Stability of Agreement in State-dependent Interaction Environments

by

Alexander Condello

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Master of Applied Science

Queen's University
Kingston, Ontario, Canada
June 2013

Copyright © Alexander Condello, 2013

Abstract

We study stochastic stability for a class of agreement dynamics and define two forms of agreement which we call *stochastic agreement* and *stochastic absolute agreement*. We identify conditions for a broad class of random, possibly state-dependent agreement processes to achieve stochastic agreement. We take the approach of applying Lyapunov drift criteria to study the behaviour of such processes. We generalize some results in the literature to the noisy and state-dependent case.

Acknowledgments

I would like most to thank Professor Serdar Yüksel for his patience and guidance. He has been a role model throughout my education and he has been unwavering in his support.

I would also like to thank all of the professors I have had in my time at Queen's University. I would not have been able to write this thesis without the knowledge they have given me.

Table of Contents

List of Figures

Chapter 1

Introduction

Agreement in decentralized systems is a classic problem dating back to the 1950s. Agreement problems can be intuitively thought of as a problem of convergence of opinion. Most results in the literature consider iterative models in which a group of agents exchange signals (such as their estimates of a real random variable, we will refer to these estimates as opinions or values) and use this information to come to some form of agreement in their values.

The processes we study in this paper have iterative dynamics defined by equations of the form:

$$X_{t+1} = \lambda X_t \cdot F(X_t) + D(X_t) + W_t$$

where $\lambda \in (0, 1]$. Processes of this form bear a great similarity to the classical model introduced by DeGroot (1974) in [9]. In this model a group of N agents whose values at time t are given by the vector $X_t \in \mathbb{R}^N$ update their beliefs according to

$$X_{t+1} = X_t A_t$$

where $\{A_t\}_{t \geq 0}$ is sequence of matrices which model the social network of interactions.

Much of the results for processes that behave according to the DeGroot model are based on the work of Hajnal and Wolfowitz on products of stochastic matrices [27], [14]. They showed that products of *stochastic, indecomposable and aperiodic* (SIA) matrices will converge to a matrix with all columns equal. Thus applying products of SIA matrices will cause the process to achieve *consensus* in the limit. We will define these concepts in Chapter 2.

More recently, such models have received significant interest perhaps starting with Tsitsiklis and Bertsekas in [4], [24], [25] and later with Jadbabaie, Lin and Morse in [15], among many other recent contributions.

One use for this type of dynamic is social networks modelled by graph theory. In this framework each agent is modelled as a node in a graph and communication happens along the edges, e.g. [17], [21]. Later in this thesis we will discuss an example based on [17].

In the literature, state-dependent update algorithms have received less attention. In [6], Blondel, Hendricks and Tsitsiklis show that for a setup in which each agent only communicates with agents close to its value (the difference is less than one in their case), the system converges to a set of clusters. In [7] Canuto, Fagnani and Tilli consider continuous distributions of agents to obtain convergence results. In [3] [8], authors also analyze the asymptotic behaviour of distributions of large numbers of agents. Friedkin and Johnson consider a biased update mechanism [12].

The literature considered so far in this thesis only considers update mechanisms which do not follow from Baye's rule. Such models are often referred to as non-Bayesian update rules. These rules are desirable since convergence rates to agreement are typically larger and such models may reflect real world settings better for a large class of interaction environments. This is especially true when a probabilistic update is computationally expensive. In contrast, a Bayesian update rules requires the agents to update their values in a Bayesian fashion using a complete knowledge of the probabilistic structure of the system. See [1], [5] for a discussion of such approaches. The results in this thesis are non-Bayesian.

In this thesis we are interested in two forms of stochastic stability. Let us consider a group of N agents whose state (or value) is given by $x \in \mathbb{R}^N$. Let $\mu : \mathbb{R}^N \to \mathbb{R}^N$ defined by $\mu(x) = (\sum_{i=1}^{N} \frac{1}{N} x(i))\mathbf{1}_N$ where $\mathbf{1}_N$ is the vector of length N with all entries equal to 1. Thus, $\mu(x)$ is a vector with all of its entries equal to the mean of x.

We say that a state process $\{X_t\}$ achieves *stochastic agreement* or *stochastically agrees* if X_t has a finite expected return time to a set of the form

$$\Gamma_A = \{x : ||x - \mu(x)||_1 \leq A\} \tag{1.1}$$

for some $A \in \mathbb{R}$.

We say that $\{X_t\}$ achieves *stochastic absolute agreement* or *absolutely stochastically agrees* if X_t has a finite expected return time to a set of the form

$$\Gamma = \{x : ||x||_1 \leq B\} \tag{1.2}$$

for some $B \in \mathbb{R}$. These concepts will be made more formal in Section 2.4.

As opposed to asymptotic agreement which is the result typically adopted in the literature, stochastic agreement is a weaker form of agreement but allows for a much more general analysis. Our definition allows for not only the addition of random zero-mean noise to the system but also drift away from the mean. In addition it is a more realistic form of consensus in that in many processes, agents will often agree, then disagree, then agree again over time. A disadvantage to stochastic agreement vs stochastic absolute agreement is that stochastic agreement does not say anything about the stability of the mean value of the system.

In this thesis, we first identify conditions based on the work of Wolfowitz [27] for a process to stochastically agree. This is done using stochastic drift arguments first in single step increments then over m steps. We then give conditions, based on the work of Dobrushin [10] and Wolfowitz, and using a theorem from [28], for a process to achieve stochastic absolute agreement. We then give an example illustrating a process that stochastically absolutely agrees.

Finally we give a theorem that shows that many results already in the field dealing with products of random matrices that achieve asymptotic agreement will achieve stochastic agreement with the addition of drift and noise terms, under further technical conditions.

The motivation for this work began as an interest in creating a model that resembles real world situations, most notably social networks. We have attempted to take into account several forces acting on an agent: the consensus which brings the opinions of agents closer together, the drift which pushes the opinions of agents further away, and the random drift which allows for noisy systems. Any of these drifts can be dependent on the current state of the system of agents.

This thesis is formatted as follows:

- In Chapter 2 we review some of the theory related to products of stochastic matrices and stochastic stability. Some important definitions and theorems from the literature of products of stochastic matrices are presented. This is followed by a discussion on Markov Chains. Finally the main definitions of agreement for this paper are formally presented.

- In Chapter 3 we discuss criteria for stochastic agreement. Further, we will discuss a simple example in which the average opinion of the agents behaves as a random walk.

- In Chapter 4 we discuss criteria for stochastic absolute agreement. We also provide a case study of a complex system based on social dynamics that achieves stochastic absolute agreement.

- Finally, in Chapter 5 we provide a different approach to determine criteria for stochastic agreement. In this chapter we take inspiration from literature results and look at whether systems that achieve asymptotic agreement under DeGroot's Model will achieve stochastic agreement with the additions of noise and drift.

Chapter 2

Preliminaries

2.1 Notation

We consider a network of N nodes. Let $(\Omega, \mathcal{F}, P)$ be the underlying probability space and let $\{X_t\}_{t\geq 0}$ be a sequence of random row vectors taking values in $\mathbb{R}^N$. We denote the state realization of the system at time t by a row vector x_t where $x_t(i)$ is the value held by the ith node at time t.

Let $\mathbf{1}_N$ denote the vector of length N with each entry equal to 1. Define $\mu_t = \frac{1}{N}\sum_{i=1}^{N} X_t(i)$ and thus μ_t is the average of the values of X_t. Define $\mu(X_t) = \mu_t \mathbf{1}_N$ thus $\mu(X_t)$ is a vector with all entries equal to the average of X_t. Note that because the values of μ_t depend on X_t, this is a function of X_t however for notational simplicity we will use μ_t. Let $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$ i.e. $\mathcal{F}_t$ is the sigma field generated by X_k for $k \leq t$. Let us denote $E_{x_t}[\,\cdot\,] := E[\,\cdot\,|X_t = x_t]$ and $P_{x_t}(\cdot) = P(\cdot|X_t = x_t)$.

2.2 Products of Stochastic Matrices

We begin with an overview of some of the main definitions regarding products of stochastic matrices, a topic we will borrow from heavily in this paper.

Let us define the following matrix properties:

***Definition* 2.2.1 (Stochastic Matrix)** *Let P be a $N \times N$ matrix. We say that P is a stochastic matrix if $\sum_j P(i,j) = 1,\ P(i,j) \geq 0\ \forall i, j$. Note that the product of two stochastic matrices is itself stochastic.*

***Definition* 2.2.2 (Doubly Stochastic Matrix)** *Let P be a $N \times N$ matrix. We say that P is a doubly stochastic matrix if P is stochastic and $\sum_i P(i,j) = 1$. Note that the product of two doubly stochastic matrices is itself doubly stochastic. Also note that a doubly stochastic matrix is average preserving. That is if P is doubly stochastic, $x_0 \in \mathbb{R}^N$ and $x_1 = x_0 P$ then $\mu_1 = \mu_0$.*

***Definition* 2.2.3 (SIA)** *A finite stochastic matrix P is stochastic, indecomposable and aperiodic or SIA if*

$$Q = \lim_{n \to \infty} P^n$$

exists and all columns of Q are equal, that is $Q(i,j) = Q(k,j)$ for all i, j, k.

***Definition* 2.2.4 (Wolfowitz's Coefficient)** *Let P be a finite square matrix. We define Wolfowitz's Coefficient, denoted $\gamma(P)$, by*

$$\gamma(P) = \max_k \max_{i,j} |P(i,k) - P(j,k)|$$

Thus, $\gamma(P)$ can be thought of as measuring the maximum distance between any pair of elements in the same column.

Wolfowitz's Coefficient is primarily important for its relation to the product of stochastic matrices as detailed below in Theorem 2.2.1 which also comes from [27].

Let $A_1, \ldots, A_k$ be finite square matrices of the same order. By a word of length m in the A's we mean a product of a sequence of matrices of length m with repetitions permitted.

***Theorem* 2.2.1** *Let $A_1, \ldots, A_k$ be finite square stochastic matrices such that any word in the A's is SIA. For any $\epsilon > 0$ there exists an integer $v(\epsilon)$ such that any word C (in the A's) of length $n \geq v(\epsilon)$ satisfies $\gamma(C) < \epsilon$.*

Many simple consensus algorithms can be shown to achieve consensus by applying this theorem.

Now we will define a similar coefficient that will be used in our theorems.

***Definition* 2.2.5 (Dobrushin's Ergodic Coefficient)** *Let P be a finite square matrix. We define Dobrushin's Coefficient, denoted $\delta(P)$, by*

$$\delta(P) = \min_{i,j} \sum_k \min(P(i,k), P(j,k))$$

Note that $\delta(P) > 0$ if and only if, for every two rows, there exists one column for which both terms are positive.

An important result using this coefficient, originally discussed by Dobrushin in [10], is as follows,

***Theorem* 2.2.2** *For any vectors π and σ of length $k \in \mathbb{N}$ such that $\sum_{i=1}^k \pi(i) = \sum_{i=1}^k \sigma(i)$,*

$$||\pi P - \sigma P||_1 \leq (1 - \delta(P))||\pi - \sigma||_1$$

Finally, let us note the relation between Dobrushin's and Wolfowitz's Coefficients. The following is stated by J. Hajnal in Lemma 3 of [14]:

***Theorem* 2.2.3** *Let A be an $N \times N$ stochastic matrix. Then,*

$$\gamma(A) \leq (1 - \delta(A))$$

and,

$$N\gamma(A) \geq (1 - \delta(A))$$

A similar result is Lemma 2 of [27] and earlier Theorem 2 of [14]:

***Lemma* 2.2.1** *For any k,*

$$\gamma(P_1 P_2 \cdots P_k) \leq \prod_{i=1}^{k} (1 - \delta(P_i)) \tag{2.1}$$

Let us consider this demonstrative example:

***Example* 2.2.1** *Let us consider a network of n nodes $\{N_1, \ldots, N_n\}$. Let $\xi = \{\xi_1, \ldots, \xi_m\}$ (i.e. ξ is a set containing m possible sets of edges) such that $(\{N_1, \ldots, N_n\}, \xi)$ form* *a connected graph, that is any two nodes are connected by a set of edges. Let communication at each time t be facilitated by a set of edges E_t which takes values in ξ such that it takes each value in ξ before repeating itself.*

At each communication, let two nodes average their values in the following sense: if agents i and j communicate then

$$X_{t+1}(i) = \epsilon X_t(j) + (1-\epsilon)X_t(i) \tag{2.2}$$

where $\epsilon \in (0, \frac{1}{N+1})$.

Thus our system dynamics can be modelled by:

$$x_{t+1} = x_t \cdot F_t$$

where x_t is the row vector of the values for each of the nodes and F_t depends on E_t in the following way:

$$F(i,j) = \begin{cases} 1 - \epsilon \cdot deg(v_i) & \text{if } i = j \\ \epsilon & \text{if } i \neq j \text{ and } v_i \text{ is adjacent to } v_j \\ 0 & \text{else} \end{cases}$$

where $deg(v_i)$ is the degree of vertex i at time t. We call a matrix formed in such a way a Laplacian Matrix.

Now, each $F_t \cdots F_{t+m}$ is such that $\delta(F_t \cdots F_{t+m}) > \alpha$ for some $\alpha > 0$. Thus, a slight extension of Theorem 2.2.1 can be used to show that

$$\lim_{t\to\infty} X_t = X_0 \lim_{t\to\infty} \prod_{i=0}^{t-1} F_i = X_0 \cdot Q$$

where Q is a matrix with equal columns. Thus $\lim_{t\to\infty} X_t$ is a constant vector almost surely.

2.3 Stability of Markov Chains

In this thesis we will be considering consensus in the sense that the vector of opinions will return infinitely often to a set that we will consider to indicate consensus. To analyze this setup we will begin with some important definitions. Let us consider a $\mathbb{R}^N$ valued Markov Chain $\{Z_t\}$.

***Definition* 2.3.1 (ψ-irreducibility)** *For a finite positive measure ψ, a Markov chain $\{Z_t\}$ is ψ-irreducible if $\forall D \in \mathcal{B}(\mathbb{R}^N)$ (where $\mathcal{B}(\mathbb{R}^N)$ is the Borel σ-field on $\mathbb{R}^N$) with $\psi(D) > 0$ and $z \in \mathbb{R}^N$, $\exists n$ such that*

$$P(z_{t+n} \in D | z_t = z) > 0$$

***Definition* 2.3.2 (Harris Recurrence)** *Let A be a set such that $\psi(A) > 0$. Let $\eta_A^T = \sum_{t=0}^{T} 1_{\{X_t \in A\}}$. That is, η_A^T is the number of times that a Markov chain X_t hits the set A until time T. Let $\eta_A = \lim_{T\to\infty} \eta_A^T$. The set A is Harris recurrent if $P_x(\eta_A = \infty) = 1$ $\forall x \in A$. The chain is Harris Recurrent if $P_x(\eta_A = \infty) = 1$ $\forall x \in \mathbb{R}^N$ and $\psi(A) > 0$.*

***Definition* 2.3.3 (positive Harris Recurrence)** *If a ψ-irreducible Markov chain is Harris Recurrent and admits an invariant probability measure, then the chain is called positive Harris Recurrent.*

***Definition* 2.3.4 (Aperiodicity)** *A ψ-irreducible Markov chain is aperiodic if for any $x \in \mathbb{R}^N$ and any $B \in \mathcal{B}(\mathbb{R}^N)$ satisfying $\psi(B) > 0$ there exists $n_0(x, B)$ such that*

$$P^n(x, B) > 0 \text{ for all } n \geq n_0$$

***Definition* 2.3.5 (Small Set)** *A set $A \in \mathcal{B}(\mathbb{R}^N)$ is called ν-small if $\exists n \in \mathbb{N}$ such that $P^n(x, D) \geq \nu(D)$ $\forall D \in \mathcal{B}(\mathbb{R}^N)$ $\forall x \in A$ for a positive measure ν, on $\mathcal{B}(\mathbb{R}^N)$.*

***Definition* 2.3.6 (Filtration)** *An increasing family $\{\mathcal{F}_n\}$ of sub σ-fields defined on a probability space $(\Omega, \mathcal{F}, P)$ with $\mathcal{F}_n \subset \mathcal{F}$, is called a filtration.*

***Definition* 2.3.7 (Supermartingale)** *Let X_t be measurable on a filtration $\mathcal{F}_t = \sigma(X_0, \ldots, X_t)$ (also called the natural filtration). $(X_n, \mathcal{F}_n)$ is said to be a supermartingale if*

$$E[|X_t|] < \infty$$

and

$$E[X_{t+1}|\mathcal{F}_t] \leq X_t$$

for all $t \geq 0$

***Definition* 2.3.8 (Stopping Time)** *A function τ from a measurable space $(\Omega, \mathcal{F}, P)$ to $(\mathbb{N}_+, \mathcal{B}(\mathbb{N}_+))$ is a stopping time if for all $n \in \mathbb{N}_+$, the event $\{\tau = n\} \in \mathcal{F}_n$, with $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$.*

We will also need the following theorem from [28].

***Theorem* 2.3.1** *Suppose X is a ψ-irreducible, aperiodic Markov chain on $\mathbb{R}^N$. Let $\tau_0 = 0$ and $\{\tau_z\}$ be a sequence of stopping times. Suppose that there is a function $V : \mathbb{R}^N \to (0, \infty)$, a small set C and a constant $b \in \mathbb{R}^+$ and $\varepsilon > 0$ such that the following hold:*

$$E[V(X_{\tau_{z+1}})|X_{\tau_x} = x] \leq V(X_{\tau_z}) - \varepsilon + b1_{\{X_{\tau_z} \in C\}}$$

$$\sup_{z \geq 0} E[\tau_{z+1} - \tau_z | \mathcal{F}_{\tau_z}] < \infty$$

Then $\{X_t\}$ is positive Harris recurrent.

2.4 Main Definitions for this Paper

Finally, we wish to formally define the concepts of *Stochastic Agreement* and *Stochastic Absolute Agreement.*

***Definition* 2.4.1 (Stochastic Agreement)** *Let $\{X_t\}_{t>0}$ be a sequence of random variables taking values in $\mathbb{R}^N$. Let us define a consensus set, Γ_A by*

$$\Gamma_A = \{x : ||x - \mu(x)||_1 \leq A\} \tag{2.3}$$

for some $A \in \mathbb{R}$. Let us define a sequence of stopping times for a process X_t by:

$$\tau_{z+1} = \min(t > \tau_z : x_t \in \Gamma_A)$$

with $\tau_0 = 0$. We say that the process achieves stochastic agreement if

$$\sup_z E[\tau_{z+1} - \tau_z | \mathcal{F}_{\tau_z}] < \infty$$

and $\forall x \in \mathbb{R}^N$

$$P_x(\min(t > 0 : x_t \in \Gamma_A) < \infty) = 1$$

***Definition* 2.4.2 (Stochastic Absolute Agreement)** *Let $\{X_t\}_{t>0}$ be a sequence of random variables taking values in $\mathbb{R}^N$. Let us define a consensus set, Γ by:*

$$\Gamma = \{x : ||x||_1 \leq C\} \tag{2.4}$$

for some $C \in \mathbb{R}$ as well as a sequence of stopping times, $\{\tau_z\}$ defined by

$$\tau_{z+1} = \min(t > \tau_z : x_t \in \Gamma)$$

with $\tau_0 = 0$. We say that $\{X_t\}_{t>0}$ achieves stochastic absolute agreement if

$$\sup_z E[\tau_{z+1} - \tau_z | \mathcal{F}_{\tau_z}] < \infty$$

and $\forall x \in \mathbb{R}^N$

$$P_x(\min(t > 0 : x_t \in \Gamma) < \infty) = 1$$

***Theorem* 2.4.1** *If $\{X_t\}_{t>0}$ is a ψ-irreducible, aperiodic Markov Chain with Γ a small set and $\{X_t\}_{t>0}$ achieves stochastic absolute agreement then $\{X_t\}_{t>0}$ is positive Harris Recurrent.*

Proof: First note that

$$\sup_{x_0 \in \Gamma} E_{x_0}[\min(t > 0 : X_t \in \Gamma)] < \infty.$$

Furthermore, $P_{x_0}(\min(t > 0 : X_t \in \Gamma) < \infty) = 1 \ \forall x_0 \in \mathbb{R}^N$. The proof then follows from Theorem 4.1 of Meyn-Tweedie [19]. $\diamond$

Even if $\{X_t\}$ is not irreducible, Definition 2.4.2 may lead to the existence of an invariant probability measure, see Theorem 2.2 in [18] and Theorem 12.3.9 in [28] for weak-Feller chains, which is a class of Markov chains which satisfy certain continuity properties.

Thus, stochastic agreement means that the process has a bounded expected return time to an agreement set Γ_A (which is not compact) and stochastic absolute agreement means that a process has a bounded expected return time to a compact set containing the origin.

Chapter 3

Stochastic Agreement and State-Dependence

Let W_t be a sequence of i.i.d. zero-mean noise vectors such that $E[||W_t||_1] < \infty$. Let us define a system that behaves according to the following equation:

$$X_{t+1} = X_t \cdot F(X_t) + D(X_t) + W_t \tag{3.1}$$

where W_t is a sequence of i.i.d. zero-mean noise vectors such that $E[||W_t||_1] < \infty$. $F(X)$ is a matrix-valued random variable such that $F(\cdot)$ defines a regular conditional probability measure from $\mathbb{R}^N$ to the set of matrices. This implies that, for every realization of x, $F(x)$ is a matrix-valued random variable. Similarly, for every realization of x, $D(x)$ is a vector-valued random variable.

We are interested in determining under what conditions a process that behaves according to the dynamics described in (3.1) will achieve stochastic agreement as defined in Definition 2.4.1. In order to do this, we need a version of Foster-Lyapunov drift criteria (see [26])

***Theorem* 3.0.2** *Let* $\tau = \min(t > 0 : X_t \in \Gamma_A)$, $V : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}_+$ *with* $\sup_{x \in \Gamma_A} V(x) < \infty$ *and* $b < \infty$. *If for some* $\epsilon > 0$ *the following holds*

$$E[V(X_{t+1})|X_t = x_t] \leq V(X_t) - \epsilon + b1_{\{X_t \in \Gamma_A\}} \tag{3.2}$$

then $\sup_{x_t \in \Gamma_A} E_{x_t}[\tau_{\Gamma_A}] < \infty$.

Proof: First let us define a sequence by $M_0 = V(x_0)$ and for $t \geq 0$:

$$M_{t+1} = V(X_{t+1}) - \sum_{k=0}^{t}(-\epsilon + b1_{\{X_k \in \Gamma\}})$$

Let us also define a stopping time as:

$$\tau^N = \min\{N, \tau\},$$

where N is an integer. Now, for any t,

$$\begin{aligned}
E[|M_t|] &= E[|V(X_t) + t\epsilon - \sum_{k=0}^{t-1}(b1_{\{X_k \in \Gamma\}})|] \\
&\leq E[V(X_t)] + t\epsilon + bt \\
&= E[E[V(X_t)|X_{t-1}]] + t\epsilon + bt \\
&\leq E[V(X_{t-1}) - \epsilon + b1_{\{X_{t-1} \in \Gamma_A\}}] + t\epsilon + bt \\
&\leq E[V(x_0)] + 2bt < \infty,
\end{aligned}$$

where the last step follows by an inductive argument. Now,

$$\begin{aligned}
E[M_{t+1}|\mathcal{F}_t] &\leq V(X_t) - \epsilon + b1_{\{X_t \in \Gamma_A\}} - \sum_{k=0}^{t}(-\epsilon + b1_{\{X_k \in \Gamma\}}) \\
&= V(X_t) - \sum_{k=0}^{t-1}(-\epsilon + b1_{\{X_k \in \Gamma\}}) \\
&= M_t
\end{aligned}$$

for every t. Thus,

$$E[M_{t+1}|\mathcal{F}_t] \leq M_t \tag{3.3}$$

Hence, $\{M_t\}_{t \geq 0}$ is a supermartingale. In addition τ^N is a bounded stopping time.

We have therefore satisfied the conditions for Doob's optional sampling theorem:

$$E[M_{\tau^N}] \leq M_0$$

Thus, we obtain for Γ_A,

$$\begin{aligned}
E_{x_0}[\sum_{i=0}^{\tau^N-1} \epsilon] &= E_{x_0}[M_{\tau^N} + \sum_{i=0}^{\tau^N-1} b1_{\{X_i \in \Gamma_A\}} - V(X_{\tau^N-1})] \\
&\leq M_0 + bE_{x_0}[\sum_{i=0}^{\tau^N-1} 1_{\{X_i \in \Gamma_A\}}] - E[V(X_{\tau^N-1})] \\
&= V(x_0) + bE_{x_0}[\sum_{i=0}^{\tau^N-1} 1_{\{X_i \in \Gamma_A\}}] - E[V(X_{\tau^N-1})] \\
&\leq V(x_0) + bE_{x_0}[\sum_{i=0}^{\tau^N-1} 1_{\{X_t \in \Gamma_A\}}]
\end{aligned}$$

Now, since for $1 \leq k \leq \tau^N - 1$, $X_k \notin \Gamma_A$,

$$\sup_{x_0 \in \Gamma_A} E_{x_0}[\tau^N] \leq \sup_{x_0 \in \Gamma_A} \frac{1}{\epsilon}(V(x_0) + b)$$

and therefore $\sup_{x_0 \in \Gamma_A} E_{x_0}[\tau^N] < \infty$

Now since $\tau^{N+1} \geq \tau^N$ for all N almost surely, we can apply the monotone convergence theorem

$$E_{x_0}[\tau] = E_{x_0}[\lim_{N\to\infty} \tau^N] = \lim_{N\to\infty} E_{x_0}[\tau^N] \leq \frac{1}{\epsilon}(V(x_0) + b)$$

which gives us

$$\sup_{X_0 \in \Gamma_A} E_{x_0}[\tau] < \infty$$

$\diamond$

3.1 Single-Step Stochastic Agreement

This brings us to our first set of criteria for the process (3.1) achieving stochastic agreement.

***Theorem* 3.1.1** *If $\forall x_t$, $E_{x_t}[\delta(F(X_t))] > \alpha$ for some $\alpha > 0$, every realization of $F(X_t)$ is doubly stochastic and $||D(x)||_1 < D < \infty$ almost surely $\forall x$, then X_t achieves Stochastic Agreement.*

Proof: Let us define a function $V : \mathbb{R}^n \to \mathbb{R}_+$ by $V(x_t) = ||x_t - \mu(x_t)||_1$.

Now,

$$\begin{aligned}
E_{x_t}[V(X_{t+1})] &= E_{x_t}[||X_{t+1} - \mu(X_{t+1})||_1] \\
&= E_{x_t}[||X_t F(X_t) + D(X_t) + W_t - \mu(X_{t+1})||_1] \\
&= E_{x_t}[||X_t F(X_t) - \mu(X_t)F(X_t) + \mu(X_t)F(X_t) \\
&\quad + D(X_t) + W_t - \mu(X_{t+1})||_1] \\
&= E_{x_t}[||(X_t - \mu(X_t))F(X_t) + D(X_t) + W_t \\
&\quad + \mu(X_t) - \mu(X_{t+1})||_1] && (3.4) \\
&\leq E_{x_t}[||(X_t - \mu(X_t))F(X_t)||_1 + ||D(X_t)||_1 \\
&\quad + ||W_t||_1 + ||\mu(X_t) - \mu(X_{t+1})||_1] \\
&\leq E_{x_t}[||(X_t - \mu(X_t))F(X_t)||_1] + D + W + U && (3.5) \\
&\leq E_{x_t}[1 - \delta(F(X_t))]||x_t - \mu(x_t)||_1 + D + W + U && (3.6) \\
&= V(X_t) - E_{x_t}[\delta(F(X_t))]||x_t - \mu(x_t)||_1 + D + W + U
\end{aligned}$$

In the above, (3.4) follows from that fact that every realization of $F(x_t)$ is doubly stochastic for every x_t and μ has constant entries. Thus, we get $\mu_t F(X_t) = \mu_t$. Also, since $||D(X_t)||_1$ is bounded $E_{x_t}[||D(X_t)||_1] < D$, and let $E_{x_t}[||W_t||_1] < W$. Here, (3.5) is given by the fact that the difference in the expected value of the mean is bounded thus $||\mu(X_t) - \mu(X_{t+1})||_1 \leq U$ (for further proof see Section A.2). Crucially, (3.6) follows from Theorem 2.2.2.

Now, by assumption $E_{x_t}[\delta(F(X_t))] > \alpha$, therefore

$$V(X_t) - E_{x_t}[\delta(F(X_t))]||x_t - \mu(x_t)||_1 + D + W + U \leq V(X_t) - \alpha||x_t - \mu(x_t)||_1 + D + W + U$$

Thus we arrive at:

$$E_{x_t}[V(X_{t+1})] \leq V(X_t) - \alpha||x_t - \mu(x_t)||_1 + D + W + U \tag{3.7}$$

for some constants $D, W, U \in \mathbb{R}$.

Let us define a set

$$\Gamma = \{x : ||x - \mu(x)||_1 \leq \frac{D + W + U + \epsilon}{\alpha}\}$$

This will be our consensus set. Now, following from (3.7) we arrive at:

$$\begin{aligned} E_{x_t}[V(X_{t+1})] &\leq V(x_t) - \alpha||x_t - \mu(x_t)||_1 + D + W + U \\ &\leq V(x_t) - \epsilon + (D + W + U + \epsilon)1_{x_t \in \Gamma} \end{aligned}$$

Thus we satisfy the conditions of Theorem 3.0.2 thus for $\tau_{z+1} = \min(t > \tau_z | X_t \in \Gamma)$, $\sup_z E[\tau_{z+1} - \tau_z | \mathcal{F}_z] < \infty$ and $\sup_{x \in \Gamma} V(x) < \infty$.

Also (by Theorem 3.0.2), $\forall x$

$$E_x[\min(t > 0 : X_t \in \Gamma)] \leq V(x) + b$$

for some $b > 0$. Therefore,

$$P_x(\min(t > 0 : X_t \in \Gamma) \geq M) \leq \frac{E_x[\min(t > 0 : X_t \in \Gamma)]}{M}$$

Now letting $M \to \infty$ we get that $P_x(\min(t > 0 : X_t \in \Gamma) < \infty) = 1$.

Thus X_t achieves stochastic agreement.

$\diamond$

3.2 m-Step Stochastic Agreement

We would like to relax our requirements for the possible values of F. Thus, we can extend Theorem 3.1.1 with the following Theorem:

***Theorem* 3.2.1** *Let X_t be a vector of length N. Let $F(X_t)$ be such that $\exists m$ such* *that for every* x_t, $E_{x_t}[\delta(\prod_{i=0}^{m-1} F(X_{t+i})] > \alpha$. *Let also, for every* x_t, $F(x_t)$ *be doubly stochastic and* $||D(X_t)||_1 < D$. *Then* X_t *achieves stochastic agreement.*

Proof: The proof of Theorem 3.2.1 follows similarly to that of Theorem 3.1.1, see Section A.1 in the Appendix for the complete proof.

Chapter 4

Stochastic Absolute Agreement

Let us consider the following example:

***Example* 4.0.1** *Let $F(\cdot) \in \mathcal{L}$ where $\mathcal{L}$ is the set of $N \times N$ Laplacian Matrices (defined in Example 2.2.1). Let $D(\cdot)$ be a zero vector. Let W_t be a vector of i.i.d, standard normal random variables. Because there are a finite number of possible Laplacian matrices we can apply Theorem 3.2.1. We see that while X_t will always return to some set centred at its average value, the average value itself will behave as a random walk since $\mu(X_{t+1}) = \mu(X_t) + \mu(W_t)$. The results of a simulation run with this setup can be seen in Figure 4.1.*

Thus we see the main issue with stochastic agreement - that is says nothing about the stability of the overall opinion of the group. In contrast, stochastic absolute agreement has X_t returning to some compact set about 0 in finite time almost surely. Thus we wish to give criteria for X_t to achieve stochastic absolute agreement.

4.1 Problem Setup

Let us define a new system as follows:

$$X_{t+1} = \lambda X_t \cdot F(X_t) + D(X_t) + W_t \tag{4.1}$$

where W_t is a sequence of i.i.d zero-mean Gaussian vectors (we also note that W_t does not have to be Gaussian, it can admit any density which is bounded and $P(W_t \in B) >$

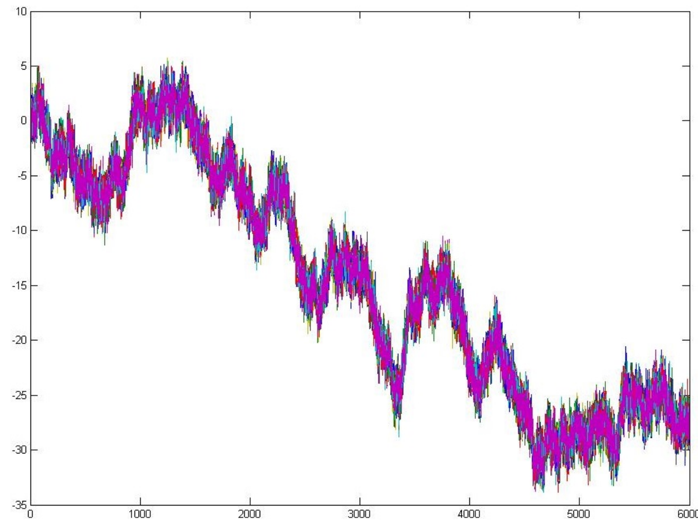

Figure 4.1

0 for all open non-empty B and $||W_t||_1 < \infty$), $\lambda \in (0,1)$. $F(X)$ is a matrix-valued random variable such that $F(\cdot)$ defines a regular conditional probability measure from $\mathbb{R}^N$ to the set of matrices. This implies that, for every realization of x, $F(x)$ is a matrix-valued random variable. Similarly, for every realization of x, $D(x)$ is a vector-valued random variablewith the additional assumption that $\sum_i D(X_t)(i) = 0$ and $||D(x)||_1$ is (uniformly) bounded almost surely.

We will show that the process achieves stochastic absolute agreement.

Figure 4.2 shows a simulation of a process that evolves according to (4.1) with $\lambda = 0.99$ and μ_t highlighted in black.

4.2 Stochastic Absolute Agreement - Main Theorem

***Theorem* 4.2.1** *Let X_t be a vector of length N. Let for every x_t, every realization of $F(x_t)$ be stochastic and for every x, let $||D(x)||_1$ be (uniformly) bounded almost surely. Let W_t be a vector of zero mean Gaussian noise. Finally let $\lambda \in (0,1)$. Then X_t absolutely stochastically agrees.*

Proof:

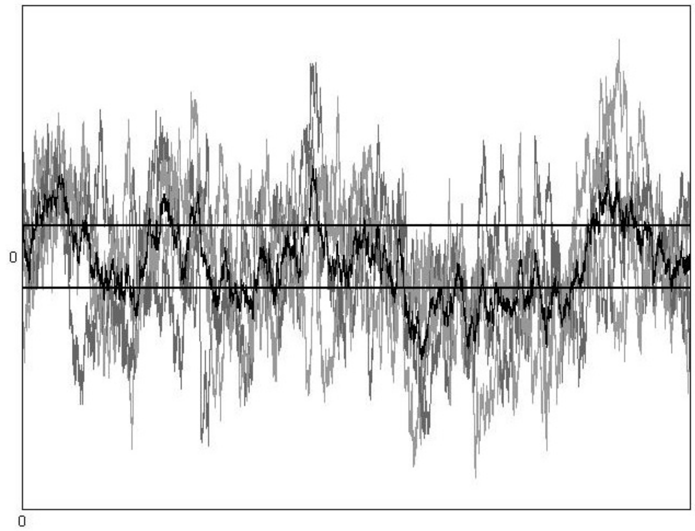

Figure 4.2

Let us define a function $V : \mathbb{R}^n \to \mathbb{R}_+$ by $V(x_t) = ||x_t||_1$.

Now,

$$\begin{aligned} E_{x_t}[V(X_{t+1})] &= E_{x_t}[||X_{t+1}||_1] \\ &= E_{x_t}[||\lambda X_t F(X_t) + D(X_t) + W_t||_1] \\ &\leq \lambda ||x_t||_1 + E_{x_t}[||D(x_t)||_1] + E[||W_t||_1] \\ &= V(x_t) - (1-\lambda)||x_t||_1 + D + W \\ &= V(x_t) - \epsilon + \Big(\epsilon + D + W - (1-\lambda)||x_t||_1\Big) \\ &= V(x_t) - \epsilon + b1_{\{||x_t|| \in \Gamma_\Lambda\}}, \end{aligned} \tag{4.2}$$

where $b = \epsilon + D + W$ and $\Gamma_\Lambda = \{x : ||x||_1 \leq (\epsilon + D + W)/(1-\lambda)\}$. The inequality above follows from Appendix A.3 and the triangle inequality.

As in the proof of Theorem 3.1.1, the above implies that the expected return times to Γ_Λ from Γ_Λ are uniformly bounded in expectation.

Now, following the same steps as in the proof of Theorem 2.1 in [28] (which is similar to Theorem 3.0.2), we obtain that for all x, the condition: $E_x[\min(t > 0 :$

$x_t \in \Gamma_\Lambda)] < \infty$ follows and, as in the proof of Theorem 3.1.1:

$$P_x(\min(t > 0 : x_t \in \Gamma_\Lambda) < \infty) = 1$$

for all x.

Thus we get that X_t stochastically absolutely agrees. $\diamond$

We have the following useful result.

***Theorem* 4.2.2** *The set* $\Gamma_\Lambda = \{x : ||x||_1 \leq C\}$ *is small.*

Proof:

Following an approach taken by Tweedie in Lemma 4 [26] and Proposition 5.5.5(iii) of [19] and in [28], we will use the following test: if a set S is such that

$$\lim_{n\to\infty} \sup_{x\in S} P(x, B_n) = 0 \tag{4.3}$$

is satisfied for every sequence $B_n \downarrow \emptyset$ and if the Markov chain is irreducible (which ours is due to our additive noise term W_t), then S is petite. By Theorem 5.5.5(iii) of Meyn-Tweedie [19] for an aperiodic irreducible chain, petiteness is equivalent to smallness. The setup in this thesis is aperiodic because, due to the presence of the Gaussian noise, $P(x_{t+1} \in S_A | x_t = x) > 0$ for all $x \in S_A$, for any compact Borel set S_A with positive Lebesgue measure [18]. This implies that a small set to be constructed will be visited in consequent time stages with positive probability (see Theorems 5.2.2 and 5.4.4 in [18]).

Let $B_n \to \emptyset$ and let $K = [-k, k]^N$ for $k \in \mathbb{R}$. Note that $B_n = (B_n \cap K) \cup (B_n \cap K^C)$ and if $B_n \to \emptyset$, then both $(B_n \cap K) \to \emptyset$ and $(B_n \cap K^C) \to \emptyset$.

Now,

$$\begin{aligned}\lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(X_{t+1} \in B_n | X_t = x) &= \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(\lambda x \cdot F(x) + D(x) + W_t \in B_n | X_t = x) \\ &\leq \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(\lambda x \cdot F(x) + D(x) + W_t \in B_n \cap K | X_t = x) \\ &+ \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(\lambda x \cdot F(x) + D(x) + W_t \in B_n \cap K^C | X_t = x)\end{aligned} \tag{4.4}$$

Let $M_A = \sup_{x\in S} ||D(x)||_1 + ||x||_1$. For the second component above, we have that:

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(\lambda x \cdot F(x) + D(x) + W_t \in B_n \cap K^C | X_t = x) \\
&\leq \sup_{x\in S} P(\lambda x \cdot F(x) + D(x) + W_t \in K^C | X_t = x) \\
&\leq P(W_t \in ([-k - M_A, k + M_A]^N)^C)
\end{aligned}$$

Note now that for any $\epsilon > 0$, there exists k such that $P(W_t \in ([-k - M_A, k + M_A]^N)^C) \leq \epsilon$.

For the first component in (4.4), we note that

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(\lambda x \cdot F(x) + D(x) + W_t \in B_n \cap K | X_t = x) \\
&= \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} \int_z P(z + W_t \in B_n \cap K) P(\lambda x \cdot F(x) + D(x) \in dz | X_t = x) \\
&= \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} \int_z P(W_t \in (B_n \cap K) - z) P(\lambda x \cdot F(x) + D(x) \in dz | X_t = x) \\
&\leq \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} \int_z \left(\sup_z \bar{\mu}(z)\right) \lambda((B_n \cap K) - z) P(\lambda x \cdot F(x) + D(x) \in dz | X_t = x) \\
&= \lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} \int_z \left(\sup_z \bar{\mu}(z)\right) \lambda(B_n \cap K) P(\lambda x \cdot F(x) + D(x) \in dz | X_t = x) \\
&= \lim_{n\to\infty} \left(\sup_z \bar{\mu}(z)\right) \lambda(B_n \cap K) \sup_{x\in\Gamma_\Lambda} \int_z P(\lambda x \cdot F(x) + D(x) \in dz | X_t = x) \\
&= \left(\sup_z \bar{\mu}(z)\right) \lim_{n\to\infty} \lambda(B_n \cap K) = 0,
\end{aligned}$$

where $\bar{\mu}$ is the density of W_t and $\lambda(\cdot)$ is the Lebesgue measure which is translation invariant (that is, $\lambda((B_n \cap K) - z) = \lambda(B_n \cap K)$). Note that $\lambda(B_n \cap K) \to 0$.

As a result, for every $\epsilon > 0$, $\lim_{n\to\infty} \sup_{x\in\Gamma_\Lambda} P(X_{t+1} \in B_n | X_t = x) \leq \epsilon$ and (4.3) holds. ◇

Hence, by invoking Theorem 2.3.1, X_t is positive Harris recurrent.

4.3 Example

Let us close the chapter by considering an example to which Theorem 4.2.1 can be applied.

Let us consider a state-dependent example inspired by social dynamics. Let there be a group of 10 political followers (10 agents) who support one of two political parties. Over the course of a year, the congress has a chance to enact 100 pieces of legislation. After each piece of legislation is enacted, the agents have a chance to talk to each-other and debate. The process values describe how favourably each agent views Party A and unfavourably they view party B, e.g. if an agent has a positive opinion, he/she approves of Party A and disapproves of Party B, if another agent has a negative opinion then he/she approves of Party B and disapproves of Party A.

Formally, let us consider 10 agents whose opinion at time t is given by $X_t \in \mathbb{R}^{10}$. At each time step the following dynamics occur:

- An agent with a non-negative opinion is said to approve of Party A, each agent with a negative opinion is said to support Party B and to be a critic of Party A
- Each person's opinion is also affected by some random variation, in our case modelled by a Gaussian noise process.
- Each agent has a 10% chance of communicating with every other agent. If both agents are supporters or both agents are critics then they average their views. If one agent is a supporter and one agent is a critic then they each move a small amount (5%) closer i.e. $X_{t+1}(i) = \lambda(.95X_t(i) + .05X_t(j)) + D(X_t)(i) + W_t(i)$ and $X_{t+1}(j) = \lambda(.95X_t(j) + .05X_t(i)) + D(X_t)(j) + W_t(j)$
- At the same time, because these political followers are so involved in politics, their opinion affects the ability of each party to enact legislation. The probability of the congress enacting legislation favourable to Party A is given by $P(A) = .6\frac{\#X>0}{10} + .2$, that is if all agents support Party A then Party A has a 80% chance of enacting favourable legislation, else Party B enacts legislation it favours.
- Which team enacts favourable legislation affects the opinion of the agents as well. If Party A enacts legislation than each agent supporting Party A's opinion

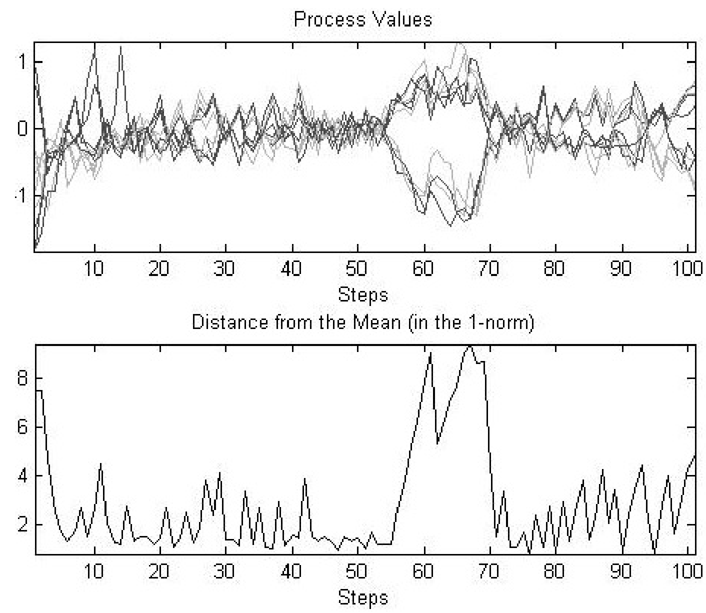

Figure 4.3

moves in the positive direction by $\frac{\delta}{\#X>0}$ and vice versa for Party B. Consider the following which accounts for the cases $\#X > 0 = 0$ and $\#X < 0 = 0$:

$$D(X)(i) = \begin{cases} \frac{\delta}{\#X>0} & \text{if } X(i) > 0 \text{ and } \#(X(j) > 0) \neq 0 \\ -\frac{\delta}{\#X\leq 0} & \text{if } X(i) \leq 0 \text{ and } \#(X(j) \leq 0) \neq 0 \\ \delta & \text{if } \max_j X(j) = X(i) \text{ and } \#(X(j) > 0) = 0 \\ -\delta & \text{if } \min_j X(j) = X(i) \text{ and } \#(X(j) \leq 0) = 0 \\ \frac{\delta}{9} & \text{if } \max_j X(j) \neq X(i) \text{ and } \#(X(j) > 0) = 0 \\ -\frac{\delta}{9} & \text{if } \min_j X(j) \neq X(i) \text{ and } \#(X(j) \leq 0) = 0 \end{cases}$$

- Finally, the total process has a factor of $\lambda = .99$ which represents the tendency for opinions to move towards neutrality in the absence of outside forces.

In Figure 4.3 we see a sample output from a process run under the above conditions. As we can see the process remains within some compact set, only leaving occasionally. We therefore see examples in the simulation of times when people are divided in their support for each party and times when the difference is small.

Chapter 5

Stochastic Agreement Under DeGroot's Model

Many results in the literature are based on the DeGroot Model [9]; that is, their dynamics are defined by equations of the form:

$$X_{t+1} = X_t A_t \tag{5.1}$$

where $\{A_t\}_{t\geq 0}$ is a sequence of stochastic (not necessarily doubly stochastic) matrices. Under these dynamics a process is considered to have achieved consensus if there exists an X such that $\lim_{t\to\infty} X_t = X$ almost surely. Often X is required to have all entries the same, or even that all entries be equal to the average of X_0.

Notably, J. Wolfowitz in [27] studies this problem for when $\{A_t\}_{t\geq 0}$ is a sequence of SIA matrices; this was discussed extensively in Chapter 2.

We would like to give conditions that a process that achieve consensus under the DeGroot model will achieve Stochastic Consensus under dynamics of the form:

$$X_{t+1} = X_t A_t + D(X_t) + W_t \tag{5.2}$$

where $D(.)$ is such that every realization of $D(X_t)$ defines a bounded random vector and W_t is a zero-mean Gaussian vector.

However, not every sequence $\{A_t\}_{t\geq 0}$ that achieves consensus under the DeGroot Model will achieve stochastic agreement under the process defined by (5.2) as the

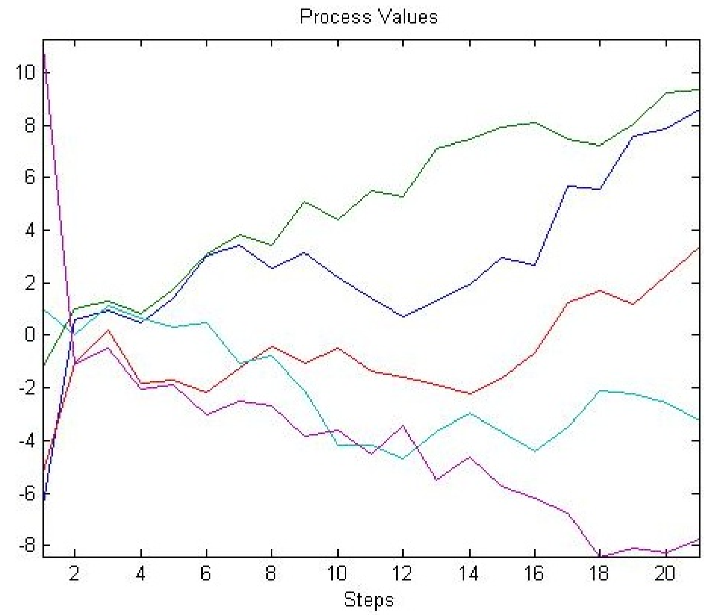

Figure 5.1

following example demonstrates:

***Example* 5.0.1** *Let $\{A_t\}_{t\geq 0}$ be defined in the following way:*

- *Let $A_0[i,j] = \frac{1}{N}$ for all i, j*
- *Let $A_t = I_{N\times N}$ for all $t > 0$ where $I_{N\times N}$ is the $N \times N$ identity matrix.*

Thus, under the DeGroot Model, for any $X_0 \in \mathbb{R}^N$, X_1 will be the vector with all entries equal to the average of X_0 and $X_t = X_1, t > 1$. Thus $\{A_t\}_{t\geq 0}$ achieves consensus under the DeGroot Model.

However, under the process defined by (5.2) and taking $D(X) = 0$ it is clear that after the first step, each agent i will behave independently with

$$X_{t+1}(i) = X_t(i) + W_t(i)$$

In other words, each agent will behave as an independent random walk. A sample output of a simulation run under these dynamics is seen in Figure 5.1. Thus, stochastic agreement is not achieved.

We therefore strengthen our requirement on $\{A_t\}_{t\geq 0}$ to requiring that for all $T \geq 0$, $\lim_{s\to\infty} \prod_{t=T}^{s} A_t$ is a matrix with all columns the same almost surely.

5.1 Stochastic Agreement - Main Theorem

Let $\{A_t\}_{t\geq 0}$ be a sequence of independent not necessarily identical $N \times N$ stochastic matrices such that for all $T \geq 0$, $\lim_{s\to\infty} \prod_{t=T}^{s} A_t$ is a matrix with equal columns almost surely. Fix $0 < \delta < 1$. For some sequence of vectors $\{X_t\}_{t\geq 0}$, let us define $\rho_0 = 0$ and for $t > 0$

$$\rho_t = \min\{s > \rho_{t-1} : ||X_{\rho_{t-1}} \prod_{t=\rho_{t-1}}^{s-1} A_t - \mu(X_{\rho_{t-1}} \prod_{t=\rho_{t-1}}^{s-1} A_t)||_1 \leq \delta ||X_{\rho_{t-1}} - \mu(X_{\rho_{t-1}})||_1\} \tag{5.3}$$

Let us note that $\sup_t E[\rho_{t+1} - \rho_t | \mathcal{F}_{\rho_t}]$ is not necessarily bounded.

***Theorem* 5.1.1** *Let $\{A_t\}_{t\geq 0}$ be a sequence of $\mathbb{R}^{N\times N}$ matrices such that for all $Z_0 \in \mathbb{R}^N$ and for all $T \geq 0$ $\lim_{s\to\infty} \prod_{t=T}^{s} A_t$ is a matrix with equal columns surely. Let also $\sup_t E[\rho_{t+1} - \rho_t | \mathcal{F}_{\rho_t}] < \infty$ where ρ_t is the sequence of stopping times defined in (5.3).*

Then the process defined by

$$X_{t+1} = X_t A_t + D(X_t) + W_t$$

where $D(.)$ is such that every realization of $D(X_t)$ defines a bounded random vector and W_t is a zero-mean, Gaussian vector achieves Stochastic Agreement.

Proof:

Let us denote by $E_{\mathcal{F}_{\rho_z}}[\cdot] = E[\cdot|\mathcal{F}_{\rho_z}]$ where $\mathcal{F}_{\rho_z} = \sigma(X_0, \ldots, X_{\rho_z})$ i.e. $\mathcal{F}_k$ is the sigma field generated by X_s for $s \leq k$. Let $T = \sup_t E[\rho_{t+1} - \rho_t | \mathcal{F}_{\rho_t}] < \infty$.

Let us define $V(X) = ||X - \mu(X)||_1$ and note the following inequalities:

$$\begin{aligned} E_{\mathcal{F}_{\rho z}}[V(X_{\rho_{z+1}})] &= E_{\mathcal{F}_{\rho z}}[||X_{\rho_{z+1}} - \mu(X_{\rho_{z+1}})||_1] \\ &= E_{\mathcal{F}_{\rho z}}[||X_{\rho_{z+1}-1} A_{\rho_{z+1}-1} + D(X_{\rho_{z+1}-1}) + W_{\rho_{z+1}-1} - \mu(X_{\rho_{z+1}-1})||_1] \end{aligned}$$

$$\leq E_{\mathcal{F}_{\rho z}}[||X_{\rho_{z+1}-1}A_{\rho_{z+1}-1} - \mu(X_{\rho_{z+1}})||_1] + D + W \tag{5.4}$$

$$\leq E_{\mathcal{F}_{\rho z}}[||(X_{\rho_{z+1}-2}A_{\rho_{z+1}-2} + D(X_{\rho_{z+1}-2})$$

$$+ W_{\rho_{z+1}-2})A_{\rho_{z+1}-1} - \mu(X_{\rho_{z+1}})||_1] + D + W$$

$$\leq E_{\mathcal{F}_{\rho z}}[||X_{\rho_{z+1}-2}A_{\rho_{z+1}-2}A_{\rho_{z+1}-1} - \mu(X_{\rho_{z+1}})||_1] + 2(D + W) \tag{5.5}$$

$$\leq E_{\mathcal{F}_{\rho z}}[||X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k - \mu(X_{\rho_{z+1}})||_1] + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W)$$

$$\leq E_{\mathcal{F}_{\rho z}}[||X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k - \mu(X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k)||_1$$

$$+ ||\mu(X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k) - \mu(X_{\rho_{z+1}})||_1] + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W)$$

$$\leq E_{\mathcal{F}_{\rho z}}[||X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k - \mu(X_{\rho_z}\prod_{k=\rho_z}^{\rho_{z+1}-1} A_k)||_1]$$

$$+ U + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W) \tag{5.6}$$

$$\leq \delta E_{\mathcal{F}_{\rho z}}[||X_{\rho_z} - \mu(X_{\rho_z})||_1] + U + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W) \tag{5.7}$$

$$= \delta V(X_{\rho_z}) + U + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W)$$

$$= V(X_{\rho_z}) - (1-\delta)V(X_{\rho_z}) + U + \sum_{k=0}^{E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W)$$

$$\leq V(X_{\rho_z}) - (1-\delta)V(X_{\rho_z}) + U + \sum_{k=0}^{\sup_z E_{\mathcal{F}_{\rho z}}[\rho_{z+1}-\rho_z]} (D + W) \tag{5.8}$$

$$\leq V(X_{\rho_z}) - \epsilon + (U + \sum_{k=0}^{T}(D + W))1_{\{X_{\rho_z}\in\Gamma\}} \tag{5.9}$$

(5.4) comes from the boundedness of $D(.)$ and the finite variance of W_t. (5.5) follows from the fact that A_t is stochastic and therefore $\exists\alpha$ such that $E||(W_t+D(X_t))A_t||_1 \leq D + W$ (See Appendix A.3). (5.7) uses the mean criteria of the theorem. (5.6) follows from proof in Appendix A.2. In (5.8) we add $\sup_z E_{\mathcal{F}_{\rho z}}[\rho_{z+1} - \rho_z]$ rather than $E_{\mathcal{F}_{\rho z}}[\rho_{z+1} - \rho_z]$ and we know that $\sup_z E_{\mathcal{F}_{\rho z}}[\rho_{z+1} - \rho_z] < \infty$.

In (5.9) we define our consensus set Γ by

$$\Gamma = \{x : ||x - \mu(x)||_1 \leq \frac{U + T(D + W) + \epsilon}{1 - \delta}\}$$

Finally, $\forall x$

$$E_x[\min(t > 0 : x_t \in \Gamma)] \leq V(x) + b$$

for some $b > 0$. Therefore,

$$P_x(\min(t > 0 : x_t \in \Gamma) \geq M) \leq \frac{E_x[\min(t > 0 : x_t \in \Gamma)]}{M}$$

Now letting $M \to \infty$ we get that $P_x(\min(t > 0 : x_t \in \Gamma) < \infty) = 1$.

Thus X_t achieves stochastic agreement.

Thus, we have satisfied the relevant criteria in Theorem 2.3.1.

$\diamond$

***Remark* 5.1.1** *In contrast to the work of Hajnal and Wolfowitz [14] [27] which makes the assumption that* $|\rho_{z+1} - \rho_x| < M$ *surely for some* $M \in \mathbb{R}$*, Theorem 5.1.1 only assumes that* $\sup_t E[\rho_{t+1} - \rho_t | \mathcal{F}_{\rho t}] < \infty$*.*

5.2 Example

Another way that Theorem 5.1.1 differs from the theorems presented in Chapters 3 and 4 is that it does not require the consensus matrix to be doubly stochastic. This allows for some more interesting examples including one-way communication as the following example taken from social dynamics illustrates.

Let us consider a small company of 8 people. This company has a power structure. There is a president, two managers and five workers. This company is beginning a new project. The opinion of each employee about the project on day t is given by a vector $x_t \in \mathbb{R}^8$. Let $x_t(1)$ represent the president's opinion, $x_t(2)$ and $x_t(3)$ represent the manager's opinions and all others represent the worker's opinions. For this model let x_0 be a standard normal random vector.

Let us model the opinion dynamics of the company in the following way

- Each employee's opinion is affected by noise which is a modelled by a standard normal random variable - see equation (5.10)
- Each worker is also likely to see new information in light of their current opinion - that is, the opinion of each worker who already see the project negatively will continue to go down, and vice versa for the workers who see the project positively. The president has seen many projects so his/her opinion is not affected by the others' opinions. For the managers, one is a pessimist and one is an optimist. These dynamics are modelled by

$$D(x_t)(i) = \begin{cases} 0 & \text{if } i = 1 \\ -.25 & \text{if } i = 2 \\ .25 & \text{if } i = 3 \\ -.25 & \text{if } i = 4,5,6,7,8 \text{ and } x_t(i) \leq 0 \\ .25 & \text{if } i = 4,5,6,7,8 \text{ and } x_t(i) > 0 \end{cases}$$

- Each day each employee sends an email to another randomly selected employee. If the recipient of the email is of equal or lower rank to the sender, their opinion will change to be 7% closer to that of the sender's. E.g if employee i who is a worker gets emails from employees j and k on day t then

$$x_{t+1}(i) = (1 - 2(.07))x_t(i) + .07x_t(j) + .07x_t(k) + D(x_t)(i) + w_t(i) \tag{5.10}$$

  Thus the maximum weight by which an employee can be influenced is 49%.

Let us note that under these dynamics the noise is a standard normal random vector and the drift term is bounded.

Because the president never changes his opinion based on signals from other

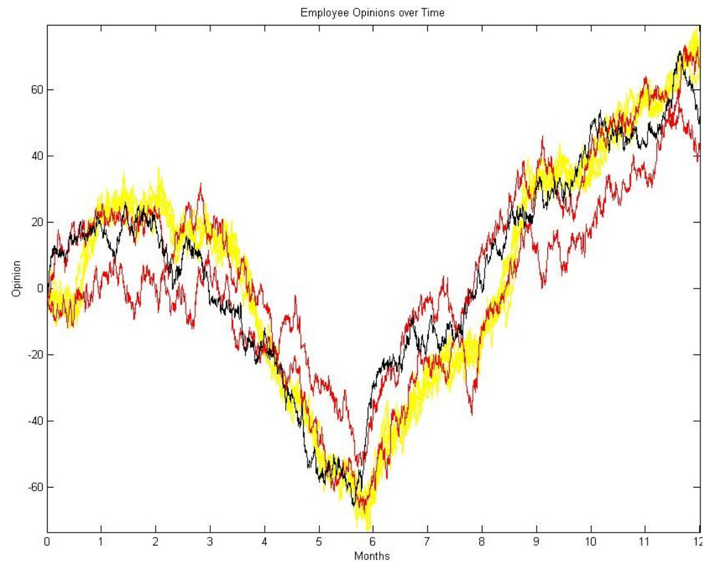

Figure 5.2

agents, it can also be seen that for all T, almost surely

$$\lim_{s\to\infty}\prod_{t=T}^{s} A_t = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Since there are finitely many possible realizations of A_t each with $\delta(A_t) > 0$, one can apply Wolfowitz's Theorem 2.2.1, and thus $\sup_t E[\rho_{t+1} - \rho_t | \mathcal{F}_{\rho_t}] < \infty$. Thus this scheme fits the criteria for Theorem 5.1.1.

The results of a simulation run under these dynamics can be seen in Figure 5.2. As can be seen, the beliefs of the agents in the simulation tend to stabilize into a tube but the average value of the tube is not itself stable. The tube moves with the president's opinion which behaves as a random walk.

Chapter 6

Conclusion

6.1 Contributions

The notions of stochastic agreement and stochastic absolute agreement allow for a more realistic and general notion of stability. The main contribution of this thesis is providing criteria for different processes to achieve stochastic agreement and stochastic absolute agreement.

6.2 Future Research Areas

This thesis concerns itself with stability. A process can be stable in the sense of stochastic agreement or even stochastic absolute agreement without actually being *practically stable*. The criteria that a process return to a consensus set in finite time with probability 1 may be useful, however for many practical applications the rate of return to agreement may be more relevant. A possible direction for further research would be establishing a *degree of consensus* for such processes; e.g. a process that has an expected return time of 1000 steps to the consensus set is in some sense less in agreement than a process that has an expected return time of 10 steps.

In Chapter 4 we provide criteria for a process to be positive Harris Recurrence, this means that it has a unique invariant distribution (the uniqueness is due to the irreducibility of the process). This invariant distribution would be a way to access the *degree of consensus* in the sense that if we know the steady state distribution on the consensus set, through Kac's Theorem (see Theorem 10.4.9 in [18]), we may obtain useful bounds.

Bibliography

[1] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian Learning in Social Networks", LIDS report 2780, 2009.

[2] D. Acemoglu and A. Ozdaglar, "Opinion Dynamics and Learning in Social Networks", *Dynamic Games and Applications*, vol.1, no.1, 2011.

[3] D. Acemoglu, G. Como, F. Fagnani and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks", *Mathematics of Operations Research*, vol.38, no.1. 2013.

[4] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

[5] V. Borkar and P. Varaiya, "Asymptotic Agreement in Distributed Estimation", *IEEE Transactions on Automatic Control*, vol.AC-27, no.3, 1982.

[6] V. Blondel, J Hendrickx, and J. Tsitsiklis, "Continuous-time Average-preserving Opinion Dynamics with Opinion-dependent Communications", *SIAM Journal on Control and Optimization*, vol.48, no.8, 2010.

[7] C. Canuto, F. Fagnani and P. Tilli, "An Eulerian Approach to the Analysis of Krause's Consensus Model", *SIAM J. Control Optim.*, vol.50, pp.243-265, 2012.

[8] G. Como and F. Fagnani, "Scaling Limits for Continuous Opinion Dynamics Systems", in *The Annals of Applied Probability*, vol.21(4), pp.1537-1567, 2011

[9] M. DeGroot, "Reaching a Consensus", *Journal of the American Statistical Association*, vol.69, pp. 118-121, 1974.

[10] R. Dobrushin, "Central Limit Theorem for Nonstationary Markov Chains. I", *Theory of Probability and its Applications*, vol.1, no.1, 1956.

[11] F. Fagnani and S. Zampieri, "Randomized Consensus Algorithms over Large Scale Networks", *IEEE Journal on Selected Areas in Communications*, vol.26, no.4, 2008.

[12] N. Friedkin and E Johnsen, "Social Influence Networks and Opinion Change", *EJ Lawler and MW Macy (eds.), Advances in Group Processes*, vol.16, 1999.

[13] F. Garin and S. Zampieri, "Mean Squared Performance of Consensus-Based Distributed Estimation Over Regular Geometric Graphs", *SIAM Journal on Control and Optimization*, vol.50, no.1, 2012.

[14] J. Hajnal, "Weak ergodicity in non-homogeneous Markov chains", *Proc. Cambridge Philos. Soc.*, vol.54, pp.233-246, 1958.

[15] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbour rules", *IEEE Trans. Automat. Contr.*, vol.48, pp. 988-1001, 2003.

[16] S. Li and T. Başar, "Asymptotic Agreement and Convergence of Asynchronous Stochastic Algorithms", *IEEE Transactions on Automatic Control*, vol.AC-32, no.7, 1987.

[17] I. Matei, N. Martins and J. Baras, "Almost Sure Convergence to Consensus in Markovian Random Graphs", in *Proc. IEEE Conference on Decision and Control*, 2008.

[18] S. P. Meyn and R. Tweedie, *Markov chains and Stochastic Stability*, Springer-Verlag, London, 1993.

[19] Meyn, S. P., and R. Tweedie, "Stability of Markovian processes I: Criteria for discrete-time chains", *Advances in Applied Probability*, vol.24, pp.542-574, 1992.

[20] M. Porfiri and D. Stilwell, "Consensus Seeking Over Random Weighted Directed Graphs", *IEEE Transactions on Automatic Control*, vol.52, no.9, 2007.

[21] A. Tahbaz-Salehi and A. Jadbabaie, "A Necessary and Sufficient Condition for Consensus Over Random Networks", *IEEE Transactions on Automatic Control*, vol.53, no.3, 2008.

[22] A. Tahbaz-Salehi and A. Jadbabaie, "On Consensus Over Random Networks", *44th Annual Allerton Conference*, 2006.

[23] B. Touri and A. Nedić, "On Ergodicity, Infinite Flow, and Consensus in Random Models", *IEEE Transactions on Automatic Control*, vol.56, no.7, 2011.

[24] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans,, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms", *IEEE Trans. Autom. Control*, vol.AC-31, no.9, pp.803-812, 1986.

[25] J. N. Tsitsiklis, *Problems in decentralized decision making and computation*, Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, 1984.

[26] R. Tweedie, "Drift conditions and invariant measures for Markov chains", *Stochastic processes and their applications*, vol.92, pp.345-354, 2001.

[27] J. Wolfowitz, "Products of indecomposable aperiodic stochastic matrices", *Proc. Am. Math. Soc.*, vol.15, pp.733-736, 1963.

[28] S. Yüksel and S. Meyn, "Random-Time, State-Dependent Stochastic Drift for Markov Chains and Application to Stochastic Stabilization Over Erasure Channels", *IEEE Transactions on Automatic Control*, vol.58, no.1, 2013.

Appendix A

Additional Proofs

A.1 Proof of Theorem 3.2.1

First since by assumption $E_{x_t}[\delta(\prod_{i=0}^{m-1} F(X_{t+i}))] > \alpha$ we know that $E_{x_t}[\gamma(\prod_{i=0}^{m-1} F(X_{t+i}))] < \frac{\alpha}{N}$

Next, let us also define a function $V : \mathbb{R}^N \to \mathbb{R}_+$ as $V(X_t) = ||X_t - \mu(X_t)||_1$.

Now,

$$
\begin{aligned}
E_{x_t}[V(X_{t+m})] &= E_{x_t}[||X_{t+m} - \mu(X_{t+m})||_1] \\
&= E_{x_t}[||X_{t+m-1}F(X_{t+m-1}) + D(X_{t+m-1}) + W_{t+m-1} - \mu(X_{t+m})||_1] \\
&\leq E_{x_t}[||X_{t+m-1}F(X_{t+m-1}) - \mu(X_{t+m})||_1] + E_{x_t}[||D(X_{t+m-1})||_1] \\
&\quad + E_{x_t}[||W_{t+m-1}||_1] \\
&\leq E_{x_t}[||X_{t+m-1}F(X_{t+m-1}) - \mu(X_{t+m})||_1] + D + W \qquad \text{(A.1)} \\
&= E_{x_t}[||(X_{t+m-2}F(X_{t+m-2}) + D(X_{t+m-2}) + W_{t+m-2})F(X_{t+m-1}) \\
&\quad - \mu(X_{t+m})||_1] + D + W \\
&= E_{x_t}[||(X_{t+m-2}F(X_{t+m-2})F(X_{t+m-1}) - \mu(X_{t+m}) \\
&\quad + D(X_{t+m-2})F(X_{t+m-1}) + W_{t+m-2}F(X_{t+m-1})||_1] + D + W \\
&\leq E_{x_t}[||(X_{t+m-2}F(X_{t+m-2})F(X_{t+m-1}) - \mu(X_{t+m})||_1] \\
&\quad + E_{x_t}[||D(X_{t+m-2})F(X_{t+m-1})||_1] \\
&\quad + E_{x_t}[||W_{t+m-2}F(X_{t+m-1})||_1] + D + W \\
&\leq E_{x_t}[||(X_{t+m-2}F(X_{t+m-2})F(X_{t+m-1}) - \mu(X_{t+m})||_1
\end{aligned}
$$

$$
\begin{aligned}
&+ ||D(X_{t+m-2})||_1 + ||W_{t+m-2}||_1 + D + W && \text{(A.2)}\\
&\leq E_{x_t}[||(X_{t+m-2}F(X_{t+m-2})F(X_{t+m-1}) - \mu(X_{t+m})||_1 + 2(D+W) && \text{(A.3)}\\
&\leq E_{x_t}[||(X_{t+m-2}F(X_{t+m-2})F(X_{t+m-1}) - \mu(X_{t+m})||_1 + 2(D+W)\\
&\leq E_{x_t}[||(X_t \prod_{k=0}^{m-1} F(X_{t+k}) - \mu(X_{t+m})||_1 + mD + mW\\
&= E_{x_t}[||(X_t \prod_{k=0}^{m-1} F(X_{t+k}) - \mu(X_t) + \mu(X_t) - \mu(X_{t+m})||_1 + mD + mW\\
&\leq E_{x_t}[||(X_t \prod_{k=0}^{m-1} F(X_{t+k}) - \mu(X_t) \prod_{k=0}^{m-1} F(X_{t+k})||_1 + U + mD + mW && \text{(A.4)}\\
&\leq E_{x_t}[1 - \delta(\prod_{k=0}^{m-1} F(X_{t+k}))]||x_t - \mu(x_t)||_1 + U + mD + mW\\
&\leq E_{x_t}[N\gamma(\prod_{k=0}^{m-1} F(X_{t+k}))]||x_t - \mu(x_t)||_1 + U + mD + mW && \text{(A.5)}
\end{aligned}
$$

Let $E_{x_t}[||D(X_t)||_1] < D$ and $E_{x_t}[||W_t||_1] < W$ which gives us (A.1) and (A.3) above. (A.2) comes from $||D(X_t)F(X_t)||_1 \leq (1 - \delta(F(X_t)))||D(X_t)||_1 \leq ||D(X_t)||_1$. (A.4) comes from $\mu(X_t)F(X_t) = \mu(X_t)$ for all $F(X_t)$ and $||\mu(X_t) - \mu(X_{t+m})||_1 \leq U$. (A.5) follows from Theorem 2.2.3.

Let us define a set:

$$\Gamma = \{x : ||x - \mu(x)||_1 \leq \frac{mD + mW + U + \epsilon}{\alpha}\}$$

where α is defined as above. We therefore arrive at:

$$
\begin{aligned}
E_{x_t}[V(X_{t+m})] &\leq E_{X_t}[N\gamma(\prod_{k=0}^{m-1} F(X_{t+k}))]||x_t - \mu(x_t)||_1 + mD + mW\\
&\leq \alpha||x_t - \mu(x_t)||_1 + mD + mW\\
&= V(X_t) - \alpha||x_t - \mu(x_t)||_1 + mD + mW + U && \text{(A.6)}\\
&\leq V(X_t) - \epsilon + (mD + mW + U + \epsilon)1_{X_t \in \Gamma}
\end{aligned}
$$

For some ϵ small. See proof of Theorem 3.1.1 for the details of A.6.

Also, $\forall x$

$$E_x[\min(t > 0 : x_t \in \Gamma)] \leq V(x) + b$$

for some $b > 0$. Therefore,

$$P_x(\min(t > 0 : x_t \in \Gamma) \geq M) \leq \frac{E_x[\min(t > 0 : x_t \in \Gamma)]}{M}$$

Now letting $M \to \infty$ we get that $P_x(\min(t > 0 : x_t \in \Gamma) < \infty) = 1$.

Thus X_t achieves stochastic agreement.

A.2 Bounding the Mean

We wish to show $E_{X_t}[||\lambda\mu(X_t) - \mu(X_{t+1})||_1] < U$ for some U when $X_{t+1} = \lambda X_t \cdot F(X_t) + D(X_t) + W_t$. Note that when $\lambda = 1$ this becomes $E_{X_t}[||\mu(X_t) - \mu(X_{t+1})||_1] < U$ for some U when $X_{t+1} = X_t \cdot F(X_t) + D(X_t) + W_t$.

Proof: Let $\mathbf{1}_{N,N}$ be the $N \times N$ matrix of all ones.

$$\begin{aligned}
E_{x_t}[||\lambda\mu(X_t) - \mu(X_{t+1})||_1] \quad &= E_{x_t}[||\lambda\mu(X_t) - \frac{1}{N}\mathbf{1}_{N\times N}X_{t+1}||_1] \\
&= E_{x_t}[||\lambda\mu(X_t) - \frac{1}{N}\mathbf{1}_{N\times N}(\lambda X_t F(X_t) + D(X_t) + W_t)||_1] \\
&= E_{x_t}[||\lambda\mu(X_t) - \lambda\mu(X_t)F(X_t) + \frac{1}{N}\mathbf{1}_{N\times N}(D(X_t) + W_t)||_1] \\
&= E_{x_t}[||\lambda\mu(X_t) - \lambda\mu(X_t) + \frac{1}{N}\mathbf{1}_{N\times N}(D(X_t) + W_t)||_1] \\
&= E_{x_t}[||\frac{1}{N}\mathbf{1}_{N\times N}(D(X_t) + W_t)||_1]
\end{aligned}$$

Which is bounded since $E[||D(X)||]$ and $E[||W_t||]$ are bounded by assumption. $\diamond$

A.3 Product of a Random Stochastic Matrix and a Random Vector

Let A_t be a possibly random $N \times N$ stochastic matrix. Let $u \in \mathbb{R}^N$ be a random vector such that there exists $U \in \mathbb{R}$ such that $E[||u||_1] < U$. We wish to show that $E[||uA_t||_1] < U$.

Proof:

$$E[||uA_t||_1] = E[\sum_{i=1}^{N} |\sum_{j=1}^{N} u(j)A(j,i)|] \leq E[\sum_{i=1}^{N}\sum_{j=1}^{N} |u(j)|A(j,i)] = E[\sum_{j=1}^{N} |u(j)|] = ||u||_1$$

◇