{i=1}^{n-1} \frac{t_i}{i+1}.$$

Proof. First of all note that since the number of merging times cab be at most $n-1$, therefore, $t_n = t_{n+1} = \ldots = 0$. Pick an arbitrary time $t \geq 0$ such that it is not a merging time. Therefore, there exists $k \in \{1, 2, \ldots, n-2\}$ such that $\sum_{i=1}^{k} t_i \leq t < \sum_{i=1}^{k+1} t_i$. We will show that the amount of decrease in $V(\cdot)$ between time steps t and $t+1$ is at least $\frac{\epsilon^2}{nk}$.

Let us define $T = \{\ell : d_\ell(t) < \frac{\epsilon}{2}\}$ and $A = \cup_{\ell \in T} N_\ell(t)$. Also, let A^c to be the complement of A. As a result of lemma (3.4) we can rewrite A as the union of some disjoint sets, i.e. $A = \cup_{j=1}^{m} N_{\ell_j}(t)$ for some m and indices ℓ_j. Now we have:

$$\begin{aligned} V(t) - V(t+1) &\geq \frac{1}{2n} \sum_{i=1}^{n} \pi_i(t+1) d_i^2(t) \\ &= \frac{1}{2n} \sum_{i \in A} \pi_i(t+1) d_i^2(t) + \frac{1}{2n} \sum_{i \in A^c} \pi_i(t+1) d_i^2(t) \\ &\geq \frac{1}{2n} \sum_{i \in A} \pi_i(t+1) d_i^2(t) + \frac{\epsilon^2}{8n} \sum_{i \in A^c} \pi_i(t+1). \end{aligned} \tag{3.16}$$

Next, we will find a lower bound for $\frac{1}{2n} \sum_{i \in A} \pi_i(t+1) d_i^2(t)$. For this purpose, we define $N_{\ell_j}^0 = \{\ell : x_\ell(t) = x_{\ell_j}(t)\}$. Now, it can bee seen that:

$$\begin{aligned} \frac{1}{2n} \sum_{i \in A} \pi_i(t+1) d_i^2(t) &= \frac{1}{2n} \sum_{j=1}^{m} \sum_{i \in N_{\ell_j}} \pi_i(t+1) d_i^2(t) \\ &= \frac{1}{2n} \sum_{j=1}^{m} \Big(\sum_{i \in N_{\ell_j} \setminus N_{\ell_j}^0} \pi_i(t+1) d_i^2(t) + \sum_{i \in N_{\ell_j}^0} \pi_i(t+1) d_i^2(t) \Big) \\ &\geq \frac{1}{2n} \sum_{j=1}^{m} \Big(\frac{\epsilon^2}{4} \sum_{i \in N_{\ell_j} \setminus N_{\ell_j}^0} \pi_i(t+1) + \sum_{i \in N_{\ell_j}^0} \pi_i(t+1) d_i^2(t) \Big), \end{aligned} \tag{3.17}$$

where in the last inequality we have used lemma this property that if $i \in N_{\ell_j} \setminus N^0_{\ell_j}$, then $d_i(t) \geq \frac{\epsilon}{2}$. Otherwise, the same as what we showed in the proof of lemma (3.4) we will have a merging time at time t which is a contradiction.

Furthermore, since we assumed that $\sum_{i=1}^{k} t_i \leq t < \sum_{i=1}^{k+1} t_i$, therefore, we had k merging time before and hence, the number of agents which could be equal to $x_{\ell_j}(t)$ is at most k. In other words, $|N^0_{\ell_j}| \leq k$. By using (3.17) we can write:

$$\begin{aligned}
\frac{1}{2n}\sum_{i\in A}\pi_i(t+1)d_i^2(t) &\geq \frac{1}{2n}\sum_{j=1}^{m}\Big(\frac{\epsilon^2}{4}\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)+\sum_{i\in N^0_{\ell_j}}\pi_i(t+1)d_i^2(t)\Big)\\
&\geq \frac{1}{2n}\sum_{j=1}^{m}\Big(\frac{\epsilon^2}{4}\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\Big)\\
&= \frac{1}{2n}\sum_{j=1}^{m}\Big(\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}|N^0_{\ell_j}|\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\\
&+\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\Big)\\
&\geq \frac{1}{2n}\sum_{j=1}^{m}\Big(\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}|N^0_{\ell_j}|\pi_{\ell_j}(t+1)+\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\Big)\\
&= \frac{1}{2n}\sum_{j=1}^{m}\Big(\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}\sum_{i\in N^0_{\ell_j}}\pi_i(t+1)+\frac{\epsilon^2}{4(|N^0_{\ell_j}|+1)}\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\Big),
\end{aligned}$$

where in the last inequality we have used lemma (3.3). Also, in the last equality we used this fact that if two agent have the same opinion at time t, then they will have the same $\pi_i(t+1)$. Finally, by replacing all the $|N^0_{\ell_j}|$ by a larger amount k in the last equality of the above relation, we get:

$$\begin{aligned}
\frac{1}{2n}\sum_{i\in A}\pi_i(t+1)d_i^2(t) &\geq \frac{\epsilon^2}{8n(k+1)}\sum_{j=1}^{m}\Big(\sum_{i\in N^0_{\ell_j}}\pi_i(t+1)+\sum_{i\in N_{\ell_j}\setminus N^0_{\ell_j}}\pi_i(t+1)\Big)\\
&= \frac{\epsilon^2}{8n(k+1)}\sum_{j=1}^{m}\sum_{i\in N_{\ell_j}}\pi_i(t+1)\\
&= \frac{\epsilon^2}{8n(k+1)}\sum_{i\in A}\pi_i(t+1) \qquad (3.18)
\end{aligned}$$

Therefore, by using (3.18) in (3.16) we get:

$$\begin{aligned}
V(t)-V(t+1) &\geq \frac{\epsilon^2}{8n(k+1)}\sum_{i\in A}\pi_i(t+1)+\frac{\epsilon^2}{8n}\sum_{i\in A^c}\pi_i(t+1)\\
&\geq \frac{\epsilon^2}{8n(k+1)}\Big(\sum_{i\in A}\pi_i(t+1)+\sum_{i\in A^c}\pi_i(t+1)\Big)\\
&= \frac{\epsilon^2}{8n(k+1)}.
\end{aligned}$$

Therefore, we have shown that for every t in the time interval $[\sum_{i=1}^{k}t_i,\sum_{i=1}^{k+1}t_i)$,

$$V(t)-V(t+1)\geq \frac{\epsilon^2}{8n(k+1)}.$$

By taking summation of the above inequality over $[0,\sum_{i=1}^{n-1}t_i]$ we get:

$$V(\sum_{i=1}^{n-1}t_i)\leq V(0)-\frac{\epsilon^2}{8n}\sum_{i=1}^{n-1}\frac{t_i}{i+1}.$$

Q.E.D.

As a result of theorem (3.5), we always have to have:

$$\begin{aligned}
V(0) &\geq \frac{\epsilon^2}{8n}\sum_{i=1}^{n-1}\frac{t_i}{i+1}\\
&\Rightarrow \sum_{i=1}^{n-1}\frac{t_i}{i+1}\leq\frac{8nV(0)}{\epsilon^2}\leq 8n^3 \qquad (3.19)
\end{aligned}$$

where in the above inequality we used this fact that $V(0) \leq n^2\epsilon^2$. Therefore, we have the following corollary.

Corollary 3.1. *In the Hegselmann-Krause model with n agents, the first merging time takes no more than $16n^3$ steps.*

Proof. It is a direct result of relation (3.19). **Q.E.D.**

Now, If we let T_n to be the termination time for the Hegselmann-Krause dynamics, according to the definition of t_i, we can write $T_n = \sum_{i=1}^{n-1} t_i$. Therefore, we are interested to find an upper bound for $T_n = \sum_{i=1}^{n-1} t_i$ under the constraint given in (3.19). In other words, the worst upper bound is given by:

$$\max \quad T_n = \sum_{i=1}^{n-1} t_i$$
$$\text{s.t} \quad \sum_{i=1}^{n-1} \frac{t_i}{i+1} \leq 8n^3,$$

which has a solution $T_n \leq 8n^3$. Therefore, we have the following theorem.

Theorem 3.6. *In the Hegselmann-Krause model with n agents the termination time is bounded from above by:*

$$T_n \leq 8n^4.$$

CHAPTER 4

TERMINATION TIME OF THE HEGSELMANN-KRAUSE MODEL IN HIGHER DIMENSIONS

In this section, we consider the Hegselmann-Krause model in higher dimensions. Once again, the same as the scalar case, we have a set of agents and a bound of confidence ($\epsilon > 0$), which in this chapter we assume to be the same for all the agents (homogenous case). In this model at each time instant, all of the agents update their value by taking the average of their own value and all the other agents which are in their ϵ-neighborhood. But this time, each agents has a vector as her opinion. For instance, if we assume that the dimension is d, each agents' opinion is a vector in R^d. It has been proven before that as in the scalar case, the dynamics will reach its steady state after finite time [17]. This motivates us to consider these dynamics and to look for an upper bound for its termination time. Therefore, our goal in this chapter is to get an upper bound for the termination time as a function of number of agents (n) and bound of confidence (ϵ). But before we start our analysis we are going to illustrate some of the differences between the dynamics in higher dimension and the dynamics in scalar case and show that these differences make our analysis much more complicated than the scalar case.

4.1 Properties and Numerical Analysis

We saw in the previous chapter that when all the agents take values on the real line, running the dynamics will not change their order. In other words, the dynamics preserves the order of the agents. In fact, this is not the case in higher dimensions because in higher dimensions each opinion is a vector and therefore we cannot consider an order for them anymore.

Another important property that was used frequently in the scalar case was that if there exists a break between two agents, then, this break will remain until the end. But this property doesn't hold in higher dimensions any more. In other words, it may happen that an agent is isolated from a group of agents by a distance more than ϵ, but after some time this agent will get back to the group again. Figure 4.1 shows an example of this type of situation in two dimensions. Note that the agent in the middle is isolated from the rest of the agents, but as the dynamics run, the other agents will get closer and closer to each other such that after some time they will capture the isolated. Figure 4.2 is an other example in three dimensions after running the dynamics 10 times and 50 times (steady state), respectively. Note that in Figure 4.2, the red points show the current location of agents after 10 and 50 iterations, respectively. Now that we have some intuition about how bad the situation can be, we are going to start our analysis. Our analysis in this chapter is again based on the Lyapunov comparison functions. Once again, in this chapter we use almost the same quadratic Lyapunov function which is defined by

$$V(t) = \sum_{i=1}^{n} \pi_i(t) \|x_i(t) - \pi'(t)X(t)\|^2, \tag{4.1}$$

. The only difference here is that instead of absolute value we use Euclidian norm. According to [17] we have the following result.

Lemma 4.1. *For the quadratic comparison function defined in (4.1) we have:*

$$V(t) - V(t+1) = \frac{1}{2} \sum_{i,j=1}^{n} H_{ij}(t) \|x_i(t) - x_j(t)\|^2 \geq \frac{1}{2n} \sum_{i=1}^{n} \frac{\pi_i(t+1) d_i^2(t)}{|N_i(t)|},$$

where $H(t) = A'(t) diag(\pi(t+1)) A(t)$ *and* $d_i(t) = \max\{\|x_p(t) - x_q(t)\| : p, q \in N_i(t)\}$.

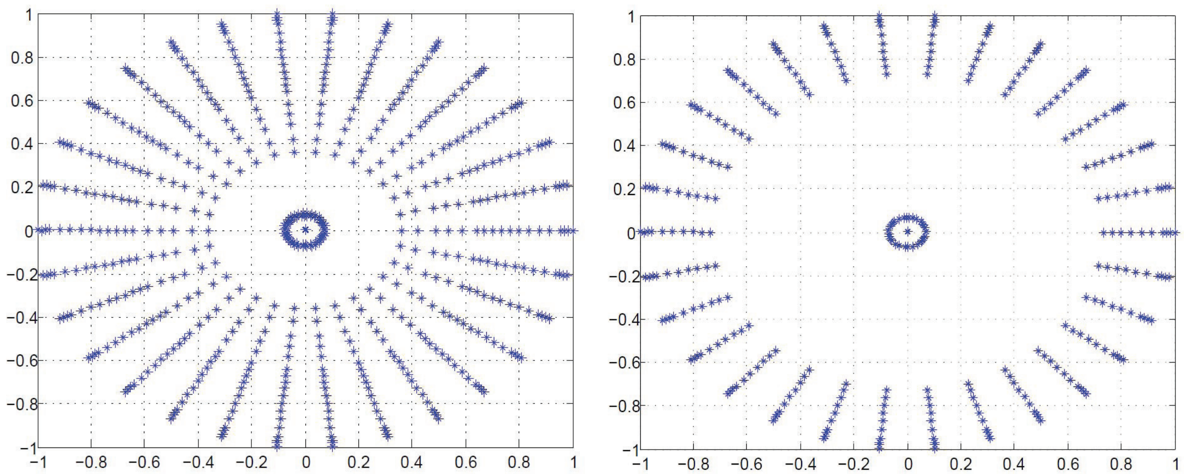

Figure 4.1: 2D, n=30. In the left Figure $\epsilon = 0.4$, t=100 and in the right Figure $\epsilon = 0.55$, t=10.

4.2 Definitions and Basic Lemmas

Here, we list some of the notation which we use to prove some of the results in this section.

Definition 4. *Suppose that we have a bound of confidence ϵ and a set of n agents. At each time instant $t \geq 0$ we define the following sets to be:*

- $N_i(t) = \{j : \|x_j(t) - x_i(t)\| < \epsilon\}$,
- $d_i(t) = \max\{\|x_p(t) - x_q(t)\| : p, q \in N_i(t)\}$,
- $S_2(t) = \{i : d_i(t) \geq \frac{\epsilon}{n}\}$,
- $S_1(t) = \{i : 0 \neq d_i(t) < \frac{\epsilon}{n}\}$,
- $S_0(t) = \{i : d_i(t) = 0\}$.

Lemma 4.2. *If $d_i(t) \leq \frac{\epsilon}{n}$ for some time instant t, then, $N_i(t) \subseteq N_i(t+1)$.*

Proof. Consider an arbitrary element $j \in N_i(t)$. We show that $j \in N_i(t+1)$. First, note that since $d_i(t) \leq \frac{\epsilon}{n}$, thus $N_i(t) \subseteq N_j(t)$. Now let us assume that $N_i(t) \setminus N_j(t)$ has r elements and we show these agents by $x_1(t), x_2(t), \ldots, x_r(t)$. Furthermore, assume $|N_i(t)| = m$ and we

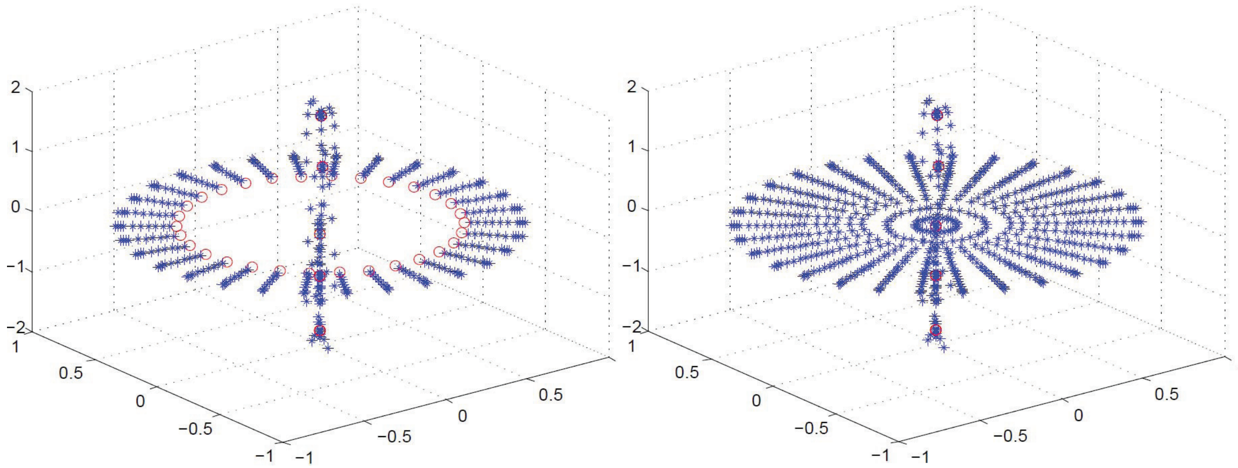

Figure 4.2: 3D, n=60, $\epsilon = 0.4$. In the left Figure t=10 and in the right Figure t=50.

show these agents by $x_i(t), x_j(t), x_{r+1}(t), \ldots, x_{r+m-2}(t)$. Therefore, by applying the dynamic rule we get:

$$x_i(t+1) = \frac{x_i(t) + x_j(t) + \sum_{k=r+1}^{r+m-2} x_k(t)}{m},$$
$$x_j(t+1) = \frac{x_i(t) + x_j(t) + \sum_{k=1}^{r+m-2} x_k(t)}{m+r}.$$

Now, we claim that $\|x_j(t+1) - x_i(t+1)\| < \epsilon$ which shows that $j \in N_i(t+1)$. To show our claim, we have:

$$\begin{aligned}\|x_j(t+1) - x_i(t+1)\| &= \frac{\|r(x_i(t) + x_j(t) + \sum_{k=r+1}^{r+m-2} x_k(t)) - m\sum_{k=1}^{r} x_k(t)\|}{m(m+r)} \\ &= \frac{1}{m(m+r)}\Big\| \sum_{\ell=1}^{r}(x_i(t) - x_\ell(t)) + \sum_{\ell=1}^{r}(x_j(t) - x_\ell(t)) + \\ &+ \sum_{k=r+1}^{r+m-2}\sum_{\ell=1}^{r}(x_p(t) - x_q(t)) \Big\|. \end{aligned} \tag{4.2}$$

On the other side, we have:

$$\sum_{\ell=1}^{r} \|(x_i(t) - x_\ell(t))\| \leq r(\epsilon + \frac{\epsilon}{n}),$$
$$\sum_{\ell=1}^{r} \|(x_j(t) - x_\ell(t))\| < \epsilon,$$
$$\sum_{k=r+1}^{r+m-2} \sum_{\ell=1}^{r} \|(x_p(t) - x_q(t))\| \leq (m-2)r(\epsilon + \frac{2\epsilon}{n}). \tag{4.3}$$

Therefore, by applying the triangle inequality on the right hand side of (4.2) and using the above inequalities (4.3), we obtain:

$$\begin{aligned}\|x_j(t+1) - x_i(t+1)\| &\leq \frac{1}{m(m+r)}[r(\epsilon + \frac{\epsilon}{n}) + r\epsilon + (m-2)r(\epsilon + \frac{2\epsilon}{n})] \\ &= [\frac{rm + \frac{r}{n}(1 + \frac{2(m-2)}{n})}{rm + m^2}]\epsilon < \epsilon.\end{aligned}$$

Q.E.D.

In the next few lemmas, we will see some of the relations among the sets S_0, S_1 and S_2.

Lemma 4.3. *Suppose that $\ell_0 \in S_0(t)$ and $\ell_0 \notin S_0(t+1)$, then, $\ell_0 \in S_2(t+1)$.*

Proof. Since by assumption $\ell_0 \in S_0(t)$, thus, $x_{\ell_0}(t+1) = x_{\ell_0}(t)$. Moreover, since $\ell_0 \notin S_0(t+1)$, hence, it has at least one neighbor at time $t+1$. Let us choose one of them arbitrarily and show it by ℓ_1. Now we can write:

$$\begin{aligned}\|x_{\ell_1}(t+1) - x_{\ell_0}(t+1)\| = \|x_{\ell_1}(t+1) - x_{\ell_0}(t)\| &= \|\frac{\sum_{k \in N_{\ell_1}(t)} x_k(t)}{|N_{\ell_1}(t)|} - x_{\ell_0}(t)\| \\ &= \frac{\|\sum_{k \in N_{\ell_1}(t)}(x_k(t) - x_{\ell_0}(t))\|}{|N_{\ell_1}(t)|} \\ &\geq \frac{\|x_{\ell_1}(t) - x_{\ell_0}(t)\|}{|N_{\ell_1}(t)|} > \frac{\epsilon}{|N_{\ell_1}(t)|} \geq \frac{\epsilon}{n}.\end{aligned}$$

where, in the second last inequality we have used this fact that since $\ell_0 \in S_0(t)$, thus, all of its neighbors have a distance more than ϵ from it. This shows that $d_{\ell_0}(t+1) \geq \frac{\epsilon}{n}$, i.e. $\ell_0 \in S_2(t+1)$. **Q.E.D.**

Lemma 4.4. *Suppose that there is no merging at time step t. Then, for every $j \in S_1(t)$ we have:*

$$N_j(t+1) \subset \{S_0(t) \cap S_2(t+1)\} \cup \{\cup_{i \in S_2(t)} N_i(t+1)\}. \tag{4.4}$$

Proof. Assume that $r \in N_j(t+1)$, then, there are two possibilities:

- Case 1)

$r \in S_0(t)$ and at time $t+1$ we have $r \in N_j(t+1)$. It means that $r \notin S_0(t+1)$. Hence, by using lemma (4.3) we can see that $r \in S_2(t+1)$ and therefore, $r \in \{S_0(t) \cap S_2(t+1)\}$. Thus, in this case r appears in both sides and the lemma (4.4) holds.

- Case 2)

$r \notin S_0(t)$. Therefore, $r \in S_1(t) \cup S_2(t)$. If $r \in S_2(t)$, again, the case is trivial and it can be seen that r appears in the right hand side of (4.4). But, if $r \in S_1(t)$, since we don't have any merging at time step t, therefore, for every $q \in N_r(t)$, $q \neq r$ we have $q \in S_2(t)$. Furthermore, by using lemma (4.2) we see that q will remain as a neighbor of r at time $t+1$. It means:

$$q \in N_r(t+1), \quad q \in S_2(t) \Rightarrow r \in N_q(t+1), \quad q \in S_2(t) \Rightarrow r \in \cup_{i \in S_2(t)} N_i(t+1).$$

Therefore, we showed that in both cases r would be in the right hand side of (4.4). This completes the proof. **Q.E.D.**

In the rest of this chapter, the same as scalar case, we assume that $\{\pi(t)\}_{t \geq 0}$ is a sequence of stochastic vectors which are defined backward in time by using the adjoint dynamic (3.2). The following lemma gives us a necessary tool to evaluate the amount of decrease in our quadratic comparison function.

Theorem 4.1. *Suppose that there is no merging at time step t, then,*

$$\sum_{j \in S_1(t)} \pi_j(t+1) \leq |S_1(t)| \Big(\sum_{i \in S_2(t)} \pi_i(t+1) + \sum_{i \in S_2(t+1)} \pi_i(t+2) \Big).$$

Proof. By using lemma (4.4) we can write:

$$
\begin{aligned}
\pi_j(t+1) &= \sum_{k \in N_j(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} \\
&\leq \sum_{k \in S_0(t) \cap S_2(t+1) \cup_{i \in S_2(t)} N_i(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} \\
&\leq \sum_{k \in S_0(t) \cap S_2(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} + \sum_{k \in \cup_{i \in S_2(t)} N_i(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} \\
&\leq \sum_{k \in S_0(t) \cap S_2(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} + \sum_{i \in S_2(t)} \sum_{k \in N_i(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} \\
&= \sum_{k \in S_0(t) \cap S_2(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} + \sum_{i \in S_2(t)} \pi_i(t+1), \qquad (4.5)
\end{aligned}
$$

where in the last equality we have used the property of adjoint dynamics (3.2). On the other side we have:

$$
\sum_{k \in S_0(t) \cap S_2(t+1)} \pi_k(t+2) \geq \sum_{k \in S_0(t) \cap S_2(t+1)} \frac{\pi_k(t+2)}{|N_k(t+1)|} \qquad (4.6)
$$

By combining (4.5) and (4.6) we can write:

$$
\begin{aligned}
\pi_j(t+1) &\leq \sum_{i \in S_2(t)} \pi_i(t+1) + \sum_{i \in S_0(t) \cap S_2(t+1)} \pi_i(t+2) \\
&\leq \sum_{i \in S_2(t)} \pi_i(t+1) + \sum_{i \in S_2(t+1)} \pi_i(t+2).
\end{aligned}
$$

Since the above inequality is true for every arbitrary $j \in S_1(t)$, therefore, by taking summation over all the $j \in S_1(t)$, we can write:

$$
\sum_{j \in S_1(t)} \pi_j(t+1) \leq |S_1(t)| \Big(\sum_{i \in S_2(t)} \pi_i(t+1) + \sum_{i \in S_2(t+1)} \pi_i(t+2) \Big). \qquad (4.7)
$$

Q.E.D.

In order to analyze the Hegselmann-Krause dynamics, once again, we consider the quadratic

Lyapunov function defined by:

$$V(t) = \sum_{i=1}^{n} \pi_i(t) \|x_i(t) - \pi'(t)X(t)\|^2, \tag{4.8}$$

where, $X(t) = (x_1(t), x_2(t), \ldots, x_n(t))'$. We have seen before that:

$$\begin{aligned} V(t) - V(t+1) &= \sum_{i=1}^{n} \frac{\pi_i(t+1)}{|N_i(t)|^2} \sum_{p,q \in N_i(t)} \|x_p(t) - x_q(t)\|^2 \\ &\geq \frac{1}{2} \sum_{i=1}^{n} \frac{\pi_i(t+1)}{|N_i(t)|} d_i^2(t) \geq \frac{1}{2n} \sum_{i=1}^{n} \pi_i(t+1) d_i^2(t). \end{aligned} \tag{4.9}$$

On the other side, one can write:

$$\begin{aligned} \sum_{i=1}^{n} \pi_i(t+1) d_i^2(t) &\geq \frac{\epsilon^2}{n^2} \sum_{i \in S_2(t)} \pi_i(t+1) + 0 \sum_{i \in S_0(t)} \pi_i(t+1) + \sum_{i \in S_1(t)} \pi_i(t+1) d_i^2(t) \\ &\geq \frac{\epsilon^2}{n^2} \sum_{i \in S_2(t)} \pi_i(t+1). \end{aligned} \tag{4.10}$$

Therefore, by combining (4.9) and (4.10) together we have the following corollary.

Corollary 4.1. *Suppose that $V(t)$ is the Lyapunov function defined in (4.8), then,*

$$V(t) - V(t+1) \geq \frac{\epsilon^2}{2n^3} \sum_{i \in S_2(t)} \pi_i(t+1).$$

4.3 Estimation on the Comparison Function Decrease

In the rest of this chapter, our goal is to find a lower bound for $\sum_{i \in S_2(t)} \pi_i(t+1)$ and thus for $V(t) - V(t+1)$.

Theorem 4.2. *Let $T \geq 2$ to be an integer and t_0 to be an arbitrary time step such that there is no merging time in the time interval $[t_0, t_0 + T]$. Also, assume that $n \geq 2$, then, we have:*

$$V(t_0) - V(t_0 + T) \geq \frac{\epsilon^2}{4n^4} \Big(\sum_{j=0}^{T-2} (1 - \sum_{i \in S_0(t_0+j)} \pi_i(t_0 + j)) \Big).$$

Proof. We prove the theorem by induction. Choose $t \in [t_0, T-2]$ and $T \geq 2$ arbitrarily and then fix them. By using the above corollary two times and adding them together, we get:

$$
\begin{aligned}
V(t)-V(t+2) &\geq \frac{\epsilon^2}{2n^3}\Big(\sum_{i\in S_2(t)}\pi_i(t+1)+\sum_{i\in S_2(t+1)}\pi_i(t+2)\Big)\\
&= \frac{\epsilon^2}{2n^3(1+|S_1(t)|)}\Big(\sum_{i\in S_2(t)}\pi_i(t+1)+\sum_{i\in S_2(t+1)}\pi_i(t+2)+\\
&\quad + |S_1(t)|[\sum_{i\in S_2(t)}\pi_i(t+1)+\sum_{i\in S_2(t+1)}\pi_i(t+2)]\Big)\\
&\geq \frac{\epsilon^2}{2n^3(1+|S_1(t)|)}\Big(\sum_{i\in S_2(t)}\pi_i(t+1)+\sum_{i\in S_2(t+1)}\pi_i(t+2)+\sum_{i\in S_1(t)}\pi_i(t+1)\Big)\\
&\geq \frac{\epsilon^2}{2n^4}\Big((1-\sum_{i\in S_0(t)}\pi_i(t+1))+\sum_{i\in S_2(t+1)}\pi_i(t+2)\Big),
\end{aligned}
$$

where, in the second last inequality we have used theorem (4.1). Moreover, the last inequality holds because of stochastisity of the vector $\pi(t+1)$ and also since $(1+|S_1(t)|) \leq n$. Therefore, we have shown so far that:

$$V(t)-V(t+2) \geq \frac{\epsilon^2}{2n^4}(1-\sum_{i\in S_0(t)}\pi_i(t+1)).$$

By taking summation on the above inequality, we get:

$$
\begin{aligned}
2(V(t_0)-V(t_0+T)) &\geq \Big(V(t_0)-V(t_0+T-1)\Big)+\Big(V(t_0+1)-V(t_0+T)\Big)\\
&= \sum_{t=t_0}^{t_0+T-2}(V(t)-V(t+2))\\
&\geq \sum_{t=t_0}^{t_0+T-2}\frac{\epsilon^2}{2n^4}(1-\sum_{i\in S_0(t)}\pi_i(t+1))\\
&= \frac{\epsilon^2}{2n^4}\Big(\sum_{j=0}^{T-2}(1-\sum_{i\in S_0(t_0+j)}\pi_i(t_0+j+1))\Big). \qquad (4.11)
\end{aligned}
$$

Also, we know that if $r \in S_0(t+j)$, then, $\pi_r(t+j) = \pi_r(t+j+1)$. Therefore, by using this fact in (4.11) we get:

$$V(t_0) - V(t_0+T) \geq \frac{\epsilon^2}{4n^4}\Big(\sum_{j=0}^{T-2}(1 - \sum_{i\in S_0(t_0+j)} \pi_i(t_0+j))\Big).$$

Q.E.D.

4.4 An Upper Bound for the Termination Time

Now we are ready to prove the main theorem of this chapter.

Theorem 4.3. *Let $T \geq 2$ to be an integer and t to be an arbitrary time step such that there is no merging time in the time interval $[t, t+T]$. Then,*

$$V(t) - V(t+T) \geq \frac{\epsilon^2}{2n^4}\Big(1 - \sum_{i\in\bigcap_{t'=t}^{t+T-2} S_0(t')} \pi_i(t)\Big).$$

Proof. By using theorem (4.2), one can write:

$$\begin{aligned} V(t) - V(t+T) &\geq \frac{\epsilon^2}{4n^4}\Big(\sum_{j=0}^{T-2}(1 - \sum_{i\in S_0(t+j)} \pi_i(t+j))\Big) \\ &= \frac{\epsilon^2}{4n^4}\Big(1 - \sum_{i\in\bigcap_{t'=t}^{t+T-2} S_0(t')} \pi_i(t)\Big) \\ &+ \frac{\epsilon^2}{4n^4}\Big(\sum_{j=1}^{T-2}(1 - \sum_{i\in S_0(t+j)} \pi_i(t+j)) - \sum_{i\in S_0(t)\setminus\bigcap_{t'=t}^{t+T-2} S_0(t')} \pi_i(t)\Big). \end{aligned} \tag{4.12}$$

Now, we show that the second term in right hand side of the above inequality is always non-negative. Let us assume that $\ell \in S_0(t) \setminus \bigcap_{t'=t}^{t+T-2} S_0(t')$. It means that there exists a $1 \leq j \leq T-2$, such that $\ell \in \bigcap_{t'=t}^{t+j-1} S_0(t'), \ell \notin S_0(t+j)$. Since $\ell \in \bigcap_{t'=t}^{t+j-1} S_0(t')$, thus,

$$\pi_\ell(t) = \pi_\ell(t+1) = \ldots = \pi_\ell(t+j-1) = \pi_\ell(t+j). \tag{4.13}$$

Also, since $\ell \notin S_0(t+j)$, therefore, $(1-\sum_{k\in S_0(t+j)} \pi_k(t+j))$ which is equal to $\sum_{k\in S_1(t+j)\cup S_2(t+j)} \pi_k(t+$

j) has $\pi_\ell(t+j)$ in its body. But note that by (4.13), $\pi_\ell(t+j) = \pi_\ell(t)$, hence, we can say that $(1-\sum_{k\in S_0(t+j)} \pi_k(t+j))$ has $\pi_\ell(t)$ in its body. Therefore, with putting all together, we have shown so far that for every $\ell \in S_0(t) \setminus \bigcap_{t'=t}^{t+T-2} S_0(t')$, $\pi_\ell(t)$ appears in at least one of the terms of the $\sum_{j=1}^{T-2}(1-\sum_{i\in S_0(t+j)} \pi_i(t+j))$ and of course they don't have overlap, because they have different indices. In other words:

$$\sum_{j=1}^{T-2}(1-\sum_{i\in S_0(t+j)} \pi_i(t+j)) \geq \sum_{i\in S_0(t)\setminus \bigcap_{t'=t}^{t+T-2} S_0(t')} \pi_i(t). \tag{4.14}$$

Using (4.14) in (4.12) gives us the proper result. **Q.E.D.**

Here, it is important to note that there is not only way of choosing the sequence of $\{\pi(t)\}_{t\geq 0}$. In other words, to construct the sequence of $\{\pi(t)\}_{t\geq 0}$, it is enough to choose an arbitrary stochastic vector $\pi(T_n)$ at the termination time T_n and then the rest of sequence will be constructed backward in time by $\pi'(t)A(t) = \pi'(t-1)$. One of the strong aspects of theorem (4.3) is that it holds true, independent of the choice of $\{\pi(t)\}_{t\geq 0}$. However, the way that the dynamics acts is independent of the choices of this stochastic sequence. But, in fact a smart choice of $\{\pi(t)\}_{t=0}^{T_n}$ will lead us to a better approximation of the decrease in $V(\cdot)$ function and hence a better estimation of termination time based on our analysis. Now, let us define:

$$\mu_n = \max\{T : \bigcap_{t'=t}^{t+T-2} S_0(t') \neq \emptyset, \forall t = 0, \ldots, T_n\}. \tag{4.15}$$

From the above definition it can be seen that:

$$\bigcap_{t'=t}^{t+\mu_n} S_0(t') = \emptyset, \quad \forall t = 0, \ldots, T_n. \tag{4.16}$$

For instant a naive upper bound for μ_n would be T_{n-1}. Because, as long as $\bigcap_{t'=t}^{t+T-2} S_0(t') \neq \emptyset$, then, the number of agents which are not in this intersection is at most $n-1$. But, since a set of $(n-1)$ agents can move at most T_{n-1} times, therefore, we have $\mu_n \leq T_{n-1}$.

Now, by using theorem (4.3) and relation (4.16) we can write:

$$\begin{aligned} V(0) - V(n^6\mu_n) &= \sum_{j=0}^{n^6-1} \Big(V(j\mu_n) - V((j+1)\mu_n) \Big) \\ &\geq \sum_{j=0}^{n^6-1} \frac{\epsilon^2}{4n^4} [1 - \sum_{i \in \bigcap_{t'=j\mu_n}^{(j+1)\mu_n - 2} S_0(t')} \pi_i(t)] \\ &= \sum_{j=0}^{n^6-1} \frac{\epsilon^2}{4n^4} [1 - 0] = \frac{1}{4} n^2 \epsilon^2. \end{aligned}$$

On the other side, a simple calculation shows that $V(0) \leq \frac{1}{4}n^2\epsilon^2$, hence, it means that $V(n^6\mu_n) \leq 0$. Therefore, we have the following theorem:

Theorem 4.4. *Suppose that T_n is the termination time for the Hegselmann-Krause model in multi dimensions with n agents, then,*

$$T_n \leq n^6 \mu_n,$$

where, μ_n was defined in (4.15).

From the above theorem and using the naive approximation of $\mu_n \leq T_{n-1}$, we get:

$$\begin{aligned} T_n \leq n^6 T_{n-1} \leq n^6 (n-1)^6 T_{n-2} \leq \ldots \leq (n!)^6 T_1 = (n!)^6 \\ \Rightarrow \qquad T_n \leq (n!)^6. \end{aligned}$$

Corollary 4.2. *The termination time for the Hegselmann-Krause model in multi dimensions with n agents is bounded by*

$$T_n \leq (n!)^6.$$

4.5 On the Continuous-Opinion Hegselmann-Krause Model

In this section we consider the continuous case of the Hegselmann-Krause dynamics which was introduced earlier in [18]. In this model the density of the agents is suppose to be a continuous index set I, while the time step t is discrete. Also, in this model, it is assumed that the opinion profile on the agent set I at tth time step is a function which is usually denoted by x_t. According to this notation the continuous form of the Hegselmann-Krause is defined by:

$$x_{t+1}(\alpha) = \frac{\int_{\beta:(\alpha,\beta)\in C_{x_t}} x_t(\beta)d(\beta)}{\int_{\beta:(\alpha,\beta)\in C_{x_t}} d(\beta)}, \tag{4.17}$$

where,

$$C_{x_t} = \{(\alpha, \beta) \in I^2 : |x_t(\alpha) - x_t(\beta)| \leq 1\}.$$

Definition 5. *For a measure μ and a measurable function x_t and set S, we define:*

$$\mu_{x_t}(S) = \mu(\{b : x_t(b) \in S\}).$$

Our goal in section is to show that applying the dynamics (4.17) on a particular profile will give us more smooth opinion profiles. In other words, we show that applying the dynamics (4.17) will not increase the points of discontinuity.

Lemma 4.5. *Suppose that x_t is a bounded function on the interval $[0, 1]$ such that $\mu_{x_t}(a) = \mu(\{b \in [0, 1] : x_t(b) = a\}) = 0$ for every a in the range of x_t. Then, if we show the points of continuity of x_t by D_t, we will have: $D_t \subseteq D_{t+1}$.*

Proof. Suppose $\alpha \in D_t$. Also, assume $\{\alpha_n\}_{n=1}^{\infty}$ is an arbitrary sequence which converges to

α. We will show that:

$$\lim_{n\to\infty}\int_{\beta:(\alpha_n,\beta)\in C_{x_t}} d(\beta) = \int_{\beta:(\alpha,\beta)\in C_{x_t}} d(\beta),$$
$$\lim_{n\to\infty}\int_{\beta:(\alpha_n,\beta)\in C_{x_t}} x_t(\beta)d(\beta) = \int_{\beta:(\alpha,\beta)\in C_{x_t}} x_t(\beta)d(\beta). \tag{4.18}$$

For this purpose, we partition $\{\alpha_n\}_{n=1}^{\infty}$ to two sub-sequences $\{\alpha_n^1\}_{n=1}^{\infty}$ and $\{\alpha_n^2\}_{n=1}^{\infty}$ such that $x_t(\alpha_n^1) \leq x_t(\alpha)$ and $x_t(\alpha_n^2) > x_t(\alpha)$. Since we know that x_t is continuous at α, therefore, $\lim_{n\to\infty} x_t(\alpha_n^1) = \lim_{n\to\infty} x_t(\alpha_n^2) = x_t(\alpha)$. We next show

$$\lim_{n\to\infty}\int_{\beta:(\alpha_n^1,\beta)\in C_{x_t}} d(\beta) = \int_{\beta:(\alpha,\beta)\in C_{x_t}} d(\beta), \tag{4.19}$$

and

$$\lim_{n\to\infty}\int_{\beta:(\alpha_n^1,\beta)\in C_{x_t}} x_t(\beta)d(\beta) = \int_{\beta:(\alpha,\beta)\in C_{x_t}} x_t(\beta)d(\beta). \tag{4.20}$$

Choose a sub-sequence of $\{\alpha_n^1\}_{n=1}^{\infty}$ and let us show that by $\{\alpha_{i_n}^1\}_{n=1}^{\infty}$ such that $x_t(\alpha_{i_n}^1) \leq x_t(\alpha_{i_{n+1}}^1)$. Note that choosing such a subsequence is possible because $x_t(\alpha_n^1) \leq x_t(\alpha)$ and also $\lim_{n\to\infty} x_t(\alpha_n^1) = x_t(\alpha)$. Now, we have:

$$\begin{aligned}
&|\int_{\beta:(\alpha_{i_n}^1,\beta)\in C_{x_t}} d(\beta) - \int_{\beta:(\alpha,\beta)\in C_{x_t}} d(\beta)| \\
&\leq \int_{\left\{\beta:\ x_t(\beta)\in(x_t(\alpha_{i_n}^1)-1,x_t(\alpha)-1]\right\}\cup\left\{\beta:\ x_t(\beta)\in[x_t(\alpha_{i_n}^1)+1,x_t(\alpha)+1)\right\}} d(\beta) \\
&\leq \mu_{x_t}\Big((x_t(\alpha_{i_n}^1)-1, x_t(\alpha)-1]\Big) + \mu_{x_t}\Big([x_t(\alpha_{i_n}^1)+1, x_t(\alpha)+1)\Big).
\end{aligned} \tag{4.21}$$

But we note that $\mu_{x_t}((x_t(\alpha_{i_n}^1)-1, x_t(\alpha)-1]) \leq \mu[0,1] = 1$ and also $(x_t(\alpha_{i_{n+1}}^1)-1, x_t(\alpha)-1] \subseteq (x_t(\alpha_{i_n}^1)-1, x_t(\alpha)-1]$, $\forall n = 1,2,\ldots$, therefore, by continuity of the measure (1.6), we get:

$$\lim_{n\to\infty} \mu_{x_t}\Big((x_t(\alpha^1_{i_n})-1, x_t(\alpha)-1]\Big) = \mu_{x_t}\Big(\cap_{n=1}^{\infty}(x_t(\alpha^1_{i_n})-1, x_t(\alpha)-1]\Big)$$
$$= \mu_{x_t}(x_t(\alpha)-1) = 0,$$

where in the last equality we have used the lemma assumption. Similarly we can see that $\lim_{n\to\infty}\mu_{x_t}([x_t(\alpha^1_{i_n})+1, x_t(\alpha)+1)) = 0$. Now fix an arbitrary $\epsilon > 0$. According to what we have shown so far, there exists $M > 0$ such that if $n > M$, the amount of relation (4.21) is less than ϵ. Also, since $\lim_{n\to\infty} x_t(\alpha^1_n) = x_t(\alpha)$ thus, there exists a N such that if $n > N$ then $x_t(\alpha^1_n)$ is sufficiently close to $x_t(\alpha)$ and hence, $x_t(\alpha^1_{i_M}) \le x_t(\alpha^1_n) \le x_t(\alpha)$. Thus,

$$\Big\{\beta : x_t(\beta) \in (x_t(\alpha^1_n)-1, x_t(\alpha)-1]\Big\} \subseteq \Big\{\beta : x_t(\beta) \in (x_t(\alpha^1_{i_M})-1, x_t(\alpha)-1]\Big\},$$
$$\Big\{\beta : x_t(\beta) \in [x_t(\alpha^1_n)+1, x_t(\alpha)+1)\Big\} \subseteq \Big\{\beta : x_t(\beta) \in [x_t(\alpha^1_{i_M})+1, x_t(\alpha)+1)\Big\}. \tag{4.22}$$

Therefore, if $n > \max\{M, N\}$ by combining (4.21) and (4.22) we can write:

$$\begin{aligned}
&|\int_{\beta:(\alpha^1_n,\beta)\in C_{x_t}} d(\beta) - \int_{\beta:(\alpha,\beta)\in C_{x_t}} d(\beta)| \\
&\le \int_{\big\{\beta:(\alpha^1_n,\beta)\in C_{x_t}\big\}\Delta\big\{\beta:(\alpha,\beta)\in C_{x_t}\big\}} d(\beta) \\
&\le \int_{\big\{\beta:\ x_t(\beta)\in(x_t(\alpha^1_n)-1,x_t(\alpha)-1]\big\}\cup\big\{\beta:\ x_t(\beta)\in[x_t(\alpha^1_n)+1,x_t(\alpha)+1)\big\}} d(\beta) \\
&\le \int_{\big\{\beta:\ x_t(\beta)\in(x_t(\alpha^1_{i_M})-1,x_t(\alpha)-1]\big\}\cup\big\{\beta:\ x_t(\beta)\in[x_t(\alpha^1_{i_M})+1,x_t(\alpha)+1)\big\}} d(\beta) \\
&\le \mu_{x_t}\Big((x_t(\alpha^1_{i_n})-1, x_t(\alpha)-1]\Big) + \mu_{x_t}\Big([x_t(\alpha^1_{i_n})+1, x_t(\alpha)+1)\Big) < \epsilon.
\end{aligned}$$

This proves (4.19).

On the other side, since x_t is bounded, therefore there is a K such that $|x_t| \le K$. Now by a similar argument, for every $n > \max\{M, N\}$ we have :

$$\begin{aligned}
&\Big| \int_{\beta:(\alpha_n^1,\beta)\in C_{x_t}} x_t(\beta)d(\beta) - \int_{\beta:(\alpha,\beta)\in C_{x_t}} x_t(\beta)d(\beta)\Big| \\
&\leq \int_{\left\{\beta:(\alpha_n^1,\beta)\in C_{x_t}\right\}\Delta\left\{\beta:(\alpha,\beta)\in C_{x_t}\right\}} |x_t(\beta)|d(\beta) \\
&\leq K \int_{\left\{\beta:(\alpha_n^1,\beta)\in C_{x_t}\right\}\Delta\left\{\beta:(\alpha,\beta)\in C_{x_t}\right\}} d(\beta) < K\epsilon.
\end{aligned}$$

Since K is a constant and $\epsilon > 0$ was arbitrary, hence, we get (4.20).

Finally, by repeating all the above argument for the sequence $\{\alpha_n^2\}$ we get (4.18). This completes the proof. **Q.E.D.**

CHAPTER 5

NECESSARY CONDITIONS FOR FINITE CONVERGENCE, SUFFICIENT CONDITIONS FOR CONSENSUS

One of the basic questions in averaging dynamics is to answer this question that when and under which conditions the agents can reach to an agreement and this motivate us to design some algorithms which guarantee the consensus. However, sometimes because of shortage of memory or because of some time limitations, not only we are interested in agreement of the agents but also in finiteness of the operation time. This brings us to some sort of questions that under which conditions designing a finite time algorithm is possible.

5.1 Finite Time Consensus

Suppose that we have a network of n nodes such that at each time instant each node updates her value to be the weighted average of her own previous value and those of her neighbors. It has been shown before that under this assumption, there exists an appropriate weight matrix such that reaching to consensus is possible in finitely many steps [19]. In this case the number of iterations can be estimated by the degree of the minimal polynomial of the weight matrix. Furthermore, if we show this degree by D, then every node has to have a memory of size at least D, such that it can update its current value by its previous values. Here, a natural question is that for which kind of networks we can find a weight matrix such that without memory agents are still able to reach an agreement in finitely many steps or equivalently suppose A is a weight matrix for the network, we are interested to know for which kind of networks there exists a m such that A^m is a consensus matrix, i.e. all of its rows are identical.

Lemma 5.1. *Suppose that for a stochastic weight matrix A we have A^m is a consensus matrix, then, it has 1 as a simple eigenvalue and all of its other eigenvalues are zero.*

Proof. Suppose that the stochastic matrix A has this property that there exists a m such that $A^m = (A_1, A_1, \ldots, A_1)'$ where A_1 is a column vector of size n. Consider the Jordan form of the matrix A and suppose $A = B^{-1}JB$. Therefore we have to have $B^{-1}J^mB = (A_1, A_1, \ldots, A_1)'$. Now suppose B_i^{-1} is the i^{th} column of B^{-1} and $\sum B_i$ is the summation of all the elements of the ith row of B. Therefore, a simple computation shows that J^m has the following form:

$$J^m = \begin{pmatrix} (A_1'B_1^{-1})\sum B_1 & (A_1'B_2^{-1})\sum B_1 & \cdots & (A_1'B_n^{-1})\sum B_1 \\ (A_1'B_1^{-1})\sum B_2 & (A_1'B_2^{-1})\sum B_2 & \cdots & (A_1'B_n^{-1})\sum B_2 \\ \vdots & \vdots & \ddots & \vdots \\ (A_1'B_1^{-1})\sum B_n & (A_1'B_2^{-1})\sum B_n & \cdots & (A_1'B_n^{-1})\sum B_n \end{pmatrix}. \tag{5.1}$$

On the other side we know that since J is the Jordan form of the matrix A, therefore, J^m has the following block diagonal form:

$$J^m = \begin{pmatrix} J_1^m & 0 & \cdots & 0 \\ 0 & J_2^m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k^m \end{pmatrix}, \tag{5.2}$$

where $J_1, J_2, \ldots J_k$ are the appropriate Jordan blocks of the matrix A. Also, since we assumed that A is a stochastic matrix, therefore it has 1 as an eigenvalue and thus, in the Jordan representation of A one of the blocks contains this eigenvalue. Without loss of generality we can assume that the first block contains eigenvalue 1. But according to the matrix representation in (5.1), we know that if one element in J^m is zero then either all the elements in the corresponding row or the corresponding column are zero. By comparing (5.1) and (5.2) we can conclude that all the elements in J^m are zeros except the first block which is related to eigenvalue 1. Therefore, J^m has just one nonzero block in its Jordan form which is related to the eigenvalue 1. With a similar argument it can be seen that the size of this block is also one. Therefore, J^m has just one nonzero entry which is 1 and hence, this shows that A has just one simple eigenvalue at 1 and all of its other eigenvalues are zero. **Q.E.D.**

Finiteness of the algorithm sometimes impose hard conditions on the properties of the network. To be more precise, let us consider an averaging gossip algorithm. Suppose that we run a gossip algorithm in a following way that in each time instant, two neighbors nodes i and j wake up and update their information according to the following rule:

$$
\begin{aligned}
x_i(k+1) = x_j(k+1) &= \frac{x_i(k)+x_j(k)}{2} \\
x_r(k+1) = x_r(k), &\ r \neq i,j
\end{aligned} \tag{5.3}
$$

In order to know more about gossip and broadcast algorithms one can refer to [20], [5], [21], [22]. In the next lemma, we show that if the above algorithm terminates in finite time in a connected network, then the number of nodes has to be some power of 2.

Lemma 5.2. *If the gossip algorithm described in (5.3) reaches to its steady state in finite time on a connected network with* n *nodes, then,* $n = 2^k$ *for some* $k \in Z^+$.

Proof. First of all we notice that if the algorithm reaches to its steady state, then, it must be consensus. Otherwise, there are two nodes p and q such that their values are not equal. Since we assumed that the network is connected, therefore, there is a path between these two nodes. But the ending points of this path have different values, therefore, we can find two neighbors in this path such that they have different values. Thus, if these two nodes wake up in some time step, then their value would be different from the previous step and this is in contradiction that we have already reached to the steady state.

Let us assume $x_1(0), x_2(0), \ldots, x_n(0)$ are the initial amounts of the agents and suppose that consensus is reached. According to the updating rule (5.3) the consensus amount has to be of the following form:

$$
x_{ave}(\infty) = \sum_{j=1}^{n} \frac{r_j x_j(0)}{2^{i_j}},
$$

for some non-negative integers $i_j,\ j = 1, \ldots, n$ and some odd numbers $r_j > 0$. On the other side since the average value of the agents at each time step t, $(x_{ave}(t))$ is the same and equal to $\frac{\sum_{j=1}^{n} x_j(0)}{n}$, therefore, we get:

$$x_{ave}(\infty) = \sum_{j=1}^{n} \frac{r_j x_j(0)}{2^{i_j}} = \frac{\sum_{j=1}^{n} x_j(0)}{n}.$$

By comparing the sides of the above equation we get $n = \frac{2^{i_j}}{r_j}$. Furthermore, since n is integer and $r_j,\ j = 1, 2, \ldots, n$ are odd numbers, this shows that $r_1 = r_2 = \ldots = r_n = 1$. Thus, there exists a k such that $i_j = k,\ \forall j = 1, 2, \ldots, n$, i.e. $n = 2^k$. **Q.E.D.**

Definition 6. *A k-cube graph is a graph which has all the o-1 sequences of length k as its nodes and there is a link between two nodes if and only if their corresponding sequences differ only in position.*

Proposition 5.1. *Suppose that a network has 2^k nodes. If the graph of this network has a k-cube as its subgraph, then there is a method which agents can wake up and update their information by (5.3) such that after finite time they all reach to agreement.*

Proof. We will show it by induction. For $n = 2$ the result is trivial. Suppose for $n = 2^{k-1}$ the proposition is correct and let $n = 2^k$. We just focus on the k-cube inside the network and ignore all the other links. By partitioning the nodes to two subsets:

$$A = \{(0, a_1, \ldots, a_{k-1}) |\ a_i \in \{0, 1\}\}$$
$$B = \{(1, a_1, \ldots, a_{k-1}) |\ a_i \in \{0, 1\}\},$$

It can be seen that the induced graphs on each of A and B are $(k-1)$-cube as well. By using the induction assumption, the agents in each of these sub graphs can reach to consensus in finite step. After that in the next time steps we let each node in A talk to her corresponding node in B (a node which is different with it just in the first coordinate). Therefore, it can be seen easily that after additional 2^{k-1} steps all the agents will have the same value. **Q.E.D.**

5.1.1 Sufficient condition for consensus of stochastic chains

In this subsection we are about to answer to this question that under which conditions we can decide about the consensus of the agents with this difference that this time we want to relay more on the configuration of the network than the weights. In other words, existence of a link in network can play a critical role in determining a consensus. We will see that if the number of links are large enough then, consensus is guaranteed under some mild assumptions. We start this part with the following definition:

Definition 7. *Suppose that Q is a non-negative matrix. We say that a permutation - matrix P exists in Q, if $P_{ij} = 1$ results that $Qij > 0$.*

As an example let assume Q to be:

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0.5 \\ 0.2 & 0 & 0.8 \end{pmatrix},$$

In this matrix, there exist two permutation matrices which are:

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

According to the above definition we have the following lemma.

Lemma 5.3. *For every integer n let $D(n) = n!(\sum_{i=0}^{n} \frac{(-1)^i}{i!})$. Then, the summation of every $D(n)$ number of different permutation matrices (n by n) will give us a matrix which has at least one positive column.*

Proof. Let $A_{n\times n}$ be a matrix which is generated by summation of at least $D(n)$ number of permutation matrices. In other words, there are permutation matrices $P_1, P_2, \ldots, P_{D(n)}$ such that $A = \sum_{i=1}^{D_n} P_i$. Since only the positivity of the entries is important, therefore, for easily when we have a positive entry we replace it by 1. Now, Suppose that the lemma is not correct. Thus, we can assume that every column of A includes at least one zero. Among all of these matrices such as A that doesn't satisfy in lemma (5.3), we assume B to be a matrix which has the minimum number of zeros and the maximum number of permutations in it. The first observation is that B has exactly one zero in each column. Otherwise, we can keep one zero in each column of B and change the other zeroes to 1. With this job, we still keep non-positivity of the columns of B and also the number of permutations in B is not decreasing. Also, note that switching the rows with each other and columns with each other doesn't change the number of permutations in B nor the positivity of the columns.

Next, we show that we can assume that each row of B has exactly one zero. Otherwise, by switching the rows and columns we can assume that B has the following representation:

$$B = \begin{pmatrix} 1 & 1 & \dots & 1 & i_1 \text{ zeros} \\ 1 & 1 & \dots & i_2 \text{ zeros} & 1 \\ \vdots & \vdots & \ddots & 1 & 1 \\ i_k \text{ zeros} & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix}. \tag{5.4}$$

Note that in the above matrix each 1 shows a block of ones of appropriate size and $\sum_{\ell=1}^{k} i_\ell = n$. Furthermore, we can assume that the last zero block with length i_k has at least two zeros. Let us show the entries of the last zero block by $b_{k1}, b_{k2}, \dots, b_{ki_k}$. We define matrix $\hat{B}$ based on B (5.4) by:

$$\hat{B}_{rs} = \begin{cases} 1, & (r,s) = (k,1) \\ 0, & (r,s) = (k+1,1) \\ B_{rs}, & \text{else} \end{cases} \tag{5.5}$$

It can be seen that $\hat{B}$ differs from B only in two entries. Also, assume that T and $\hat{T}$ denotes all the permutation matrices which are in B and $\hat{B}$, respectively. Our goal is to show that the cardinality of $\hat{T}$ is not less than the cardinality of T. This shows that instead of considering representation (5.4) for B we can consider $\hat{B}$ (5.5). But, $\hat{B}$ has zero in one more row in compare with B. With repeating this argument we can consider a representation for B which has exactly one zero in each row.

To show that $|T| \leq |\hat{T}|$, we define a one-to-one mapping F from T to $\hat{T}$. Let us Suppose that C is an arbitrary permutation matrix in T such that $C_{(k+1)1} = 1$. Since C is a permutation matrix, thus, $C_{k1} = 0$ and there exists a j such that $C_{kj} = 1$ and $C_{(k+1)j} = 0$. Now we define another permutation matrix $\hat{C}$ to be the same as matrix C with this difference that we switch $C_{kj} = 1$ and $C_{(k+1)j} = 0$ with $C_{k1} = 0$ and $C_{(k+1)1} = 1$, respectively. In other words:

$$\hat{C}_{rs} = \begin{cases} C_{kj}, & (r,s) = (k,1) \\ C_{(k+1)j}, & (r,s) = (k+1,1) \\ C_{k1}, & (r,s) = (k,j) \\ C_{(k+1)1}, & (r,s) = (k+1,j) \\ C_{rs}, & \text{else} \end{cases} \tag{5.6}$$

It can be seen that if $C \in T$, then, $\hat{C} \in \hat{T}$. By defining:

$$F : T \to \hat{T}$$
$$F(C) = \begin{cases} C, & C_{(k+1)1} = 0 \\ \hat{C}, & C_{(k+1)1} = 1 \end{cases},$$

it is not hard to see that F is a one-to-one mapping from T to $\hat{T}$ and this shows that $|\hat{T}| \geq |T|$.

So far we have shown that we can consider B to be a matrix which has exactly one zero in each row and each column. By switching the rows together and columns together we can put all the zeros on the diagonal. Therefore, we have the following configuration for the matrix B:

$$\begin{pmatrix} 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \vdots & 1 & 1 \end{pmatrix}, \tag{5.7}$$

and its not hard to see that the number of permutations which are in (5.7) is equal to $D(n)$. But B was a matrix which had the most number of permutations in it and didn't satisfy to lemma (5.3). Therefore, every matrix with more than $D(n)$ permutation in it has at least one positive column. This completes the proof. **Q.E.D.**

Lemma 5.4. *Suppose that A_1 and A_2 are two subsets of a finite group G. If $1 + \varphi(|A_1|) < |A_2|$ then, $|A_1A_2| = |\{r_1r_2|r_1 \in A_1, r_2 \in A_2\}| > |A_1|$, where φ is the Euler function.*

Proof. Let $A_1 = \{a_1, a_2, \dots, a_k\}$ and $A_2 = \{b_1, b_2, \dots, b_s\}$. It is clear that $|A_1A_2| \geq |A_1b_1| = |\{a_ib_1|a_i \in A_1\}| = k$. To show that $|A_1A_2| > |A_1|$, we will show that $A_1A_2 \neq A_1$. Otherwise for every two arbitrary elements of A_2 such as b_1 and b_2 we have: $\{a_ib_1|a_i \in A_1\} = \{a_ib_2|a_i \in A_1\}$. Without loss of generality we can assume that

$$a_1b_1 = a_2b_2, \quad a_2b_1 = a_3b_2, \quad \dots, \quad a_{k-1}b_1 = a_kb_2, \quad a_kb_1 = a_1b_2,$$

and therefore,

$$a_1^{-1}a_2 = a_2^{-1}a_3 = \dots = a_{k-1}^{-1}a_k = a_k^{-1}a_1 = b_1b_2^{-1} = c, \tag{5.8}$$

which for easily we show all of the above equal elements by c. According to (5.8) we can write:

$$
\begin{aligned}
&a_1^{-1}a_2 = c \\
&a_1^{-1}a_3 = a_1^{-1}a_2a_2^{-1}a_3 = c^2 \\
&a_1^{-1}a_4 = a_1^{-1}a_2a_2^{-1}a_3a_3^{-1}a_4 = c^3 \\
&\vdots \\
&a_1^{-1}a_k = a_1^{-1}a_2a_2^{-1}a_3 \dots a_{k-1}^{-1}a_k = c^{k-1} \\
&a_1^{-1}a_1 = a_1^{-1}a_2a_2^{-1}a_3 \dots a_{k-1}^{-1}a_ka_k^{-1}a_1 = c^k = e
\end{aligned}
$$

Therefore c is an element of order k and according to the above equalities we can write $a_1^{-1}A_1 =< c >$ such that $< c >$ shows the cyclic group with generator c. Because of the symmetry by repeating the same argument we get: $a_1^{-1}A_1 = a_2^{-1}A_1 = \ldots = a_k^{-1}A_1 =< c >$. Also, from (5.8) we had $c = b_1b_2^{-1}$. Since b_1 and b_2 had been chosen arbitrarily, therefore, with the same discussion we can say that all the different elements $b_1b_2^{-1}, b_1b_2^{-1}, \ldots, b_1b_s^{-1}$ have to be a generator for the cyclic group $a_1^{-1}A_1$ of order k. On the other side by lemma (1.5) we know that the number of different generators of a cyclic group of order k is $\varphi(k)$. Therefore, $s-1 \leq \varphi(k)$ or $|A_2| \leq 1+\varphi(|A_1|)$ which is a contradiction. Therefore, $A_1A_2 \neq A_1$ and this completes the proof. **Q.E.D.**

Now let us consider a chain of stochastic matrices, namely $A_1, A_2, A_3, \ldots$. Also, let T_r be the set of permutations which exists in $A(1,r) = A_rA_{r-1}\ldots A_1$. Furthermore, assume that σ_A denotes the set of permutation matrices which exists in A. Therefore, $T_r = \sigma_{A_r}\sigma_{A_{r-1}}\ldots\sigma_{A_1}$. Also, note that $\sigma_{A_j}, \forall j = 1, 2, \ldots$ are subsets of the permutations group, thus, by multiplying subsets of a group the cardinality of them will not decrease and we have $|T_1| \leq |T_2| \leq \ldots \leq |T_r|$. Now, we have enough tools to prove the following theorem.

Theorem 5.2. *Suppose that Σ is a set of n by n stochastic matrices. Also, suppose that there exists a $\delta > 0$ such that* $\min^+ A > \delta \; \forall A \in \Sigma$. *If for every $A \in \Sigma$, we have $|\sigma_A| > 1+\varphi(D(n))$, then every infinite product of the elements from Σ will converge to consensus.*

Proof. For every $r \in N$ we have either $|T_r| > D(n)$ which by lemma (5.3), $A(1,r)$ has at least one positive column or $|T_r| \leq D(n)$. In the second case we have:

$$1 + \varphi(|T_r|) \leq 1 + \varphi(D(n)) < |\sigma_{A_{r+1}}|.$$

Now according to lemma (5.4), $|T_{r+1}| = |T_r \sigma_{A_{r+1}}| > |T_r|$. Therefore, by increasing r, $|T_r|$ will keep increasing until $|T_r| > D(n)$. It means that there exists a sequence of time steps $\{t_s\}, s \in N$ such that $t_{s+1} - t_s < D(n), \forall s \in N$ and $A(t_s, t_{s+1}) = A_{t_s} \dots A_{t_{s+1}}$ has at least one positive column. Also, since the length of time steps is bounded by $D(n)$ and $\min^+ A > \delta$, for all $A \in \Sigma$, therefore, all the conditions of lemma (1.2) are satisfied for $\{A(t_s, t_{s+1})\}_{s=1}^{\infty}$ and this shows that the chain will converge to consensus. **Q.E.D.**

Theorem 5.3. *Suppose that we have a chain of stochastic matrices and a sequence of time steps* $(t_s)_{s \in N}$ *and* $T \in N$ *such that* $t_{s+1} - t_s \leq T \log(\log(s+2))$. *If the number of matrices which satisfy* $|\sigma_A| > \varphi(D(n)) + 1$ *is greater than* $D(n)$ *in every time slot* $[t_s, t_{s+1}]$, *then, the consequent product of the chain will converge to consensus.*

Proof. This theorem is a direct result of the theorem (5.2) and theorem 3.2.37 in [15]. **Q.E.D.**

5.1.2 Sufficient condition for consensus of stochastic chains with prime dimension

We saw before in lemma (5.4) that if A_1 and A_2 are two subsets of a group such that $|A_2| > 1 + \varphi(|A_1|)$, then, $|A_1 A_2| > |A_1|$. What we are going to show next is that when the dimension of the matrices in the chain is a prime number, then, we have additional advantages such that with even less knowledge about the network links we can have the same results. To show that, we first prove another version of the lemma (5.4).

Lemma 5.5. *Suppose that* p *is a prime number and* $< a >$ *is a cyclic group of order* p. *Also, assume that* A_1 *and* A_2 *are two proper subsets of* $< a >$ *such* $|A_1| > 1$ *and* $|A_2| > 1$. *Then,* $|A_2 A_1| > |A_1|$.

Proof. Let $A_1 = \{a^{i_1}, a^{i_2}, \dots, a^{i_k}\}$, where, $0 \leq i_1 < i_2 < \dots < i_k < p$. Also, since $|A_2| > 1$ we can pick two different elements $a^{j_1}, a^{j_2} \in A_2$. Now, if we assume that $|A_2 A_1| \not> |A_1|$, therefore, we must have $|A_2A_1| = |A_1|$. In particular, $a^{j_1}A_1 = a^{j_2}A_1$, i.e.

$$\{a^{i_1+j_1}, a^{i_2+j_1}, \ldots, a^{i_k+j_1}\} = \{a^{i_1+j_2}, a^{i_2+j_2}, \ldots, a^{i_k+j_k}\}.$$

But the above equation is equivalent to show that:

$$\{i_1 + j_1, i_2 + j_1, \ldots, i_k + j_1\} = \{i_1 + j_2, i_2 + j_2, \ldots, i_k + j_2\} \text{ mode } p,$$

which is impossible. Otherwise, we will have:

$$\sum_{\ell=1}^{k}(i_\ell + j_1) = \sum_{\ell=1}^{k}(i_\ell + j_2) \text{ mode } p$$
$$\Rightarrow j_1 = j_2 \text{ mode } p,$$

which is in contradiction with this fact that a^{j_1} and a^{j_2} are two different elements. This shows that $|A_2A_1| > |A_1|$. **Q.E.D.**

Theorem 5.4. *Suppose that p is a prime number. Let us consider a chain of p by p stochastic matrices* $\{A(t)\}_{t=1}^{\infty}$ *such than* $\min^+(A(t)) > \delta > 0$ *for all* $t \in N$*. Also, assume that G is a subgroup of permutations of size p. If every* $A(t)$ *has more than one permutation from G, then, the left product of* $\prod_{t=1}^{\infty} A(t)$ *will converge to a consensus matrix.*

Proof. Let σ_i to be the set of permutations in A_i which lie in G. Note that since G is a cyclic group of order p, therefore, the sum of all the matrices in G will give us a matrix with positive entries. Therefore, if a stochastic matrix has all the elements of G in it, then it has to be a positive matrix. We show that for every j, $A(j, j+p) = A_{j+p}A_{j+p-1}\ldots A_j$ is a scrambling matrix. For this purpose we show that $\sigma_{j+p}\sigma_{j+p-1}\ldots\sigma_j$ is equal to G .

If σ_j is equal to G, then obviously $\sigma_{j+p}\sigma_{j+p-1}\ldots\sigma_j = G$. Otherwise, by using lemma (5.5), $|\sigma_{j+1}\sigma_j| > |\sigma_j|$. Therefore, after every step the number of permutations will increase by at least 1 and hence, $\sigma_{j+p}\sigma_{j+p-1}\ldots\sigma_j$ has cardinality at least equal to p. It means $\sigma_j\sigma_{j+1}\ldots\sigma_{j+p} = G$. Therefore, since $\min^+(A(t)) > \delta > 0$ and also $A(j, j+p)$ is scrambling for all $j \in N$, then lemma (1.2) can be applied. **Q.E.D.**

REFERENCES

[1] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, 1985.

[2] J. Shen, "A geometric approach to ergodic non-homogeneous markov chains," *Wavelet Analysis and multiresolution methods*, pp. 341–366, 2000.

[3] J. C. Delvenne, R. Carli, and S. Zampieri, "Optimal strategies in the average consensus problem," *Wavelet Analysis and multiresolution methods*, vol. 56, pp. 759–765, 2009.

[4] B. Touri and A. Nedić, "On ergodicity, infinite flow and consensus in random models," *IEEE Transactions on Automatic Control*, vol. 56, no. 7, pp. 1593–1605, 2011.

[5] T. C. Aysal, M. E. Yildiz, A. D. Sarwate, and A. Scaglione, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, pp. 2508–2530, 2006.

[6] I. Daubechies and J. Lagarias, "Sets of matrices all infinite products of which converge," *Linear Algebra and its Applications*, vol. 161, pp. 227–263, 2002.

[7] J. Lorenz, "Consensus strikes back in the Hegselmann-Krause model of continuous opinion dynamics under bounded confidence," *Journal of Artificial Societies and Social Simulation*, vol. 9, pp. 227–263, 2006.

[8] J. D. Weston, "A short proof of Zorn's lemma," *Archiv der Mathematik*, vol. 8, no. 4, 1957.

[9] I. N. Herstein, *Topics in algebra*. Wiley India Pvt. Ltd, 2006.

[10] T. M. Apostol, *Mathematical analysis*. Wesley Publishing Company, 1981.

[11] D. P. Bertsekas, A. Nedić, and A. Ozdaglar, *Convex analysis and optimization*. Athena Scientific, Belmont, MA, 2003.

[12] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions. Autom. Control*, vol. 48, pp. 988–1001, 2003.

[13] L.Wang and F. Xiao, "Finite-time consensus problems for networks of dynamic agents," *IEEE Transaction On Automatic Control*, vol. 55, no. 5, 2010.

[14] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Artificial Societies and Social Simulation*, vol. 5, pp. 1–33, 2002.

[15] J. Lorenz, "Repeated averaging and bounded-confidence, modeling, analysis and simulation of continuous opinion dynamics," *PhD Thesis, University of Bremen*, 2007.

[16] B. Touri, "Product of random stochastic matrices and distributed averaging," *PhD Thesis, Department of Industrial Engineering, UIUC*, 2011.

[17] B. Touri and A. Nedić, "Hegselmann-Krause dynamics: An upper bound on termination time," *to appear in CDC*, 2012.

[18] J. M. Hendrickx, "Graphs and networks for the analysis of autonomous agent systems," *Ph.D. Thesis, Universite Catholique de Louvain*, 2011.

[19] S. Sundaran and C. N. Hadjicostis, "Finite-time distributed consensus in graphs with time-invariant topologies," *American Control Conference*, pp. 711–716, 2007.

[20] P. F. R. Carli, F. Fagnani, and S. Zampieri, "Gossip consensus algorithms via quantized communication," *Automatica*, vol. 46, pp. 70–80, 2010.

[21] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Transactions on Automatic Control*, vol. 54, pp. 2506–2517, 2009.

[22] T. C. Aysal, A. D. Sarwate, and A. G. Dimakis, "Reaching consensus in wireless networks with probabilistic broadcast," *Allerton'09 Proceedings of the 47th annual Allerton conference on communication, control, and computing*, pp. 732–739, 2009.