Titre: Title:	Linear Consensus Algorithms: Structural Properties and Connections with Markov Chains
Auteur: Author:	Sadegh Bolouki
Date:	2014
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Bolouki, S. (2014). Linear Consensus Algorithms: Structural Properties and Connections with Markov Chains [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1551/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/1551/
Directeurs de recherche: Advisors:	Roland P. Malhamé
Programme: Program:	génie électrique

UNIVERSITÉ DE MONTRÉAL

LINEAR CONSENSUS ALGORITHMS: STRUCTURAL PROPERTIES AND CONNECTIONS WITH MARKOV CHAINS

SADEGH BOLOUKI
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE ÉLECTRIQUE)
AOÛT 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

LINEAR CONSENSUS ALGORITHMS: STRUCTURAL PROPERTIES AND CONNECTIONS WITH MARKOV CHAINS

présentée par : BOLOUKI Sadegh
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. LE NY Jerôme, Ph.D., président
M. MALHAMÉ Roland P, Ph.D., membre et directeur de recherche
M. CAINES Peter E, Ph.D., membre
Mme NEDICH Angelia, Ph.D., membre

DEDICATION

To My Loving Parents

ACKNOWLEDGEMENTS

I owe a deep debt of gratitude to my adviser, Roland Malhamé, for his invaluable guidance, tireless efforts, and patience in supervising this thesis. None of the work presented here would have been possible without his insightful ideas and his constant support. What I have learned from him in research, as well as in real life, is absolutely indescribable.

I would also like to express my sincere gratitude to my thesis committee members, Angelia Nedić, Peter E. Caines, and Jérôme Le Ny, for carefully reading this thesis and providing feedback that helped me significantly improve it.

I would like to thank our collaborators, Nader Motee and Milad Siami, my colleague and wonderful friend, for their significant contributions in a chapter of this thesis.

I am grateful to Fatima, Thibault, and Mehdi with whom I have shared my office and had many interesting conversations. I am also thankful to Suzanne and Rachel of secretary staff of Poly for their great work.

Furthermore, I would like to thank my family members who, together with my friends, remain the meaning of my life outside work. My parents, Gholamhossein and Mahbeigom, and my siblings, Azadeh, Ehsan, and Adeleh, have been a limitless source of love and support during my life.

I am especially thankful to Abolfazl, Ahmad, Nakisa, Saeid J, and Mozhdeh for always being there for me. I am also thankful to my friends who turned my life in Montreal to a fantastic journey, to name a few, Sara–Saviz, Majid Sheykh, Gool, Rezanejat, Fabian, Diana, Parisa, K-mand, Ghasemi, Matin, Farshadi, Maya, Nazanin, Mansour, Jalal, and Farhadi! Finally, I would like to thank Reza Arastoo for his great suggestions in writing parts of this thesis.

RÉSUMÉ

Nous considérons un réseau d'agents multiples en interaction, et tel que chaque agent est supposé posséder un état concernant une certaine quantité d'intérêt. Selon le contexte, les états d'agents peuvent correspondre à des opinions, des valeurs, des estimés, des croyances, des positions, des vitesses, etc. Ces états sont mis à jour selon un algorithme ou protocole qui consiste en une règle d'interaction dictant la manière par laquelle les états d'un agent donné influencent ou sont influencés par ceux de ses voisins. Les voisins sont définis à partir d'un graphe sous-jacent de communication, ce dernier évoluant dans le temps de manière soit endogène ou exogène. Un consensus du système est défini comme la convergence de tous les états vers une valeur commune, lorsque le temps croît indéfiniment. La notion de consensus apparaît dans de multiples domaines de recherche. En biologie, le consensus est lié aux comportements émergents d'un ensemble d'oiseaux en vol, des bancs de poissons, etc. En robotique et en automatique, les problèmes de consensus se présentent lorsque l'on cherche à réaliser la coordination et la coopération d'agents mobiles (ex. robots et capteurs). Cette question est particulièrement importante dans la mise en réseau de capteurs avec nombreuses applications, soit en contrôle de l'environnement, ou dans un contexte militaire. En économie, la recherche de consensus par rapport à un mécanisme commun d'ajustement des prix constitue un autre exemple. En sociologie, l'émergence d'une langue commune dans une société primitive est un comportement collectif au sein d'un système complexe. Un autre comportement limite d'intérêt dans un système est celui où les états, plutôt que de converger vers une seule valeur, se fractionnent en groupes distincts, avec des limites communes dans le groupe mais distinctes d'un groupe à l'autre. Un tel comportement est appelé dans notre thèse, consensus multiple.

Dans cette thèse, nous adressons deux objectifs de recherche en relation avec le comportement asymptotique des états d'agents dans un système multi-agent, possédant une dynamique mise à jour via un algorithme distribué de calcul de moyenne, de caractère général, en temps continu ou discret. Le premier objectif visé est celui de l'identification de conditions aussi faibles que possible, pour lesquelles le consensus unique ou multiple est garanti inconditionnellement, c'est-à-dire pour toute valeur du temps initial ou encore des valeurs initiales attribuées aux états. Contrairement au premier objectif centré sur la recherche de convergence inconditionnelle, notre second objectif de recherche est celui de l'identification d'ensembles particuliers de conditions initiales, non triviales toutefois, pour lesquelles un consensus global est possible.

En particulier, nous nous intéressons à la caractérisation de coalitions d'agents dites

“coalitions à éminences grises” (EGC). Un EGC est un regroupement possiblement très limité d’individus dans un réseau, en mesure de mener “naturellement” la totalité du groupe à converger vers un état commun arbitraire, simplement par un choix adéquat et concerté de leurs états initiaux au sein de la coalition. Par “naturellement”, il est entendu que les membres de la coalition ne peuvent avoir recours à des manipulations de la structure du réseau, y compris celle de leurs propres interactions avec leurs voisins. Ils peuvent être considérés comme des leaders dans l’ombre, non identifiables a priori par un titre ou une position particuliers. Ils détiennent un potentiel d’influence totale sur le comportement des autres individus dans le réseau. Notre investigation des EGC dans un réseau d’opinions est centrée avant tout sur la caractérisation de la taille de la plus petite coalition EGC possible.

ABSTRACT

We consider a network of multiple interacting agents, whereby each agent is assumed to hold a state regarding a certain quantity of interest. Depending on the context, states may be referred to as opinions, values, estimates, beliefs, positions, velocities, etc. Agent states are updated based on an algorithm or protocol which is an interaction rule specifying the manner in which individual agent states influence and are influenced by neighboring states. Neighbors are defined via an underlying exogenously or endogenously evolving communication graph. *Consensus* in the system is defined as convergence of all states to a common value, as time grows large. The notion of consensus arises in many research areas. In biology, consensus is linked with the emergent behavior of bird flocks, fish schools, etc. In robotics and control, consensus problems arise when seeking coordination and cooperation of mobile agents (e.g., robots and sensors). This is, particularly, an important issue in sensor networking with wide applications in environmental control, military applications, etc. In economics, seeking an agreement on a common belief in a price system is another example of consensus. In sociology, the emergence of a common language in primitive societies is a collective behavior within a complex system. Another important limiting behavior of the system is one whereby agents, instead of all converging to the same value, separate into multiple clusters with a uniform limiting value within each cluster. Such behavior, in this thesis, is called *multiple consensus*.

In this thesis, we address two research objectives relating to the asymptotic behavior of agent states in a multi-agent system, with dynamics updated via a general distributed averaging algorithm in either continuous time or discrete time. The first issue is that of identifying conditions, as weak as possible, under which consensus or multiple consensus is guaranteed to occur unconditionally, i.e., irrespective of the time or values that states are initialized at. In contrast to the first research objective centered on unconditional consensus, our second research objective is that of identifying sets of particular, yet non-trivial, initial agent conditions such that global consensus occurs.

In particular, we are interested in characterizing so-called éminence grise coalitions (EGC): An EGC is a possibly small group of individuals in the network who are, "naturally", capable of leading the whole group to eventually agree on any desired value, by *only* choosing their own initial values properly. What is meant by "naturally" is that the group in question does not need to manipulate the nature of the network, and in particular, leaves all the interactions between any two individuals including members of the group themselves untouched. They could be thought of as hidden leaders, not specifically identifiable by title or position, but who hold the potential of perfectly influencing the asymptotic behavior of individuals in

the network. In investigating EGCs in a network of opinions, the size of its smallest EGC is the main focus of our analysis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background Information

A multi-agent system, in the most general sense, is a network of multiple interacting agents. Each agent is assumed to hold a state regarding a certain quantity of interest. Depending on the context, states may be referred to as opinions, values, beliefs, positions, velocities, or etc. States of agents are updated based on an algorithm or protocol which is an interaction rule specifying the interaction between each agent and its neighbors. *Global consensus*, or simply *consensus*, in the system is defined as convergence of all states to a common value over time. Among all update algorithms in multi-agent systems, distributed averaging algorithms are of great importance and have been discussed the most in the literature. Such algorithms impose that the state of each agent is updated according to a convex combination of the current states of its neighbors and its own.

The notion of consensus arises in many research areas. In biology, consensus dynamics are at the heart of the emergent behavior of bird flocks, fish schools, etc. Couzin *et al.* (2005); Cucker and Smale (2007); Flierl *et al.* (1999). Consensus models can be employed to interpret, analyze, and predict flocking aggregation behavior. In robotics and control, consensus problems arise in relation to coordination objectives and cooperation of mobile agents (e.g., robots and sensors) Jadbabaie *et al.* (2003); Tsitsiklis *et al.* (1986), which are important issues in sensor networks for environmental applications or potentially space exploration. In economics, seeking an agreement on a common belief in a price system is another example of consensus. In sociology, the emergence of a common language in primitive societies Cucker *et al.* (2004) is a collective behavior within a complex system, while in the area of social networks, consensus algorithms can shed light on the dynamics of opinion formation. Consensus algorithms also have a rich history within computer science community Lynch (1996), while formal study of consensus problems has been carried within the management science community (see DeGroot (1974) and references therein).

In a multi-agent system, it is possible that agents separate into several clusters such that consensus occurs within each cluster. In this case, *multiple consensus* is said to have occurred. In other words, occurrence of multiple consensus is equivalent to the existence of individual limits for the agent states such that the limits are not necessarily equal. Although not as extensively studied as global consensus, multiple consensus has also been widely discussed in the literature (see Chatterjee and Seneta (1977); Lorenz (2005) for example).

1.2 Problem Definition

Consider a system composed of N agents that are labeled by numbers $1, \ldots, N$. Let $x_i(t)$ be the scalar state of agent i at time $t \geq 0$. Distributed averaging algorithms can be defined in both continuous and discrete times. A general continuous time distributed averaging algorithm is defined by:

$$\dot{x}(t) = A(t)x(t),\ t \geq 0. \tag{1.1}$$

where $x(t)$ is the vector of states at each time instant $t \geq 0$, and $\{A(t)\}$ is the *underlying chain* of the system. It is assumed that each matrix of underlying chain $A(t)$ has zero row sum and non-negative off-diagonal entries and each entry $a_{ij}(t)$ of $A(t)$ is a measurable function. Consensus is now defined by the convergence of vector $x(t)$ to a vector with equal components as $t \to \infty$. Multiple consensus is also defined as the existence of a limit for $x_i(t)$ for each agent i as time grows large. Indeed, the limits may differ for different agents. Coefficients $a_{ij}(t)$ can be either purely exogenous time-varying functions, or endogenous functions dictated by the evolution of x_i's. The focus of this thesis is, first and foremost, on exogenously defined underlying chains, although some of our results will also apply to endogenously evolving consensus algorithms.

This work is motivated by the following two fundamental questions regarding the issue of consensus:

Q.1 Under what conditions on the underlying chain of the system, consensus or multiple consensus is guaranteed irrespective of the time and values that states are initialized?

Q.2 For a general underlying chain, having fixed the initial time, what is the set of initial conditions resulting in the occurrence of consensus in the system?

Answering to Q.1 is our main objective in Chapters 3 and 4. Indeed, finding an answer to Q.1 has been a challenge for several decades now. As suggested in Chatterjee and Seneta (1977), the occurrence of consensus as sought in Q.1 is equivalent to a property of the underlying chain called *ergodicity*. We now recall the definition of ergodicity of a chain. Let $\Phi(t, \tau)$, $t \geq \tau \geq 0$, represent the state transition matrix of the system, i.e.,

$$x(t) = \Phi(t, \tau)x(\tau),\ \forall t \geq \tau \geq 0. \tag{1.2}$$

Chain $\{A(t)\}$ is said to be *ergodic* if for every $\tau \geq 0$, $\Phi(t, \tau)$ converges to a matrix with equal rows as $t \to \infty$. Similarly, the occurrence of multiple consensus as desired in Q.1

is equivalent to another property of underlying chain $\{A(t)\}$ called *class-ergodicity*. Chain $\{A(t)\}$ is *class-ergodic* if for every $\tau \geq 0$, $\lim_{t\to\infty} \Phi(t, \tau)$ exists, but with in general distinct rows.

The importance of Q.2, on the other hand, is not as clear as Q.1, at first sight, although it is arguably a more fundamental question. Q.2 is a natural concern when the underlying chain is endogenously generated in flocking models for example (see Cucker and Smale (2007)). However, remember that we deal with exogenous systems in this work.

Indeed, the question arises as to whether it is possible, for a limited number of key agents, to set their initial opinion/parameter assessment, in such a way that the (exogenously evolving) network converges to a global consensus. Such an issue is important in negotiations, or even the possible shaping or manipulation of public opinion by clever campaigning. In Chapter 5, the notion of *éminence grise coalition* is developed to begin to answer Q.2.

As mentioned above, distributed averaging algorithms can also be defined in discrete time as well. A general discrete time distributed averaging algorithm is defined by:

$$x(t+1) = A(t)x(t),\ t \geq 0, \tag{1.3}$$

where $\{A(t)\}$, the underlying chain of the network, is a chain of row-stochastic matrices, i.e., for each time instant $t \geq 0$, all elements of $A(t)$ are non-negative and the elements of each row of $A(t)$ add up to 1. The same questions can also be asked in the context of discrete time consensus algorithms, and, in general, similar results will hold for both. However, as it will become evident, the methods of proof can differ significantly.

1.3 Contributions

Consensus problems for distributed time-varying averaging algorithms have gained increasing attention in various research communities. One of the fundamental problems related to consensus is the *unconditional* occurrence of consensus or multiple consensus via distributed time-varying averaging algorithms, where by "unconditional", we mean irrespective of time or values at which states are initialized. Such problem turn out to be equivalent to ergodicity or class-ergodicity of the underlying chain of the system. Discovering necessary and/or sufficient conditions for ergodicity and class-ergodicity of a time-varying chain of matrices has been the aim of a significant body of literature. One of our two main objectives in this thesis has been to extend, as far as possible, the existing results regarding the proposed problem. Our contributions to this problem can be summarized as the following.

Balanced asymmetric chains. Balanced asymmetry is a property of a chain of stochastic matrices defined in discrete time. Balanced asymmetry is a hybrid of notions of subsymmetry Bolouki and Malhamé (2011a), and cut-balance Hendrickx and Tsitsiklis (2011) which were already very much discussed in the literature, and which are essentially point-wise verifiable properties of the underlying chain. We found, for balanced asymmetric chains, necessary and sufficient conditions for ergodicity and class-ergodicity based on a dynamic notion proposed by Touri and Nedić, that of absolute infinite flow, which is a property that can be verified only when considering the chain as a whole. The notion of balanced asymmetry, on its own, helped us subsume and generalize virtually all known convergence results thus far, albeit not convergence rate issues which is thoroughly a different concern.

Applications to known models. We showed that our techniques, which are employed to derive the convergence results on the exogenous averaging algorithms, together with the results themselves, can also be applied to some well-known nonlinear models, such as the Cucker-Smale model Cucker and Smale (2007) and the Hegselmann-Krause model Hegselmann and Krause (2002). These nonlinear models can be viewed as endogenous averaging algorithms, i.e., averaging algorithms with coefficients dynamically changing according to the evolution of states in the network.

Connection to Sonin's Decomposition-Separation Theorem. Our basic conviction that the theory of inhomogeneous Markov chains could help understand the convergence properties of consensus algorithms, which essentially depended on the properties of the underlying chain of the system, led us to employ the Sonin's Decomposition-Separation Theorem Sonin *et al.* (2008). The D-S Theorem together with the intuitions of Touri and Nedić (2014) about the importance of Kolmogorov's notion of absolute probability sequence, helped us obtain a meaningful generalization of the notion of absolute infinite flow to so-called *infinite jet-flow*.

A geometric framework. Attempts to understand the convergence mechanisms of inhomogeneous Markov chains led us to our first geometric insights of the Markov chain convergence as the intersection of decreasing convex hulls of appropriate sets of vertices. The vertices of each set correspond to the rows of the state transition matrix of the system at a certain time. This geometric interpretation was employed to extend our theorems, obtained based on the D-S Theorem, to the continuous time case.

Centered on these geometric insights, we then explored a question which is often raised for endogenously evolving consensus algorithms, such as the celebrated Cucker-Smale model Cucker and Smale (2007): Are there particular sets of initial conditions which will guarantee that the resulting consensus algorithm will converge unconditionally? Instead, the question is raised here for an exogenously generated sequence of update matrices. The geometric insights and the proposed question led us to defining the following notions and addressing their related issues which has been our second main objective in this thesis:

Éminence Grise Coalitions. It turns out that there exists a minimal subset of agents, which by mere setting of their initial conditions (under the rather idealized condition that they know where everyone else stands initially and the evolution of the network update chain), can steer the complete set of agents towards a global consensus. Such agents may be viewed as the consummate negotiators in a polarized environment, and we believe that such results are important for the study of opinion dynamics. A subset of agents with that property, even if it is not minimal, is called an *éminence grise coalition*, or simply EGC, in this thesis. We extensively investigated the size of the minimal EGC in a system.

Rank of a chain. We extended the notion of rank, as defined for a matrix, to a chain of matrices in both continuous and discrete time. We proved that the rank of the underlying chain of a multi-agent system is equal to the size of the minimal EGC that the system admits.

1.4 Manuscript Overview

In this thesis, we analyze via two different points of view, the asymptotic behavior of the states (or opinions) in a multi-agent system (an opinion network) with dynamics evolving via a predefined distributed averaging algorithm evolving in either discrete or continuous time. There are in general three items to consider when dealing with the limiting behavior of agents in such a system: (i) the initial time at which the system starts to update, (ii) the initial conditions (state values at the initial time), and (iii) the predefined update algorithm which is uniquely characterized by a time-varying chain of stochastic matrices.

In Chapters 3 and 4, our point of view towards achieving global or multiple consensus is via item (iii) and irrespective of items (i) and (ii). In other words, we seek chains of stochastic matrices that guarantee occurrence of global or multiple consensus in the network irrespective of the initial time or values at which states are initialized. More precisely, in Chapter 3, based on the notions of "balanced asymmetry" and "unbounded interactions graph" that we introduce in Section 3.3, and taking advantage of the notion of absolute infinite flow defined in Touri and Nedić (2012a), we obtain a class of chains guaranteeing global (or multiple) consensus to occur in the system. This class of (class-) ergodic chains is characterized in Section 3.4. In Chapter 4, we address the same objective via a different approach. This approach is based on exploring the connections between linear consensus algorithms and the Decomposition-Separation (DS) Theorem Sonin *et al.* (2008) of Markov chains. Using the D-S Theorem and the notion of "infinite jet-flow" defined in Section 4.3, we introduce a larger class of (class-) ergodic chains in Section 4.5, leading to theoretical results subsuming, to the best of our knowledge, all previous results in the literature as detailed in Section 4.6.

From a different point of view, in Chapter 5, we assume that there is absolutely no assumption on the so-called underlying chain of the network, i.e., item (iii). Instead, given an arbitrary time-varying underlying chain, and having fixed the initial time, we aim to identify sets of initial state values leading to consensus (conditional consensus). The set of such initial state vectors forms a vector space which has a close relationship with subgroups of agents that, in a way, act as potential hidden leaders of the network. Such a subgroup is called an “éminence grise coalition” (EGC) as explicitly defined in Section 5.3. The main results of Chapter 5 are concerned with the development of a geometric insights leading to the characterization of tight upper and lower bounds on the size of a minimal EGC.

CHAPTER 2

LITERATURE SURVEY

The notion of consensus arises in many research areas. In biology, consensus dynamics are at the heart of the emergent behavior of bird flocks, fish schools, etc. Couzin *et al.* (2005); Cucker and Smale (2007); Flierl *et al.* (1999). Consensus models can be employed to interpret, analyze, and predict flocking aggregation behavior. In robotics and control, consensus problems arise in relation to coordination objectives and cooperation of mobile agents (e.g. robots and sensors) Jadbabaie *et al.* (2003); Tsitsiklis *et al.* (1986), which are important issues in sensor networks for environmental applications or potentially space exploration. In economics, seeking an agreement on a common belief in a price system is another example of consensus. In Sociology, the emergence of a common language in primitive societies Cucker *et al.* (2004) is a collective behavior within a complex system, while in the area of social networks, consensus algorithms can shed light on the dynamics of opinion formation. Consensus algorithms also have a rich history within computer science community Lynch (1996), while formal study of consensus problems has been carried within the management science community (see DeGroot (1974) and references therein). Synchronization of coupled oscillators, i.e., reaching consensus on frequency of coupled oscillators, has been studied in physics, biophysics, and neurobiology for decades now Ermentrout (1992); Graver *et al.* (1984); Strogatz (2001). Since the body of work studying consensus problems is huge, in our literature review, we focus on those publications which have been most directly relevant to our work. We review, in the following, only a part of the literature on distributed averaging algorithms and their variations.

Distributed averaging algorithms, as a kind of consensus algorithms, can be defined in both discrete and continuous time domains. The discrete time version was first introduced in DeGroot (1974). The author considered a group of k individuals in a team or committee, where each individual has his own subjective probability distribution for the unknown value of some parameter θ. Individuals seek an agreement on a subjective distribution of θ. The algorithm is as the following. Let $F_i(t)$ denote the distribution of θ believed by individual i at discrete time $t \geq 0$. For any $t \geq 0$, individual i updates his distribution according to the formula $F_i(t+1) = \sum_j a_{ij} F_j(t)$, where interaction rates a_{ij}'s, $1 \leq i, j \leq k$ are non-negative constants satisfying $\sum_{j=1}^{k} a_{ij} = 1$, $\forall i = 1, \ldots, k$. Using Markov chains properties Doob (1953), DeGroot obtained a sufficient condition for convergence of individuals' distributions to the same distribution, which is the average of initial distributions, due to the symmetry of the model. Three year later, Chatterjee and Seneta (1977) considered the same consensus problem but with time-varying interaction rates. The authors found sufficient conditions for (global) consensus via backward product of stochastic matrices. Beside that, the issue of multiple consensus was also investigated in Chatterjee and Seneta (1977). Results of Chatterjee and Seneta (1977), regarding occurrence of consensus, were generalized in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986), i.e., more general conditions for consensus to occur were given. Unlike Chatterjee and Seneta (1977); DeGroot (1974), in the model considered in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986), communication links between individuals are not necessarily bidirectional. Sufficient conditions for convergence in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986) can briefly be described as non-vanishing interaction rates and repeated connectivity of communication graph.

In Jadbabaie *et al.* (2003), the authors proposed a simple consensus algorithm which remains one of the main motivations for the current extensive research on consensus. The model considered in Jadbabaie *et al.* (2003), which is a linearized version of the well-known Viscek model Vicsek *et al.* (1995), describes the evolution of a system of k agents moving in the plane with the same speed but different headings. Each agent updates his heading to the average of its own heading and the headings of its neighbors. Hence, in this model, if we define $N_i(t)$ and $n_i(t)$ as the set and the number of agent i's neighbors at time $t \geq 0$, we have $a_{ij}(t) = 1/(1 + n_i(t))$ if agent j is a neighbor of agent i at time t or $j = i$. Otherwise, $a_{ij}(t) = 0$. The authors of Jadbabaie *et al.* (2003) showed that consensus occurs if there exists an infinite sequence of contiguous, nonempty, bounded, time-intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each interval, the k agents are linked together (via a chain of neighbors). The authors employed Wolfowitz's Theorem Wolfowitz (1963) on product of stochastic matrices belonging to a finite set to show consensus. Therefore, the fact that a_{ij}'s belong to a finite set, and the boundedness of time-intervals $[t_i, t_{i+1})$, $i = 1, 2, \ldots$, play an important role in their proof. Although Jadbabaie *et al.* (2003) also addressed a leader-follower model in both discrete time and continuous time, its main discovery turned out to be a special case of Tsitsiklis (1984); Tsitsiklis *et al.* (1986) (see Bertsekas and Tsitsiklis (2007)). After Jadbabaie *et al.* (2003) was published, many authors aimed to generalize its consensus results by employing different techniques Blondel *et al.* (2005); Hendrickx (2008); Hendrickx and Blondel (2006); Li *et al.* (2004); Lorenz (2005); Moreau (2005). It was proved in Blondel *et al.* (2005) that for consensus to occur, a_{ij}'s are not required to belong to a finite set, and a uniform lower bound for non zero a_{ij}'s is sufficient. The authors employed an extension of Wolfowitz Theorem (see Blondel *et al.* (2005)). In Moreau (2005), the author showed that the boundedness of time intervals $[t_i, t_{i+1})$ is not necessary either. However, a uniform lower bound for non zero interaction rates still seemed to be necessary.

Recently, Hendrickx and Tsitsiklis (2013), and series of papers Touri and Nedić (2011, 2012a,b, 2014), independently generalized the previous results by introducing a class of chains of stochastic matrices, the so-called *cut-balanced* chains [1]. In the work of both groups, the multiple consensus problem was also considered. Although the authors of Touri and Nedić (2014) focus mainly on random chains, one can consider Corollary 4 of Touri and Nedić (2014) as, prior to this manuscript, the most general result for deterministic chains, where a uniform positive lower bound for all diagonal entries still appeared to be necessary.

Continuous time consensus protocols are more recent than their discrete time counterparts. According to Fax and Murray (2004); Jadbabaie *et al.* (2003); Lin *et al.* (2004); Moreau (2003, 2004); Moreau *et al.* (2003); Olfati-Saber and Murray (2004); Ren and Beard (2004); Ren *et al.* (2005a,b), a continuous time-varying consensus algorithm can be summarized as:

$$x_i'(t) = \sum_{j \neq i} a_{ij}(t)(x_j(t) - x_i(t)). \tag{2.1}$$

This algorithm can be considered as a linear version of more complex synchronization models Aeyels and Rogge (2004); Jadbabaie *et al.* (2004); Strogatz (2000). Some general results on consensus for agents using protocol (2.1) were given in Moreau (2003, 2004); Moreau *et al.* (2003); Ren and Beard (2004); Ren *et al.* (2005b). Sufficient conditions for consensus derived in Moreau (2004), is the existence of a uniform lower bound for non-zero interaction rates, and also a repeated connectivity of the communication graph, which is not bidirectional in general. This sufficient condition for consensus is weaker than the one obtained in Ren and Beard (2004); Ren *et al.* (2005b), since interaction rates do not necessarily belong to a finite set. Prior to the current work, and to the best of our knowledge, the cut-balance assumption appears to be the weakest assumption on the underlying chain of a linear continuous time consensus protocol to guarantee the occurrence of consensus (see Hendrickx and Tsitsiklis (2013)).

The literature on the subject of averaging algorithms has expanded in a number of directions. Here, we give a brief overview of a few of those research directions:

— **Bounded confidence models.** We have assumed, so far, that the set of each agent's neighbors in the network is determined by a predefined sequence of agent sets. In the bounded confidence model, an agent's neighbors, at any time instant, are those that lie in the agent's area of confidence. Two well-known algorithms, based on the explained neighboring rule, are the HK model Hegselmann and Krause (2002); Krause (1997, 2000) and the DW model Deffuant *et al.* (2000); Weisbuch *et al.* (2002). The reader is

1. The cut-balance property is referred as the balancedness property in Touri and Nedić (2014).

referred to Blondel *et al.* (2007, 2009) for an interesting study of the HK model, and Lorenz (2007) for a thorough survey on the both models and their extensions.

— **Gossip algorithms.** In a gossip model, the frequency of information exchange is controlled by an internal clock ticking according to a timing model. In each step, each agent transmits its information (state) to another agent which is chosen randomly from a predefined set of neighbors or the entire network. Then, the states are updated via a gossip protocol (see Karp *et al.* (2000); Kempe and Kleinberg (2002)), which can be viewed as an averaging algorithm. It is usually assumed that each agent can only handle a single incoming transmission at a time Boyd *et al.* (2006); Dimakis *et al.* (2010).

— **Models with imperfect information exchange.** In more realistic models, one has to take into account the presence of noise and disturbance in the system. There are several ways to represent noise and disturbance in a model. One instance is consensus algorithms over networks with noisy links (see Acemoglu *et al.* (2008); Aysal and Barner (2010); Cucker and Mordecki (2008); Huang and Manton (2007, 2009)). Another example is consensus algorithms under quantization effects. Quantized effects appear when the agents can only store and transmit quantized numbers instead of real numbers (see Aysal *et al.* (2008); Carli *et al.* (2007, 2008); Kashyap *et al.* (2007); Nedic *et al.* (2009)).

— **Models with time delays.** There exists, in reality, a communication delay in the network. The communication delay is described as the difference between the time at which some information is transmitted and the time at which updates according to the transmitted information take effect. Consensus in models with time delays has been widely studied in literature. For instance, see Lin and Jia (2009a,b) for second-order discrete and continuous models with time delays, Sun and Wang (2009b) for fixed time delay, and Sun and Wang (2009a) for time-varying delays.

CHAPTER 3

ARTICLE 1: LINEAR CONSENSUS ALGORITHMS BASED ON BALANCED ASYMMETRIC CHAINS

Sadegh Bolouki and Roland P. Malhamé

Submitted to *IEEE Transactions on Automatic Control*, September 2013

3.1 Abstract

Multi-agent consensus algorithms, with update steps based on so-called balanced asymmetric chains, are analyzed. For such algorithms, it is shown that (i) the empirical distribution of state values converges asymptotically, (ii) the occurrence of consensus or multiple consensus is directly related to the property of absolute infinite flow of the underlying update chain. An example is provided to illustrate the novelty of the results.

3.2 Introduction

Consensus problems in multi-agent systems have gained increasing attention in various research communities. Many of the consensus algorithms in the literature can be described by linear update equations:

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

where $x(t)$ is a vector whose components are, without loss of generality, the scalar agent states (the value of an unknown parameter or probability), and $A(t)$ for every discrete time instant $t \geq 0$ is a row-stochastic matrix, i.e., elements of $A(t)$ are all non-negative and each row of $A(t)$ sums to 1. For simplicity, in this paper, stochastic matrices refer to row-stochastic matrices. Matrix $A(t)$, $t \geq 0$, is referred to as the matrix of interaction rates. Distributed averaging algorithms were first introduced in DeGroot (1974). Later, Chatterjee and Seneta (1977) considered the same class of consensus problems with time-varying interaction rates. The authors found sufficient conditions for consensus by analyzing backward products of stochastic matrices. Results of Chatterjee and Seneta (1977) were generalized in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986), whereby more general sufficient conditions for consensus to occur were provided. Unlike Chatterjee and Seneta (1977); DeGroot (1974), in the model considered in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986), communication links between individuals are not necessarily bidirectional. Briefly stated, sufficient conditions for the convergence in Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986) are described by non-vanishing interaction rates and continuously repeated connectivity of the integrated communication graph. As an alternative model, Vicsek *et al.* (1995) considered a system of multiple agents moving in the plane with the same constant speed but different headings which are updated according to an averaging algorithm. Consensus was observed in simulations. The authors of Jadbabaie *et al.* (2003) analyzed a linearized version of the Viscek model and provided conditions under which consensus occurs. The authors showed that consensus occurs exponentially fast if there exists an infinite sequence of contiguous, non-empty, bounded time-intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each such interval, any pair of agents are linked together via a chain of neighbors. Following Jadbabaie *et al.* (2003), many authors tried to generalize the consensus results using different techniques (see Hendrickx and Tsitsiklis (2013) and references therein).

Recently, Hendrickx and Tsitsiklis (2013) and series of papers Touri and Nedić (2011, 2012a,b, 2014) independently generalized the previous results by introducing a class of chains of stochastic matrices, the so-called *cut-balanced* chains[1]. In the work of both groups, the multiple consensus problem was also considered. Although the focus of Touri and Nedić (2014) is mainly on random chains, one can consider Corollary 4 of Touri and Nedić (2014) as, by far, the most general consensus result for deterministic chains. However, a uniform positive lower bound for on-diagonal elements still appeared to be necessary.

In this note, by introducing a property of stochastic chains, herein called *balanced asymmetry*, we derive equivalent conditions for consensus and multiple consensus to occur in a class of multi-agent systems with dynamics (3.1). As will be shown, our results subsume Corollary 4 of Touri and Nedić (2014) since no uniform positive lower bound for non-zero interaction rates or self-interaction rates is required. In the process, we also establish that if the balanced asymmetry property is satisfied, the histogram of state values asymptotically converges to a fixed discrete distribution.

The rest of this paper is organized as follows. Essential notions that are required to state the main results are defined and illustrated in Section 3.3. Main results on consensus and multiple consensus are presented in Section 3.4. An example illustrating the specificity of our results is discussed in Section 3.5. Concluding remarks end the paper in Section 3.6.

1. The cut-balance property is referred as the balancedness property in Touri and Nedić (2014).

3.3 Useful Notions and Terminology

Throughout this article, we adopt the following notation. $\mathcal{V}$ is the set of agents and $N = |\mathcal{V}|$ is the number of agents. The letter t stands for the discrete time index. $x(t) = [x_1(t) \cdots x_N(t)]'$, $t \geq 0$, is the vector of agent states, where prime ($'$) indicates the transposition. For every $t \geq 0$, $(1_t, 2_t, \ldots, N_t)$ is a permutation of $\{1, 2, \ldots, N\}$ such that agent i_t $(1 \leq i \leq N)$ has the ith least state value among all agents at time t. $z_i(t) = x_{i_t}(t)$ is the ith least number among $x_1(t), \ldots, x_N(t)$. In particular, $z_1(t)$ and $z_N(t)$ are the state values of agents associated with the least and the greatest state values at time t respectively. $A(t)$, $t \geq 0$, is the matrix of interaction rates $a_{ij}(t)$, $1 \leq i, j \leq N$, and $\{A(t)\}$ is the underlying chain of the system of interest. The overbar ($\bar{\ }$) on a subset indicates complementation of the subset in the universal set of interest.

Definition 3.1. *Consider a multi-agent system with dynamics (3.1). By consensus in system (3.1), we mean that, irrespective of the time instant or values at which states are initialized, all $x_i(t)$'s, $i = 1, \ldots, N$, converge to identical values as t goes to infinity.*

We now define ergodicity according to Chatterjee and Seneta (1977). Let $\{A(t)\}$ be a chain of stochastic matrices. For $t > \tau \geq 0$, following Touri and Nedić (2011), denote $A(t, \tau) \triangleq A(t-1)A(t-2)\ldots A(\tau)$.

Definition 3.2. *Chatterjee and Seneta (1977) A chain $\{A(t)\}$ of stochastic matrices is said to be ergodic if and only if for every $\tau \geq 0$, $\lim_{t\to\infty} A(t, \tau)$ exists and is equal to a matrix with identical rows.*

It is possible to show that the occurrence of consensus in a multi-agent system is equivalent to ergodicity of the underlying chain of the system. This is how consensus and ergodicity are related. Beside consensus, there is another important notion, multiple consensus, that constitutes our focus in this work.

Definition 3.3. *For a multi-agent system with dynamics (3.1), multiple consensus is said to have occurred, if for every i, $1 \leq i \leq N$, $\lim_{t\to\infty} x_i(t)$ exists, irrespective of the time instant or values at which states are initialized.*

To formulate multiple consensus as a property of chains of stochastic matrices, we introduce class-ergodicity as follows.

Definition 3.4. *A chain $\{A(t)\}$ of stochastic matrices is class-ergodic if and only if for every $\tau \geq 0$, $A(t, \tau)$ converges as $t \to \infty$. Moreover, $i, j \in \mathcal{V}$ are said to belong to the same ergodic class of chain $\{A(t)\}$ if the ith and the jth rows of $\lim_{t\to\infty} A(t, \tau)$ are identical for every $\tau \geq 0$.*

Note that ergodic classes form an equivalence class on $\mathcal{V}$. Note also that if $\{A(t)\}$ in dynamics (3.1) is class-ergodic, multiple consensus occurs. The converse is also true since the ith column of $A(t,\tau)$ is equal to $x(t)$ when states vector x is initialized at time τ by the initial value e_i, in which all of the components equal to zero, but the ith one equal to 1. Therefore, multiple consensus occurs in a system with dynamics (3.1) if and only if chain $\{A(t)\}$ is class-ergodic. In the rest of this section, we provide essential notions that are employed to state our main results.

3.3.1 l_1-approximation

The following is an equivalent definition of l_1-approximation first defined in Touri and Nedić (2012a).

Definition 3.5. *Chain $\{A(t)\}$ is said to be an l_1-approximation of chain $\{B(t)\}$ if $\sum_{t=0}^{\infty} \|A(t)-B(t)\|$ is finite, where for convenience only, the norm, throughout this note, refers to the max norm, i.e., the maximum of the absolute values of the matrix elements.*

It is not difficult to show that l_1-approximation is an equivalence relation in the set of chains of stochastic matrices.

Proposition 3.1. *Touri and Nedić (2012a) Let chain $\{A(t)\}$ be an l_1-approximation of chain $\{B(t)\}$. Then, the two chains have the same ergodic classes. In particular, $\{A(t)\}$ is ergodic (class-ergodic) if and only if $\{B(t)\}$ is.*

3.3.2 Absolute Infinite Flow

Definition 3.6. *Touri and Nedić (2012b) A chain $\{A(t)\}$ of stochastic matrices is said to have the absolute infinite flow property if the following holds:*

$$\sum_{t=0}^{\infty}\Big(\sum_{i\in S(t+1)}\sum_{j\in \bar{S}(t)} a_{ij}(t) + \sum_{i\in \bar{S}(t+1)}\sum_{j\in S(t)} a_{ij}(t)\Big) = \infty, \tag{3.2}$$

for every sequence $\{S(t)\}$ of proper subsets of $\mathcal{V} = \{1,\ldots,N\}$ with the same cardinality. Note that if $A(t)$ is a matrix of order 1, i.e., $N = 1$, the absolute infinite flow property is trivially satisfied.

In Touri and Nedić (2012b), the authors showed that the absolute infinite flow property is a necessary condition for ergodicity. In addition, they prove the necessity and sufficiency of the absolute infinite property for ergodicity of chains of doubly stochastic matrices.

3.3.3 Balanced Asymmetry

Definition 3.7. *Chain $\{A(t)\}$ of stochastic matrices is said to be balanced asymmetric if there exists $M \geq 1$ such that for any two non-empty subsets $\mathcal{V}_1$ and $\mathcal{V}_2$ of $\mathcal{V}$ with the same cardinality:*

$$\sum_{i\in\mathcal{V}_1}\sum_{j\in\bar{\mathcal{V}}_2} a_{ij}(t) \leq M \sum_{i\in\bar{\mathcal{V}}_1}\sum_{j\in\mathcal{V}_2} a_{ij}(t),\ \forall t \geq 0. \tag{3.3}$$

We now mention the following non-trivial subclasses of balanced asymmetric chains.

1. *Chains of doubly stochastic matrices.* It can be shown that all chains of doubly stochastic matrices are balanced asymmetric with $M = 1$.
2. Chains possessing the following two properties:
   self-confidence. There exists $\delta > 0$ such that $a_{ii}(t) \geq \delta$, for every $i = 1, \ldots, N$, and $t \geq 0$.
   cut-balance. Hendrickx and Tsitsiklis (2013)Touri and Nedić (2014) There exists $K \geq 1$, such that for every $S \subset \mathcal{V}$:

$$\sum_{i\in S}\sum_{j\in\bar{S}} a_{ij}(t) \leq K \sum_{i\in\bar{S}}\sum_{j\in S} a_{ij}(t),\ \forall t \geq 0. \tag{3.4}$$

   Indeed, inequalities (3.4) and (3.3) are equivalent when $\mathcal{V}_1$ is identical to $\mathcal{V}_2$, while if $\mathcal{V}_1 \neq \mathcal{V}_2$, then $\mathcal{V}_1 \cap \bar{\mathcal{V}}_2$ and $\bar{\mathcal{V}}_1 \cap \mathcal{V}_2$ are both non-empty. As a result, and given the assumed self-confidence property, both sums in inequality (3.3) are bounded below by δ. In addition, both sums are bounded above by $N-1$ for any non-empty $\mathcal{V}_i$, $i = 1, 2$. Thus, the chain is balanced asymmetric with $M = \max\{K, (N-1)/\delta\}$.

Remark 3.1. *Balanced asymmetry is a stronger condition than cut-balance although the latter, together with self-confidence, becomes stronger than the former.*

Remark 3.2. *For those chains that are l_1-approximation of balanced asymmetric chains, the absolute infinite flow property can be simplified as the following:*

$$\sum_{t=0}^{\infty}\sum_{i\in\bar{S}(t+1)}\sum_{j\in S(t)} a_{ij}(t) = \infty, \tag{3.5}$$

for any sequence $\{S(t)\}$ of subsets of $\mathcal{V}$, with the same cardinality. This can be easily seen by combining relations (3.2) and (3.3).

3.3.4 Unbounded Interactions Graph

The unbounded interactions graph induced by a chain is an important notion in this article, especially in the analysis of class-ergodicity. The following is the discrete time version of the definition of the unbounded interactions graph given in Hendrickx and Tsitsiklis (2013).

Definition 3.8. *For underlying chain $\{A(t)\}$ of linear algorithm (3.1), we form a directed graph $G_A = \{\mathcal{V}, E\}$, in which $(i,j) \in E$ if and only if $\sum_{t=0}^{\infty} a_{ij}(t) = \infty$. G_A is called the unbounded interactions graph induced by $\{A(t)\}$.*

Noticing that balanced asymmetry is a stronger condition than cut-balance, following a proof quite similar to that of Theorem 1 (b) of Hendrickx and Tsitsiklis (2013), one can establish the following proposition.

Proposition 3.2. *Let $\{A(t)\}$ be a chain with unbounded interactions graph G_A. If $\{A(t)\}$ is balanced asymmetric, every weakly connected component of G_A is strongly connected.*

According to Proposition 3.2, under the balanced asymmetry condition, the unbounded interactions graph can be partitioned into strongly connected components, herein called *islands*.

3.4 Convergence Results

Recalling the definition of $z_i(t)$ as the ith least number among $x_1(t), \dots, x_N(t)$, we first state a theorem on the limiting behavior of states in a multi-agent system associated with a balanced asymmetric chain.

Theorem 3.1. *Consider a multi-agent system with dynamics (3.1). Assume that chain $\{A(t)\}$ is balanced asymmetric. Then, $\lim_{t\to\infty} z_i(t)$ exists for every $i \in \mathcal{V}$.*

Proof. For future needs, we shall prove the existence of $\lim_{t\to\infty} z_i(t)$ for a more general case, i.e., when $\{A(t)\}$ is an l_1-approximation of a balanced asymmetric chain with an arbitrary fixed bound M. To this aim, we use a technique similar to the one we adopted previously in proving Theorem 2 of Bolouki and Malhamé (2011b). Note that this technique was also independently discovered by Hendrickx and Tsitsiklis (see Hendrickx and Tsitsiklis (2013)). According to the definition of $z_i(t)$, we have $z_1(t) \le z_2(t) \le \cdots \le z_N(t)$, $\forall t \ge 0$. Moreover, since agent states are updated via a convex combination of their current states, $z_1(t)$ is a non-decreasing function of t, and $z_N(t)$ is a non-increasing function of t. Thus,

$$z_1(0) \le z_i(t) \le z_N(0), \forall i \in \mathcal{V}, \forall t \ge 0. \tag{3.6}$$

As a result, both $z_i(t)$ and $x_i(t)$ are uniformly bounded from above and below. Defining $L \triangleq z_N(0) - z_1(0)$, we have:

$$x_i(t) - x_j(t) \le L, \forall i, j \in \mathcal{V}, \forall t \ge 0. \tag{3.7}$$

Now, let $\{B(t)\}$ be a balanced asymmetric chain that is an l_1-approximation of $\{A(t)\}$. Let $A(t) = B(t) + P(t)$, $\forall t \geq 0$. Denote $\|P(t)\| \triangleq p_t$, $t \geq 0$, and $p'_t \triangleq \sum_{k=0}^{t-1} p_k$, $t > 0$ with $p'_0 = 0$. Note that p'_t remains bounded, according to the definition of l_1-approximation. Set $M' = 2M$, and recalling $L \triangleq z_N(0) - z_1(0)$, define function $s_r(t)$ for every r, $1 \leq r \leq N$, by:

$$s_r(t) \triangleq \sum_{i=1}^{r} M'^{-i}(z_i(t) + Np'_tL). \tag{3.8}$$

In the following, we show that $\lim_{t\to\infty} s_r(t)$ exists for every $r = 1, \ldots, N$. Since s_r is a linear combination of z_i's with bounded coefficients, and p'_t is bounded, it is bounded. Moreover,

$$s_r(t+1) - s_r(t) \geq M'^{-N} \sum_{k=1}^{r-1} \left[\left(\sum_{i=k+1}^{N} \sum_{j=1}^{k} b_{i_{t+1}j_t} \right) (z_{k+1}(t) - z_k(t)) \right] \geq 0 \tag{3.9}$$

(see (29) and the argument leading to (29) in Bolouki and Malhamé (2012a) for details). Hence, $s_r(t)$ is non-decreasing. From boundedness and monotonic increasing behavior of s_r, we obtain that $\lim_{t\to\infty} s_r(t)$ exists for every $r = 1, \ldots, N$. Furthermore, defining $s_0 \equiv 0$, (3.8) implies:

$$z_i(t) = M'^i(s_i(t) - s_{i-1}(t)) - Np'_tL. \tag{3.10}$$

Thus, the convergence of z_i's is immediately implied from the convergence of s_i, s_{i-1}, and p'_i. $\square$

The convergence of $z_i(t)$'s in Theorem 3.1 implies that the histogram of state values asymptotically converges to a fixed discrete distribution. In the next two theorems, we address the issues of consensus (ergodicity) and multiple consensus (class-ergodicity).

Theorem 3.2. *If chain $\{A(t)\}$ is balanced asymmetric, then $\{A(t)\}$ is ergodic if and only if it has the absolute infinite flow property.*

Proof. The necessity of the absolute infinite flow property has been proved in Touri and Nedić (2012b). Here, we show that if chain $\{A(t)\}$ has the absolute infinite flow and the balanced asymmetry properties, then $\{A(t)\}$ is ergodic, or equivalently, consensus occurs in system (3.1). Without loss of generality, we assume that states are initialized at $t = 0$. The main part of the proof is common with the proof of Theorem 3.1. According to Theorem 3.1, we know that $\lim_{t\to\infty} z_i(t)$ exists for every $i \in \mathcal{V}$. Let us define $\forall i \in \mathcal{V}$: $z_i^* = \lim_{t\to\infty} z_i(t)$. From the definition of z_i's, we have:

$$z_1^* \leq z_2^* \leq \cdots \leq z_N^*. \tag{3.11}$$

Since $z_1(t)$ and $z_N(t)$ are respectively the least and the greatest values of states at time t, consensus occurs if and only if $\lim_{t\to\infty}(z_N(t) - z_1(t)) = 0$, or equivalently $z_1^* = z_N^*$. Assume that this does not happen, or equivalently, $z_1^* < z_N^*$. We aim to show that applying the absolute infinite flow property in inequality (3.9) when $r = N$ (while identifying a with b in (3.9)), leads to an unbounded $s_N(t)$, which would be a contradiction. Since $z_1^* < z_N^*$, from inequalities (3.11), we conclude that there exists p, $1 \le p \le N-1$, such that $z_p^* < z_{p+1}^*$. If we define $\epsilon \triangleq (z_{p+1}^* - z_p^*)/2 > 0$, there exists $T \ge 0$ such that:

$$z_{p+1}^*(t) - z_p^*(t) > \epsilon, \forall t \ge T. \tag{3.12}$$

On the other hand, for balanced asymmetric chains, the absolute infinite flow property reduces to (3.5). From (3.5), we conclude that for any sequence $S(t)$ of subsets of $\mathcal{V}$ of the same cardinality:

$$\sum_{t=T}^{\infty} \sum_{i\in\bar{S}(t+1)} \sum_{j\in S(t)} a_{ij}(t) = \infty, \tag{3.13}$$

since $\sum_{t=0}^{T-1} \sum_{i\in\bar{S}(t+1)} \sum_{j\in S(t)} a_{ij}(t)$ is finite. If in (3.13) we set $S(t) = \{1_t, 2_t, \ldots, r_t\}$, we obtain:

$$\sum_{t=T}^{\infty} \sum_{i=r+1}^{N} \sum_{j=1}^{r} a_{i_{t+1}j_t} = \infty. \tag{3.14}$$

On the other hand, we note that according to the proof of Theorem 3.1, $\lim_{t\to\infty} s_r(t)$ exists for every $r = 1, \ldots, N$. Therefore, we can write:

$$\lim_{t\to\infty} s_r(t) - s_r(0) = \sum_{t=0}^{\infty} \left(s_r(t+1) - s_r(t)\right). \tag{3.15}$$

Relations (3.15) and (3.9) yield:

$$\begin{aligned} \lim_{t\to\infty} s_r(t) - s_r(0) &\ge \textstyle\sum_{t=0}^{\infty} \left\{M'^{-N} \sum_{k=1}^{r-1} \left[\left(\sum_{i=k+1}^{N} \sum_{j=1}^{k} a_{i_{t+1}j_t}\right)(z_{k+1}(t) - z_k(t))\right]\right\} \\ &= M'^{-N} \textstyle\sum_{k=1}^{r-1} \left[\sum_{t=0}^{\infty} \left(\sum_{i=k+1}^{N} \sum_{j=1}^{k} a_{i_{t+1}j_t}\right)(z_{k+1}(t) - z_k(t))\right]. \end{aligned} \tag{3.16}$$

Setting $r = N$ we obtain:

$$\lim_{n\to\infty} s_N(t) - s_N(0) \ge M'^{-N} \sum_{k=1}^{N-1} \left[\sum_{t=0}^{\infty} \left(\sum_{i=k+1}^{N} \sum_{j=1}^{k} a_{i_{t+1}j_t}\right)(z_{k+1}(t) - z_k(t))\right]. \tag{3.17}$$

From the above inequality, recalling that $z_{k+1}(t) \ge z_k(t)$, and keeping only terms corresponding to $k = p$ and $t \geq T$ in the RHS, we obtain:

$$\lim_{t\to\infty} s_N(t) - s_N(0) \geq M'^{-N} \sum_{t=T}^{\infty} \left(\sum_{i=p+1}^{N} \sum_{j=1}^{p} a_{i_{t+1}j_t} \right) (z_{p+1}(t) - z_p(t)) \,. \tag{3.18}$$

Inequalities (3.12) and (3.18) imply:

$$\lim_{t\to\infty} s_N(t) - s_N(0) \geq \epsilon . M'^{-N} \sum_{t=T}^{\infty} \sum_{i=p+1}^{N} \sum_{j=1}^{p} a_{i_{t+1}j_t}. \tag{3.19}$$

From (3.14), we know that the RHS of inequality (3.19) is unbounded. Thus, the LHS is unbounded, and so is $s_N(t)$, which is a contradiction. This completes the proof. □

Theorem 3.3. *Let chain $\{A(t)\}$ be balanced asymmetric. Then, $\{A(t)\}$ is class-ergodic if and only if the absolute infinite flow property holds over each island of the unbounded interactions graph induced by $\{A(t)\}$. Furthermore, in case of class-ergodicty, the islands form the ergodic classes of $\{A(t)\}$.*

Proof. To prove the sufficiency of the condition, we adopt the same technique as used in Touri and Nedić (2012a) and form a new chain $\{B(t)\}$ of the unbounded interactions graph G_A by eliminating interactions between each agent within an island and agents of other islands at all times. From the definition of islands, it is immediately implied that $\{B(t)\}$ is an l_1-approximation of $\{A(t)\}$. According to Proposition 3.1, it suffices to prove that $\{B(t)\}$ is class-ergodic. The system with $\{B(t)\}$ as its underlying chain can be decomposed into subsystems corresponding to islands, as there is no communication between islands at all. It is straightforward to verify that each subchain of $\{B(t)\}$ corresponding to a subsystem is an l_1-approximation of a balanced asymmetric chain with the absolute infinite flow property. Thus, Theorem 3.2 and Proposition 3.1 imply that each subchain is ergodic, and as a result, $\{B(t)\}$ is class-ergodic.

We now prove the converse property. More specifically, we assume that $\{A(t)\}$ is class-ergodic and balanced asymmetric, and prove that the absolute infinite flow property holds inside each island of G_A. Once again we form chain $\{B(t)\}$ from $\{A(t)\}$ by eliminating all interactions between agents of distinct islands. Since $\{B(t)\}$ is an l_1-approximation of $\{A(t)\}$, Proposition 3.1 implies that $\{B(t)\}$ is class-ergodic as well. It is sufficient now to show that the absolute infinite flow property holds inside islands of the unbounded interactions graph induced by chain $\{B(t)\}$. Define subchains of $\{B(t)\}$ corresponding to islands. We shall show that each island subchain is ergodic. Thus, consider an arbitrary initial state for each subsystem and by concatenating these states, form an initial vector $y(0)$ for the original system:

$$y(t+1) = B(t)y(t),\ t \geq 0. \tag{3.20}$$

Since $\{B(t)\}$ is assumed class-ergodic, multiple consensus occurs in system (3.20). Let I be an arbitrary island. We wish to show that agents of I belong to the same consensus cluster. Assume that, on the contrary, there exists an island I containing agents corresponding to distinct consensus clusters. We proceed with the exact same proof of Theorem 3.2, identifying this time y with x in the theorem, and taking advantage of inequality (3.19) by setting p as follows: since members of island I do not belong to the same cluster, I can be partitioned into non-empty I_1 subsets and $\bar{I}_1$ such that

$$\lim_{t\to\infty} y_i(t) < \lim_{t\to\infty} y_j(t),\ \forall i \in I_1,\ j \in \bar{I}_1. \tag{3.21}$$

Recalling that $\{B(t)\}$ is an l_1-approximation of a balanced asymmetric chain, the ordered limits $\{z_k^*\}_{1\leq k\leq N}$ in Theorem 3.1 exist. Set p equal to the maximum index k such that:

$$z_k^* \leq \max\{\lim_{t\to\infty} y_i(t) | i \in I_1\}, \tag{3.22}$$

and follow steps (3.15) to (3.19) in the proof of Theorem 3.2. Since, by the definition of the island I,

$$\sum_{t=0}^{\infty} \sum_{i\in\bar{I}_1, j\in I_1} b_{ij}(t) = \infty, \tag{3.23}$$

the RHS of inequality (3.19) is unbounded as in the proof of Theorem 3.2, which is a contradiction. Therefore, all agents contained in every island end up in the same consensus cluster. Since the initial state was arbitrary, we obtain that every subchain is ergodic. From ergodicity and balanced asymmetry of each subchain, we conclude that the absolute infinite flow property holds for each subchain, i.e., inside each island. □

As a result of Theorem 3.3, the following result, stated and proved previously in Touri and Nedić (2014), provides a sufficient condition for class-ergodicity of a chain of row stochastic matrices. Recall definitions of self-confidence and cut-balance properties from Part 3.3.3.

Theorem 3.4. *If chain $\{A(t)\}$ is self-confident and cut-balanced, it is also class-ergodic.*

Proof. See Theorem 3 of Bolouki and Malhamé (2012b) for the proof. □

3.5 An Illustrative Example

We first note that most of known models, such as the HK model Hegselmann and Krause (2002), the JLM model Jadbabaie *et al.* (2003), and the Cucker-Smale model Cucker and

Smale (2007) are self-confident and cut-balanced. Therefore, from Theorem 3.4, they are class-ergodic. Further results on ergodicity of these models can be found in our earlier work Bolouki and Malhamé (2012b). In the rest of the section, using our Theorem 3.3, we discuss a less restricted version of the JLM model Jadbabaie *et al.* (2003), whereby self-confidence no longer holds. Indeed, the parameter considered in Jadbabaie *et al.* (2003) is the heading of each agent. If we write $\theta_i(t)$ as the heading of an arbitrary agent i at moment t, the classical JLM model describing evolution of headings can be formulated as:

$$\theta_i(t+1) = \frac{1}{d_i(t)} \sum_{j \in D_i(t)} \theta_j(t), \tag{3.24}$$

where $D_i(t)$ and $d_i(t)$ denote respectively the set of neighbors and their number for agent i at time t. It is also assumed that (i) $i \in D_i(t)$, $\forall t \geq 0$, which guarantees the model has the self-confidence property with $\delta = 1/N$, (ii) interaction links are undirected, which guarantees the cut-balance property, with bound $M = 2/N$, of the model. These two properties immediately result in the occurrence of multiple consensus with no further assumptions. However, we wish to investigate a the JLM model in which the self-confidence assumption, i.e., $i \in D_i(t)$, is removed. Note that the links are still assumed to be undirected. Therefore, for every $i, j \in \mathcal{V}$, $a_{ij}(t) = 1/d_i(t)$ if $j \in D_i(t)$, and $a_{ij}(t) = 0$ otherwise. Let undirected graph $G_t(\mathcal{V}, E_t)$ represent the interactions of the system at time $t \geq 0$. For a simple graph, a *2-factor* is defined as a spanning subgraph made of a collection of vertex-disjoint cycles of the graph. Although $G_t(\mathcal{V}, E_t)$ may not be a simple graph, as it may contain self-loops, we use the same definition of 2-factor, while treating self-loops as cycles of length 1.

Theorem 3.5. *For the JLM model (A.1), where i is not necessarily in $D_i(t)$, multiple consensus occurs if $G_t(\mathcal{V}, E_t)$ has a 2-factor for every $t \geq 0$. Moreover, in case of multiple consensus, consensus clusters are the islands of the unbounded interactions graph induced by the underlying chain of the system.*

Proof. Assume that G_t has a 2-factor for every $t \geq 0$. For every $S \subset \mathcal{V}$, let $D_S(t)$, $t \geq 0$, denote the set of agents each of which is connected to at least one member of S at time t. It is easy to verify that the existence of a 2-factor in G_t implies that for every $S \subset \mathcal{V}$, either $D_S(t) = S$ or $|D_S(t)| > |S|$. We wish to take advantage of Theorem 3.3 to establish the result by showing that the balanced asymmetry property is satisfied. Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be two subsets of $\mathcal{V}$ with the same cardinality. If $D_{\mathcal{V}_1}(t) = \mathcal{V}_2$, then $\mathcal{V}_1 = \mathcal{V}_2$. Therefore, $D_{\mathcal{V}_2}(t) = \mathcal{V}_1$ as well. Thus, inequality (3.3) holds for any M, as both sides are zero. If $D_{\mathcal{V}_1}(t) \neq \mathcal{V}_2$, then $D_{\mathcal{V}_2}(t) \neq \mathcal{V}_1$ as well. Therefore, both summations in (3.3) are non-zero. Moreover, the summations are bounded below by $1/N$ and bounded above by N (total sum of interaction rates). Thus, (3.3) is satisfied for $M = N^2$. To complete the proof, keeping Theorem 3.3 in mind, we show that the absolute infinite flow property holds over each island of the unbounded interactions graph induced by the underlying chain of the system. Let I be an arbitrary island. We know that there is some finite time $T > 0$ past which all islands are isolated since the interaction weights, whenever non-zero, are bounded below by $1/N$. Let $\{S(t)\}_{t\geq 0}$ be an arbitrary sequence of equal cardinality subsets of $\mathcal{V}$ inside I. Since I is isolated for $t > T$, $D_{S(t)}(t) \subset I$ for $t > T$. On the other hand, the existence of a 2-factor in G_t implies $D_{S(t)}(t) = S(t)$ or $D_{S(t)}(t)| > |S(t)|$. The absolute infinite flow property over I is now proved considering the following two cases:

Case I: $D_{S(t+1)}(t+1) \neq S(t)$ occurs infinitely often. As a result, $\bar{S}(t) \cap D_{S(t+1)}(t+1) \neq \emptyset$ happens infinitely many times after time T, and every time $\sum_{i\in S(t+1)} \sum_{j\in I\setminus S(t)} a_{ij}(t)$ is bounded below by $1/N$. This leads to the satisfaction of (3.2) over I.

Case II: $D_{S(t+1)}(t+1) \neq S(t)$ does not occur infinitely often. Thus, there exists a finite time $T_1 > 0$ such that $D_{S(t+1)}(t+1) = S(t)$, $\forall t > T_1$. Therefore, $S(t+1) = S(t) = S$, $\forall t > T_1$. Otherwise, $|S(t+1)| < |S(t)|$ which is impossible by the assumption on the sequence cardinality. In this case, the absolute flow property holds due to the connectivity of island I. Otherwise, S would be an island inside I, which would be a contradiction. $\square$

Noticing that the existence of a single island in Theorem 3.5 results in consensus, we have the following corollary.

Corollary 3.1. *For the JLM model (A.1), where i is not necessarily in $D_i(t)$, consensus occurs if both the followings are satisfied: (i) $G_t(\mathcal{V}, E_t)$ has a 2-factor for every $t \geq 0$, (ii) there exist an infinite sequence of non-empty, bounded time-intervals $[t_i, t_{i+1})$, $i \geq 0$, starting at $t_0 = 0$, with the property that across each such interval, any pair of agents are linked together via a chain of neighbors.*

Remark 3.3. *If we made the self-confidence assumption in the JLM model, there would exist a trivial 2-factor (consisting of N disjoint self-loops) in $G_t(\mathcal{V}, E_t)$. As a result, the occurrence of multiple consensus would immediately be implied from Theorem 3.5.*

Remark 3.4. *One can define a 2-factor of a directed graph as a spanning subgraph made of a collection of vertex-disjoint directed cycles of the graph. By this definition, Theorem 3.5, with its current proof, also holds for the case in which $G_t(\mathcal{V}, E_t)$ is directed.*

3.6 Conclusion

In this note, we have focused on a class of linear distributed averaging algorithms in discrete time, such that the underlying non-homogeneous update Markov chain satisfies a property called balanced asymmetry. Under the balanced asymmetry assumption, we established that, asymptotically, states of agents involved in the consensus algorithm keep taking their values within a fixed set of limiting values of cardinality at most N.

We then considered the unbounded interactions graph and its islands as introduced in Hendrickx and Tsitsiklis (2013) for continuous time consensus algorithms. under the balanced asymmetry assumption, we obtained a necessary and sufficient condition for the above limiting values to become limits of individual agent states. We established that the number of potential consensus clusters is equal to the number of islands, and consensus over an island occurs if and only if the absolute infinite flow property (Touri and Nedić Touri and Nedić (2012b)) holds over that island. Finally, we displayed the applicability of our results to a number of well-known consensus models in the literature and developed a generalization of the JLM model requiring the tools in this note for its analysis. In future work, we shall investigate the impact of the number of agents increasing to infinity on all of our results.

CHAPTER 4

ARTICLE 2: CONSENSUS ALGORITHMS AND THE DECOMPOSITION-SEPARATION THEOREM

Sadegh Bolouki and Roland P. Malhamé

Submitted to *IEEE Transactions on Automatic Control*, September 2014

4.1 Abstract

Convergence properties of time inhomogeneous Markov chain based discrete and continuous time linear consensus algorithms are analyzed. Provided that a so-called infinite jet flow property is satisfied by the underlying chains, necessary conditions for both consensus and multiple consensus are established. A recent extension by Sonin of the classical Kolmogorov-Doeblin decomposition-separation for homogeneous Markov chains to the inhomogeneous case is then employed to show that the obtained necessary conditions are also sufficient when the chain is of class $\mathcal{P}^*$, as defined by Touri and Nedić. It is also shown that Sonin's theorem leads to a rediscovery and generalization of most of the existing related consensus results in the literature.

4.2 Introduction

Linear consensus algorithms and their convergence properties have gained increasing attention in the past decade. They were first introduced in DeGroot (1974), where the author considered the case when the interactions rates between any two agents are time-invariant. Later, more general cases were considered in Bertsekas and Tsitsiklis (1989); Blondel *et al.* (2005); Chatterjee and Seneta (1977); Hendrickx (2008); Hendrickx and Blondel (2006); Jadbabaie *et al.* (2003); Li *et al.* (2004); Moreau (2005); Tsitsiklis (1984); Tsitsiklis *et al.* (1986). The authors aimed at identifying sufficient conditions for consensus to occur, i.e., for states to asymptotically converge to the same value. Beside consensus, multiple consensus has been the subject of many articles, e.g., Bolouki and Malhamé (2012a); Hendrickx and Tsitsiklis (2013); Lorenz (2005); Touri and Nedić (2012a, 2014). Multiple consensus refers to the case when each agent state converges, as time grows large, to an individual limit which may or may not be different from the individual limits of other agent states. Considering the work on linear consensus algorithms, Bolouki and Malhamé (2012a); Hendrickx and Tsitsiklis (2013);

Touri and Nedić (2014) appear to provide the most general sufficient conditions for the occurrence of consensus or multiple consensus in a multi-agent system with dynamics described by a linear consensus algorithm.

In this paper, we deal with the limiting behavior of a general linear consensus algorithm in both discrete and continuous time. Let $\mathcal{V} = \{1, \ldots, N\}$ be the set of agents. In discrete time, we consider an N-agent system with linear update equation:

$$x(t+1) = A(t)x(t), \forall t \geq 0. \tag{4.1}$$

In (4.1), t indicates the discrete time index, $x(t) = [x_1(t) \cdots x_N(t)]'$, $t \geq 0$, is the vector of agent states, where prime ($'$) indicates the transposition, $A(t)$, $t \geq 0$, is the matrix of *interaction rates* $a_{ij}(t)$, $1 \leq i, j \leq N$, and $\{A(t)\}$ is the *underlying chain* or *transition chain* of the system, which is a chain of $(N \times N)$ row-stochastic matrices, i.e, for every $t \geq 0$, all elements of $A(t)$ are non-negative and each row of $A(t)$ sums up to 1. Throughout the paper, for simplicity, we refer to a row-stochastic matrix as a stochastic matrix. Since $A(t)$ is a stochastic matrix for every $t \geq 0$, sequence $\{x(t)\}$, by definition, forms a *backward Markov chain* with transition chain $\{A(t)\}$ (notice the evolution is described by a right hand multiplication by a column vector instead of the usual left hand multiplication by a row vector). Although we mainly focus on the discrete time case in this work, we shall extend our results to the continuous time case.

If all components of $x(t)$ asymptotically converge to the same limit, irrespective of the time index t or the values at which they are initialized, *unconditional global consensus*, or simply, *unconditional consensus*, is said to occur. Furthermore, if there exists a fixed partition of the N agents such that unconditional consensus occurs for the corresponding subvectors of $x(t)$, then *unconditional multiple consensus* is said to occur. The subsets in the partition are then said to form consensus clusters. It is well known that under dynamics (4.1), unconditional consensus is equivalent to *ergodicity* of chain $\{A(t)\}$ (see Chatterjee and Seneta (1977)), i.e., the property that backward products converge to matrices with identical rows. Furthermore, Bolouki and Malhamé (2012a) and Touri and Nedić (2012b) establish that a consensus algorithm with update chain $\{A(t)\}$ will induce multiple consensus if $\{A(t)\}$ is so-called *class-ergodic*, i.e., for every $t_0 \geq 0$, the product $A(t)A(t-1) \cdots A(t_0)$ converges, as $t \to \infty$. For class-ergodic chains, set $\mathcal{V}$ can be partitioned into *ergodic classes*, whereby i, j in $\mathcal{V}$ belong to the same ergodic class if the difference between the ith and jth rows of matrix product $A(t)A(t-1) \cdots A(t_0)$ vanishes, as $t \to \infty$. Under multiple consensus, the agent indices within the ergodic classes are the same as those within consensus clusters.

Sonin, in his so-called Decomposition-Separation (D-S) Theorem Sonin *et al.* (2008), suggests an elegant and illuminating physical interpretation of the dynamics in (4.1), which we now report for completeness: Start with a forward propagating Markov chain with $(N \times N)$ transition matrices $P(t)$ and associated sequence of probability distribution vectors $m(t)$:

$$m'(t+1) = m'(t)P(t), \forall t \geq 0. \tag{4.2}$$

Interpret $m_i(t)$, $i \in \mathcal{V}$, $t \geq 0$, as the *volume* of some liquid, say water for example, in a cup i (out of N cups), at time $t \geq 0$, while $p_{ij}(t)m_i(t)$ is the volume of liquid transferred from cup i to cup j at time $t \geq 0$ (see Fig. 4.1).

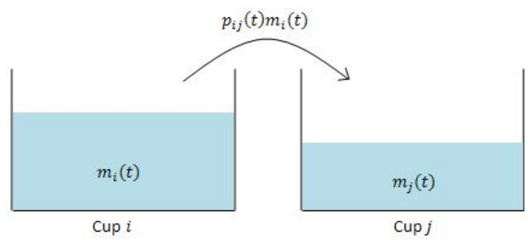

Figure 4.1 A physical interpretation of a Markov chain.

The volume of liquid in cup i, $\forall i \in \mathcal{V}$, is assumed to be initialized as $m_i(0)$ at time zero. Now, let $x_i(t)$, $i \in \mathcal{V}$, $t \geq 0$, be the *concentration* of a certain substance, such as sugar, alcohol, etc., within the liquid of cup i at time t. We first assume that the volume of each cup is non-zero at all times in order to make the concentration well-defined. Moreover, assume, for every $i \in \mathcal{V}$, that $x_i(t)$ is initialized as $x_i(0)$ at time zero. It is not difficult to show that, for every $i \in \mathcal{V}$, and $t \geq 0$:

$$x_i(t+1) = \frac{\sum_{j \in \mathcal{V}} p_{ji}(t)m_j(t)x_j(t)}{m_i(t+1)}. \tag{4.3}$$

Let:

$$x(t) \triangleq \left[x_1(t) \cdots x_N(t)\right]', \tag{4.4}$$

and $(N \times N)$ matrix $A(t)$, with elements $a_{ij}(t)$, $i, j \in \mathcal{V}$, be defined by:

$$a_{ij}(t) = p_{ji}(t)m_j(t)/m_i(t+1), \forall t \geq 0. \tag{4.5}$$

From (4.3), (4.4), and (4.5), we conclude that:

$$x(t+1) = A(t)x(t), \forall t \geq 0. \tag{4.6}$$

Since $A(t)$ is stochastic for every $t \geq 0$ (check (4.5)), $\{x(t)\}$ forms a backward Markov chain, with transition chain $\{A(t)\}$, as in (4.1). Removing the non-zero volume assumption, $\{A(t)\}$ is constructed in such a way that elements of $A(t)$, $t \geq 0$, satisfy:

$$m_i(t+1)a_{ij}(t) = m_j(t)p_{ji}(t), \forall i, j \in \mathcal{V}, \forall t \geq 0. \tag{4.7}$$

The D-S Theorem, Sonin *et al.* (2008), describes the limiting behavior of both $m(t)$ and $x(t)$, as t grows large. However, to take advantage of the D-S Theorem in a general consensus algorithm (4.1), one has to, first, answer the following questions: Starting with a backward propagating Markov chain generated by $\{A(t)\}$, is it always possible to find an associated forward propagating Markov chain, with distribution vector $\{m(t)\}$, generated by a transition chain $\{P(t)\}$, satisfying an equation of the form (4.7)? And how, if so? As discussed in this paper, due to the existence of a so-called *absolute probability sequence* for $\{A(t)\}$, as proved in fundamental work Kolmogoroff (1936), one could show the existence of the desired chains satisfying (4.7). More specifically, any absolute probability sequence $\{m(t)\}$ admitted by $\{A(t)\}$, would help construct a forward propagating sequence of transition matrices, via (4.7).

In this paper, it is established that, based on the D-S Theorem, all these previous results can be subsumed. Furthermore, inspired by Touri and Nedić (2012b), and recalling the notion of jets in Markov chains from Blackwell (1945), we introduce a property of chains resulting in necessary conditions for the unconditional occurrence of consensus or multiple consensus in (4.1). We also establish that, under an additional assumption, that is the chain being in the so-called Class $\mathcal{P}^*$ Touri and Nedić (2014), these necessary conditions also become sufficient.

In addition to the notation defined in the beginning of this section, we adopt the following notation throughout the paper. Letter t stands for either discrete or continuous time indices according to context. $\Phi(t,\tau)$, $t, \tau \geq 0$, represents the state transition matrix of the considered system, which can be defined in either the discrete time domain, as in (4.1), or the continuous time domain, as we will see later on. Moreover, $\Phi_i(t,\tau)$ and $\Phi_{i,j}(t,\tau)$, $1 \leq i, j \leq N$, denote the ith column and the (i,j)th element (intersection of ith row and jth column) of $\Phi(t,\tau)$ respectively, while $\Phi'_i(t,\tau)$ refers to the ith column of $\Phi'(t,\tau)$ (the prime acts first), which is also the transpose of the ith row of $\Phi(t,\tau)$. For an arbitrary vector $v \in \mathbb{R}^N$, and $1 \leq i \leq N$, v_i denotes the ith element of v. The overline $(^-)$ on a subset indicates complementation of the subset in the universal set of interest.

The rest of the paper is organized as follows. In Section 4.3, we state necessary conditions for class-ergodicity and ergodicity of a chain. The D-S Theorem, and its application in a general linear consensus algorithm, are discussed in Section 4.4. In Section 4.5, based on the

D-S Theorem, we analyze the convergence properties of chains in Class $\mathcal{P}^*$. It is shown, in Section 4.6, that this analysis leads to a generalization of most of the existing results in the literature on convergence properties of linear consensus algorithms. A geometric approach is introduced in Section 4.7 that applies to both discrete and continuous time consensus protocols. From the geometric framework built, we extend our analysis to the continuous time case in Section 4.8. Concluding remarks end the paper in Section 4.9.

4.3 The Infinite Jet-Flow Property

Inspired by Blackwell (1945), as reported in Sonin *et al.* (2008) and Touri and Nedić (2012b), in this section, we introduce a property of chains of stochastic matrices, herein called the *infinite jet-flow property*, leading to necessary conditions for ergodicity and class-ergodicity of the chain.

Definition 4.1. *For a given subset $\mathcal{V}'$ of finite set $\mathcal{V} = \{1, \ldots, N\}$, a jet J in $\mathcal{V}'$ is a sequence $\{J(t)\}$ of subsets of $\mathcal{V}'$. A jet J in $\mathcal{V}'$ is called proper if $\emptyset \neq J(t) \subsetneq \mathcal{V}'$, $\forall t \geq 0$ (see Fig. 4.2). Moreover, for a jet J, jet-limit J^* denotes the limit of the sequence $\{J(t)\}$, as t grows large, if it exists, in the sense that the sequence becomes constant after a finite time. When the elements of the sequence are all identical to a subset S of $\mathcal{V}$, the jet will be referred to as jet S.*

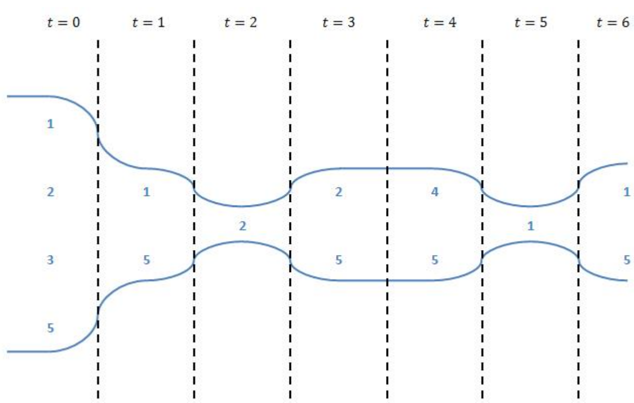

Figure 4.2 Example of a proper jet J in $\mathcal{V} = \{1, 2, 3, 4, 5\}$: $J(0) = \{1, 2, 3, 5\}$, $J(1) = \{1, 5\}$, $J(2) = \{2\}$, $J(3) = \{2, 5\}$, ...

Definition 4.2. *A tuple of jets $(J^1, \ldots, J^c)$ is a jet-partition of $\mathcal{V}$, if $(J^1(t), \ldots, J^c(t))$ is a partition of $\mathcal{V}$ for every $t \geq 0$.*

Definition 4.3. *Let chain $\{A(t)\}$ of stochastic matrices be given. For any two disjoint jets J^s and J^k in $\mathcal{V}$, $U_A(J^s, J^k)$, or simply $U(J^s, J^k)$, when no ambiguity results, denotes the total interactions between the two jets over the infinite time interval, as defined by:*

$$U(J^s, J^k) = \sum_{t=0}^{\infty} \Big[\sum_{i\in J^s(t+1)} \sum_{j\in J^k(t)} a_{ij}(t) + \sum_{i\in J^k(t+1)} \sum_{j\in J^s(t)} a_{ij}(t) \Big]. \tag{4.8}$$

Moreover, $U_{A(t)}(J^s, J^k)$, or simply, $U_t(J^s, J^k)$, denotes the interactions between the two jets at time t. More specifically,

$$U_t(J^s, J^k) = \sum_{i\in J^s(t+1)} \sum_{j\in J^k(t)} a_{ij}(t) + \sum_{i\in J^k(t+1)} \sum_{j\in J^s(t)} a_{ij}(t). \tag{4.9}$$

Definition 4.4. *The complement of jet J in $\mathcal{V}$, denoted by $\mathcal{V}\backslash J$, or simply, $\bar{J}$, is the jet defined by the set sequence $\{\mathcal{V}\backslash J(t)\}$.*

Definition 4.5. *A chain $\{A(t)\}$ of stochastic matrices is said to have the infinite jet-flow property over subset $\mathcal{V}'$ of $\mathcal{V}$ if, for every proper jet J in $\mathcal{V}'$, $U(J, \mathcal{V}'\backslash J)$ is unbounded. If $\mathcal{V}' = \mathcal{V}$, chain $\{A(t)\}$ is simply said to have the infinite jet-flow property.*

Example 4.1. *The following chain $\{A(t)\}_{t\geq 0}$ is an example of chains with the infinite jet-flow property:*

$$A(t) = \begin{bmatrix} 1 & 0 & 0 \\ 1-\frac{1}{t+1} & 0 & \frac{1}{t+1} \\ 0 & 0 & 1 \end{bmatrix}, \text{ if } t \text{ is even}, \tag{4.10}$$

and

$$A(t) = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{t+1} & 0 & 1-\frac{1}{t+1} \\ 0 & 0 & 1 \end{bmatrix}, \text{ if } t \text{ is odd}. \tag{4.11}$$

It is not easy, at this stage, to show that chain $\{A(t)\}$ defined by (4.10–4.11) has the infinite jet-flow property. In Lemma 4.2 stated later in the paper, we suggest a way to check the infinite jet-flow property of a chain that implies the infinite jet-flow property of $\{A(t)\}$ defined by (4.10–4.11).

Example 4.2. *Chain* $\{A(t)\}_{t\geq 0}$ *defined by:*

$$A(t) \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 1-\frac{1}{(t+1)^2} & 0 & \frac{1}{(t+1)^2} \\ 0 & 0 & 1 \end{bmatrix}, \text{ if } t \text{ is even,} \tag{4.12}$$

and

$$A(t) \;=\; \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{(t+1)^2} & 0 & 1-\frac{1}{(t+1)^2} \\ 0 & 0 & 1 \end{bmatrix}, \text{ if } t \text{ is odd,} \tag{4.13}$$

is an example of chains for which the infinite jet-flow property is not satisfied. More specifically, if we define jet J *by:*

$$J(t) \;=\; \begin{cases} \{1\} & \text{if } t \text{ is even} \\ \{1,2\} & \text{if } t \text{ is odd} \end{cases} \tag{4.14}$$

then we have:

$$U(J, \mathcal{V}\backslash J) \;=\; \sum_{t=0}^{\infty} \frac{1}{(t+1)^2} \;<\; \infty, \tag{4.15}$$

which shows that the infinite jet-flow property does not hold.

In the following proposition, we state a sufficient condition for the infinite jet-flow property to hold.

Definition 4.6. *Touri and Nedić (2014) For a chain* $\{A(t)\}$ *of stochastic matrices, we define its infinite flow graph,* $G_A(\mathcal{V}, E)$*, by an undirected graph of size* N*, such that:*

$$E = \{(i,j)|i,j \in \mathcal{V},\, i \neq j,\, \sum_{t=0}^{\infty}(a_{ij}(t) + a_{ji}(t)) = \infty\}. \tag{4.16}$$

The set of nodes of each connected component of $G_A(\mathcal{V}, E)$ *is called an island of* $\{A(t)\}$*. Moreover, chain* $\{A(t)\}$ *is said to have the infinite flow property if and only if* $G_A(\mathcal{V}, E)$ *is connected.*

The following theorem states a necessary condition for class-ergodicity of chain $\{A(t)\}$ of stochastic matrices.

Theorem 4.1. *A chain* $\{A(t)\}$ *of stochastic matrices is class-ergodic only if the infinite jet-flow property holds over each island of* $\{A(t)\}$.

Proof. Assume that, on the contrary, $\{A(t)\}$ is class-ergodic, yet some proper jet J, in an island I of $\{A(t)\}$, is such that $U_A(J, I\backslash J)$ is bounded. Recall, from Definition 4.1, that by a proper jet in I, we mean $\emptyset \neq J(t) \subsetneq I$, $\forall t \geq 0$. Since $U_A(J, I\backslash J)$ is bounded and I is an island of $\{A(t)\}$, we conclude that $U_A(J, \mathcal{V}\backslash J)$ is bounded as well. Recalling the definition of l_1-approximation from Touri and Nedić (2012a), a chain $\{B(t)\}$ is an l_1-approximation of chain $\{A(t)\}$ if:

$$\sum_{t=0}^{\infty} \|A(t) - B(t)\| < \infty, \tag{4.17}$$

where for convenience only, the norm refers to the *max norm*, i.e., the maximum of the absolute values of the matrix elements. We now form chain $\{B(t)\}$, an l_1-approximation of chain $\{A(t)\}$, by eliminating interactions between J and $\mathcal{V}\backslash J$ at all times. From (Touri and Nedić, 2012a, Lemma 1), it is known that l_1-approximations do not influence the ergodic classes of a chain. Therefore, $\{B(t)\}$ will remain class-ergodic with the same ergodic classes as $\{A(t)\}$. Also, the islands of B(t) are the same as those of A(t). On the other hand, $U_B(J, \mathcal{V}\backslash J) = 0$. Given two distinct arbitrary constants, α_1 and α_2, let states of a multi-agent system, $y_i(t)$, $i \in \mathcal{V}$, evolve via dynamics $y(t+1) = B(t)y(t)$, $\forall t \geq 0$, and be initialized at: $y_i(0) = \alpha_1$ if $i \in J(0)$, and $y_i(0) = \alpha_2$ otherwise. Since there is no interaction between J and $\mathcal{V}\backslash J$ at any time, we conclude that for every $t \geq 0$, we have: $y_i(t) = \alpha_1$ if $i \in J(t)$, and $y_i(t) = \alpha_2$ otherwise. Since $\{B(t)\}$ is class-ergodic, $\lim_{t\to\infty} y_i(t)$ exists for every $i \in \mathcal{V}$ and the consensual agents can be grouped into clusters sharing the same limit and forming an ergodic class. Since the elements in $\{J(t)\}$ are always associated with the same value of y for any t, they will asymptotically belong to a fixed limiting cluster S^* , namely agents for which $y_i(t)$ converges to α_1. Since J is a proper jet in I, we have: $\emptyset \neq S^* \subsetneq I$. Consider, now, jet S^* on island I. S^* is essentially the limiting jet J^* of J. Since the island structure is common for chains $\{A(t)\}$ and $\{B(t)\}$, we know that $U_B(J^*, I\backslash J^*)$ is unbounded. This is in contradiction with $U_B(J, I\backslash J) \leq U_B(J, \mathcal{V}\backslash J) = 0$, which completes the proof. □

Later in this paper, we shall establish the sufficiency of the infinite jet-flow property in Theorem 4.1, provided $\{A(t)\}$ is in Class $\mathcal{P}^*$, as defined in Touri and Nedić (2014). We now note that the infinite flow property of $\{A(t)\}$, which is a necessary condition for ergodicity of $\{A(t)\}$ according to Touri and Nedić (2011), is equivalent to the existence of a single island. Thus, Theorem 4.1 immediately results in the following corollary which is a necessary condition for ergodicity of chain $\{A(t)\}$ of stochastic matrices.

Corollary 4.1. *A chain $\{A(t)\}$ of stochastic matrices is ergodic only if it has the infinite jet-flow property.*

Corollary 4.1 provides a more restrictive necessary condition for ergodicity of a chain than Theorems 1 and 2 of Touri and Nedić (2012b). For instance, from Corollary 4.1, we conclude that the chain of Example 4.2 is not ergodic since it does not have the infinite jet-flow property. However, this cannot be concluded from Theorem 1 and 2 of Touri and Nedić (2012b).

On the other hand, we notice that the infinite jet-flow property is not sufficient for ergodicity. For instance, one can verify that the chain of Example 4.1 is not ergodic while the infinite jet-flow property holds.

Definition 4.7. *A jet J in $\mathcal{V}$ is called an independent jet if the total influence of $\bar{J}$ on J is finite over the infinite time interval, i.e.,*

$$\sum_{t=0}^{\infty} \sum_{i \in J(t+1)} \sum_{j \in \bar{J}(t)} a_{ij}(t) < \infty. \tag{4.18}$$

The following theorem, which is a generalization of Corollary 4.1, states yet another necessary condition for ergodicity of chain $\{A(t)\}$ of stochastic matrices.

Theorem 4.2. *A chain $\{A(t)\}$ of stochastic matrices is ergodic only if no two disjoint independent jets in $\mathcal{V}$ exist.*

Proof. Assume that on the contrary, there exist two disjoint independent jets J^1 and J^2 in $\mathcal{V}$. Similar to the proof of Theorem 4.1, form chain $\{B(t)\}$, an l_1-approximation of $\{A(t)\}$, by eliminating the influence of $\bar{J}^s$ on J^s, $s = 1, 2$, at all times. Recall that $\{A(t)\}$ and $\{B(t)\}$ will share the same ergodicity properties. Let states of a multi-agent system, $y_i(t)$, $1 \leq i \leq N$, evolve via dynamics $y(t+1) = B(t)y(t)$, $\forall t \geq 0$, and be initialized such that for every $i \in J^s(0)$ ($s = 1, 2$), $y_i(0) = \alpha_s$, where $\alpha_1 \neq \alpha_2$. Then, for every $t \geq 0$, we have: $y_i(t) = \alpha_s$, $\forall i \in J^s(t)$ ($s = 1, 2$). Since $\alpha_1 \neq \alpha_2$, consensus does not occur. Consequently, chain $\{B(t)\}$ and thus $\{A(t)\}$ could not possibly be ergodic. $\square$

As an example, for chain $\{A(t)\}$ of Example 4.1, jet $\{1\}$ and jet $\{3\}$ are two disjoint independent jets in $\mathcal{V} = \{1, 2, 3\}$. Thus, Theorem 4.2 implies that $\{A(t)\}$ is not ergodic.

Remark 4.1. *The following argument explains why Theorem 4.2 generalizes Corollary 4.1. Without the infinite jet-flow property, there exists a jet J such that $U(J, \mathcal{V}' \backslash J)$ is bounded. Thus, both jets J and $\mathcal{V} \backslash J$ are independent jets. On the other hand, jet J and $\mathcal{V} \backslash J$ are disjoint. Thus, infinite jet-flow is a weaker condition than the non-existence of any two disjoint independent jets.*

4.4 Relationship to the D-S Theorem

Consider a multi-agent system with states evolving according to linear algorithm (4.1), where $\{A(t)\}$ is a chain of stochastic matrices. Based on the work of Kolmogorov in Kolmogoroff (1936), we know that for every chain $\{A(t)\}_{t\geq 0}$, there exists a sequence $\{\pi(t)\}_{t\geq 0}$ of probability distribution vectors, called *absolute probability sequence*, such that

$$\pi'(t+1)A(t) = \pi'(t), \forall t \geq 0. \tag{4.19}$$

The transition chain $\{P(t)\}$ of the forward propagating chain associated with $\{A(t)\}$ and $\{\pi(t)\}$ as in (4.7), must be such that:

$$\pi_i(t)p_{ij}(t) = \pi_j(t+1)a_{ji}(t), \forall i, j \in \mathcal{V}, \forall t \geq 0. \tag{4.20}$$

More specifically, if $\pi_i(t) \neq 0$, then:

$$p_{ij}(t) = \pi_j(t+1)a_{ji}(t)/\pi_i(t), \tag{4.21}$$

while if $\pi_i(t) = 0$, for some i and $t \geq 0$, we choose $p_{ij}(t)$'s non-negative, arbitrarily such that:

$$\sum_{j=1}^{N} p_{ij}(t) = 1. \tag{4.22}$$

Note that in the former case ($\pi_i(t) \neq 0$), (4.22) is automatically satisfied, implying that $P(t)$ is a stochastic matrix for every $t \geq 0$. It is easy to see that:

$$\pi'(t)P(t) = \pi'(t+1), \forall t \geq 0. \tag{4.23}$$

Thus, $\{\pi(t)\}$ forms the probability distribution vector of an inhomogeneous forward propagating Markov chain. Let $V(J^s, J^k)$ denote the total *flow* between two arbitrary jets J^s and J^k in $\mathcal{V}$ over the infinite time interval as defined by:

$$\begin{aligned} V(J^s, J^k) = \textstyle\sum_{t=0}^{\infty} \Big[& \textstyle\sum_{i\in J^k(t)} \sum_{j\in J^s(t+1)} r_{ij}(t) \\ & + \textstyle\sum_{i\in J^s(t)} \sum_{j\in J^k(t+1)} r_{ij}(t) \Big], \end{aligned} \tag{4.24}$$

where

$$r_{ij}(t) = \pi_i(t)p_{ij}(t) = \pi_j(t+1)a_{ji}(t). \tag{4.25}$$

Value $r_{ij}(t)$ can be interpreted as the absolute joint probability of being in i at time t and j at time $t+1$. Recalling U from (4.8), we note that for every J^s, J^k in $\mathcal{V}$, $V(J^s, J^k) \leq U(J^s, J^k)$. Sonin, in his elegant work Sonin *et al.* (2008), characterizes the limiting behavior of the two sequences $\{\pi(t)\}$ and $\{x(t)\}$ (evolving via (4.1)) in the so-called D-S Theorem as the following.

Theorem 4.3. *(Sonin's D-S Theorem) There exists an integer c, $1 \leq c \leq N$, and a decomposition of $\mathcal{V}$ into jet-partition $(J^0, J^1, \ldots, J^c)$, $J^k = \{J^k(t)\}$, $0 \leq k \leq c$, such that irrespective of the particular time or state at which x_i's are initialized,*

(i) For every k, $1 \leq k \leq c$, there exist constants π_k^* and x_k^*, such that $\lim_{t\to\infty} \sum_{i\in J^k(t)} \pi_i(t) = \pi_k^*$ and $\lim_{t\to\infty} x_{i_t}(t) = x_k^*$ for every sequence $\{i_t\}$, $i_t \in J^k(t)$. Furthermore,

$$\lim_{t\to\infty} \sum_{i\in J^0(t)} \pi_i(t) = 0. \tag{4.26}$$

(ii) For every distinct k, s, $0 \leq k, s \leq c$: $V(J^k, J^s) < \infty$.

(iii) This decomposition is unique up to jets $\{J(t)\}$ such that $\lim_{t\to\infty} \sum_{i\in J(t)} \pi_i(t) = 0$ and $V(J, \mathcal{V}\backslash J) < \infty$ for any $\{\pi(t)\}$.

We shall take advantage of the Sonin's D-S Theorem to characterize the asymptotic behavior of a class of chains of stochastic matrices in the following section.

4.5 Convergence in Class $\mathcal{P}^*$

In this section, we apply Sonin's D-S Theorem to chains in class $\mathcal{P}^*$ as first defined in Touri and Nedić (2014).

Definition 4.8. *(Touri and Nedić, 2014, Definition 3) Chain $\{A(t)\}$ is said to be in class $\mathcal{P}^*$ if it admits an absolute probability sequence uniformly bounded away from zero, i.e., there exists $p^* > 0$ such that*

$$\pi_i(t) \geq p^*, \forall i \in \mathcal{V}, \forall t \geq 0. \tag{4.27}$$

For chains in Class $\mathcal{P}^*$, it is immediately implied that in the jet decomposition of the D-S Theorem, there is no jet J^0. Otherwise, $\lim_{t\to\infty} \sum_{i\in J^0(t)} \pi_i(t)$ would be bounded away from zero by at least p^*, which is in contradiction with the D-S Theorem. Therefore, there is a jet-partition of $\mathcal{V}$ into jets $J^1, \ldots, J^c$, such that for every $k = 1, \ldots, c$, $\lim_{t\to\infty} x_{i_t}(t) = x_k^*$, for every sequence $\{i_t\}$, where $i_t \in J^k(t)$. Thus, we have the following proposition for chains in Class $\mathcal{P}^*$.

Proposition 4.1. *Consider a multi-agent system with dynamics (4.1), where chain $\{A(t)\}$ is in Class $\mathcal{P}^*$. Then, the set of accumulation points of states is finite.*

Proof. Obvious if we note that $\{x_k^* | 1 \leq k \leq c\}$ form the set of accumulation points of states. $\square$

Lemma 4.1. *If $\{A(t)\} \in \mathcal{P}^*$, then for every two jets J^1 and J^2 in $\mathcal{V}$, $V(J^1, J^2) = \infty$ if and only if $U(J^1, J^2) = \infty$.*

Proof. The result is obvious if one notes that

$$p^* U(J^1, J^2) \le V(J^1, J^2) \le U(J^1, J^2). \tag{4.28}$$

□

Theorem 4.4. *A chain $\{A(t)\}$ in Class $\mathcal{P}^*$ is class-ergodic if and only if the infinite jet-flow property holds over each island of $\{A(t)\}$. In case of class-ergodicity of $\{A(t)\}$, islands are the ergodic classes of $\{A(t)\}$, and constitute the jet limits in the jet decomposition of $\{A(t)\}$. Moreover, these limits are attained in finite time.*

Proof. We first assume that chain $\{A(t)\}$ in $\mathcal{P}^*$ is class-ergodic. Then, Theorem 4.1 implies that the infinite jet-flow property holds over each island of the chain. We now show that if $\{A(t)\} \in \mathcal{P}^*$ is class-ergodic, islands are the ergodic classes of $\{A(t)\}$. Let us call an agent $i \in \mathcal{V}$, a *prime member* of jet J^k if $i \in J^k(t)$ for infinitely many times. Having defined the prime membership, there exists some Sonin's jet-decomposition of $\{A(t)\}$ such that each agent becomes the prime member of a unique jet. To obtain such a jet-decomposition, start with an arbitrary jet-decomposition and let any two jets with a common prime member merge. The merging process results in a Sonin's jet-decomposition with the desired property. Jets of such decomposition have the property that they become time-invariant after a finite time. Thus, the jet-limits exist for each jet and are ergodicity classes of $\{A(t)\}$. If i and j belong to the same jet-limit, they are in the same island since they are in the same ergodic class of $\{A(t)\}$ (Touri and Nedić (2012a), Lemma 2). Conversely, assume that i and j are neighbors in the infinite flow graph, i.e., $\sum_{t=0}^{\infty}(a_{ij}(t) + a_{ji}(t)) = \infty$. If i and j were to belong to different jet-limits J^{s^*}, J^{k^*}, then $U(J^s, J^k)$ would be unbounded. Thus, based on Lemma 4.1, $V(J^s, J^k)$ would be unbounded as well, which contradicts property (ii) in the D-S theorem. Therefore, every two neighbors in the infinite flow graph belong to the same jet-limit. Consequently, every i and j in the same island must be in the same jet-limit.

To prove the sufficiency, we assume that the infinite jet-flow property holds over each island. Let $(J^1, \ldots, J^c)$ be a Sonin's jet-decomposition, and $\lim_{t\to\infty} x_{i_t}(t) = x_k^*$ for every sequence $\{i_t\}$, where $i_t \in J^k(t)$ (for every k, $1 \le k \le c$). Let I be an arbitrary island. We aim to show that, for every $i \in I$, $\lim_{t\to\infty} x_i(t)$ exists. To this aim, keeping in mind that the aim is achieved is one of jets $J^1, \ldots, J^c$ contains island I after some finite time, we follow three steps. Pick an arbitrary jet J^k among $J^1, \ldots, J^c$.

Step 1: We show that, infinitely often, we have: $I \cap J^k(t) = \emptyset$ or $I \cap J^k(t) = I$, where $\emptyset$ denotes the empty set. Indeed, assume instead that this behavior occurs only a finite number

r of times, denoted $t_1, \ldots, t_r$. We form a proper jet J in I such that:

$$J(t) = I \cap J^k(t), \text{ if } t \neq t_i, 1 \leq i \leq r. \tag{4.29}$$

Since the infinite jet-flow property holds over I, $U(J, I \backslash J)$ is unbounded. On the other hand, except for a finite number of time indices $t = t_i$, $1 \leq i \leq r$, $U_t(J, I \backslash J) \leq U_t(J^k, \mathcal{V} \backslash J^k)$. This implies that $U(J^k, \mathcal{V} \backslash J^k)$ is unbounded, and, according to Lemma 4.1, so is $V(J^k, \mathcal{V} \backslash J^k)$. This is in contradiction with the D-S Theorem. Therefore, $I \cap J^k(t) = \emptyset$ or I happens infinitely many times. This means that either one or both of the events $I \cap J^k(t) = \emptyset$ and $I \cap J^k(t) = I$ occurs infinitely often.

Step 2: We show that there are at most a finite number of times such that $I \subseteq J^k(t)$ and $I \nsubseteq J^k(t+1)$. Indeed, denote:

$$\epsilon \triangleq \frac{1}{3} \min\{|x_s^* - x_l^*| \,|1 \leq s \neq l \leq c\}, \tag{4.30}$$

there exists $T_\epsilon \geq 0$ such that:

$$|x_i(t) - x_l^*| < \epsilon, \forall l = 1, \ldots, c, \forall i \in J^l(t), \forall t \geq T_\epsilon. \tag{4.31}$$

For some given $t \geq T_\epsilon$ assume that: $I \subseteq J^k(t)$ and $I \nsubseteq J^k(t+1)$. Then, there exists $i \in I$ such that $i \in J^k(t) \backslash J^k(t+1)$. In view of (4.1), (4.30), and (4.31), we then have:

$$|\sum_{j \notin J^k(t)} a_{ij}(t)(x_j(t) - x_i(t))| \geq \epsilon. \tag{4.32}$$

On the other hand,

$$\begin{aligned} &|\textstyle\sum_{j \notin J^k(t)} a_{ij}(t)(x_j(t) - x_i(t))| \\ &\qquad \leq \textstyle\sum_{j \notin J^k(t)} a_{ij}(t)|x_j(t) - x_i(t)| \\ &\qquad \leq L \textstyle\sum_{j \notin J^k(t)} a_{ij}(t), \end{aligned} \tag{4.33}$$

where

$$L \triangleq \max\{x_j(0) - x_i(0), |i, j \in \mathcal{V}\}. \tag{4.34}$$

Note that L remains an upper bound of $|x_j(t) - x_i(t)|$, $\forall t \geq 0$, since states are updated via a convex combination of previous states. Eqs. (4.32) and (4.33) imply:

$$\sum_{j \notin J^k(t)} a_{ij}(t) \geq \epsilon / L. \tag{4.35}$$

Therefore, since $i \in I$:

$$\sum_{l\in I}\sum_{j\notin I} a_{lj}(t) \geq \sum_{j\notin I} a_{ij}(t) \geq \sum_{j\notin J^k(t)} a_{ij}(t) \geq \epsilon/L. \tag{4.36}$$

Since $U(I,\mathcal{V}\backslash I) < \infty$, inequality (4.36) can only occur for finitely many times t. This shows that if $I \subseteq J^k(t)$ happens infinite times, then there exists T such that $I \subseteq J^k(t)$ for every $t \geq T$. Consequently, $\lim_{t\to\infty} x_i(t)$ exists, $\forall i \in I$, and is equal to x_k^*. Therefore, assume that for a fixed island I, $I \subseteq J^k(t)$ happens only a finite number of times for every k, $1 \leq k \leq c$. Thus, from the result of *Step 1*, $I \cap J^k(t) = \emptyset$ must happen infinite times, for every k, $1 \leq k \leq c$.

Step 3: We show that if $I \cap J^k(t) = \emptyset$ happens infinite times, for every k, $1 \leq k \leq c$, then, the following contradiction occurs: For every k, $1 \leq k \leq c$, there exists $T_k \geq 0$ such that $I \cap J^k(t) = \emptyset$, $\forall t \geq T_k$. The proof is established by induction on k. With no loss of generality, assume that $x_1^* < \cdots < x_k^*$. ($k = 1$): Recalling ϵ and T_ϵ from (4.30) and (4.31), assume that for a fixed $t \geq T_\epsilon$ we have $I \cap J^1(t) = \emptyset$ and $I \cap J^1(t+1) \neq \emptyset$. Thus, there exists $i \in I$ such that $i \in J^1(t+1)\backslash J^1(t)$. Therefore,

$$\sum_{j\in J^1(t)} |a_{ij}(t)(x_j(t) - x_i(t))| \geq \epsilon. \tag{4.37}$$

Noting that $J^1(t) \subseteq \mathcal{V}\backslash I$, by repeating steps (4.32)-(4.36), we conclude that there are at most finitely many times at which $I \cap J^1(t) = \emptyset$ and $I \cap J^1(t+1) \neq \emptyset$. This together with the fact that $I \cap J^1(t) = \emptyset$ happens infinite times, shows that there exists $T_1 \geq 0$ such that $I \cap J^1(t) = \emptyset$, $\forall t \geq T_1$.
$k-1 \to k$ ($1 < k \leq c$): Assume that for a fixed $t \geq \max\{T_l | 1 \leq l < k\}$, we have $I \cap J^k(t) = \emptyset$ and $I \cap J^k(t+1) \neq \emptyset$. Thus, there exists $i \in I$ such that $i \in J^k(t+1)\backslash J^k(t)$. Therefore,

$$\sum_{j\in\bigcup_{l=1}^k J^l(t)} |a_{ij}(t)(x_j(t) - x_i(t))| \geq \epsilon. \tag{4.38}$$

Once again, we note that $\bigcup_{l=1}^k J^l(t) \subseteq \bar{I}$, and repeat steps (4.32)-(4.36) to show that there exists $T_k \geq 0$ such that $I \cap J^k(t) = \emptyset$, $\forall t \geq T_k$. □

Corollary 4.2. *A chain $\{A(t)\} \in \mathcal{P}^*$ is ergodic if and only if it has the infinite jet-flow property.*

Since convergence of states occurs inside each jet J^k, $1 \leq k \leq c$, for multiple consensus to occur unconditionally (class-ergodicity of $\{A(t)\}$), it suffices that for each jet of the D-S Theorem jet decomposition, its jet-limit exists.

4.6 Relationship to Previous Work

4.6.1 Weakly Aperiodic Chains in Class $\mathcal{P}^*$

In this section of the paper, we see how the weak aperiodicity property, as defined in Touri and Nedić (2014), guarantees that the infinite jet-flow property holds over each island. In accordance with Touri and Nedić (2014), weak aperiodicity of a chain is defined as follows:

Definition 4.9. *A chain $\{A(t)\}$ of stochastic matrices is said to be weakly aperiodic if there exists $\gamma > 0$ such that for every distinct $i, j \in \mathcal{V}$ and each $t \geq 0$, there exists $l \in \mathcal{V}$ such that*

$$a_{li}(t).a_{lj}(t) \geq \gamma a_{ij}(t). \tag{4.39}$$

Lemma 4.2. *Let $\{A(t)\}$ be a chain of stochastic matrices in Class $\mathcal{P}^*$ that is weakly aperiodic. Then, the infinite jet-flow property holds over each island of $\{A(t)\}$. In particular, in presence of a single island, the infinite jet-flow property holds for chain $\{A(t)\}$.*

Proof. Let $\{A(t)\}$ be weakly aperiodic, I be an arbitrary island of $\{A(t)\}$, and J be an arbitrary jet in I. If jet-limit J^* exists, since I is a connected component of the infinite flow graph, $U(J^*, I\backslash J^*)$ is unbounded. Consequently, $U(J, I\backslash J)$ is unbounded and the lemma holds. Thus instead, assume that for jet J, the jet-limit does not exist. Therefore, for infinitely many times t, we must have: $J(t+1) \not\subseteq J(t)$. Let t be fixed and $J(t+1) \not\subseteq J(t)$. Thus, there exists $i \in J(t+1)\backslash J(t)$. From the weak aperiodicity property of $\{A(t)\}$ (see (4.39)), for every $j \in J(t)$, there exists $l \in \mathcal{V}$ such that:

$$\begin{aligned} \gamma a_{ij}(t) \leq a_{li}(t).a_{lj}(t) &\leq \min\{a_{li}(t), a_{lj}(t)\} \\ &\leq U_t(J, \mathcal{V}\backslash J), \end{aligned} \tag{4.40}$$

where U_t is defined in (4.9). The reason for the last inequality is that, whether $l \in J(t+1)$ or $l \notin J(t+1)$, one of $a_{li}(t), a_{lj}(t)$ appears in $U_t(J, \mathcal{V}\backslash J)$. Hence,

$$\sum_{j \in J(t)} \gamma a_{ij}(t) \leq |J(t)| U_t(J, \mathcal{V}\backslash J). \tag{4.41}$$

On the other hand,

$$\begin{aligned} \textstyle\sum_{j \in J(t)} \gamma a_{ij}(t) &= \gamma \textstyle\sum_{j \in J(t)} a_{ij}(t) \\ &= \gamma \left(1 - \textstyle\sum_{j \notin J(t)} a_{ij}(t)\right) \\ &\geq \gamma \left(1 - U_t(J, \mathcal{V}\backslash J)\right). \end{aligned} \tag{4.42}$$

Relations (4.41) and (4.42) imply:

$$U_t(J, \mathcal{V} \backslash J) \geq \gamma/(\gamma + |J(t)|) > \gamma/(\gamma + N). \tag{4.43}$$

Since (4.43) holds for infinitely many times t, $U(J, \mathcal{V} \backslash J) = \sum_{t=0}^{\infty} U_t(J, \mathcal{V} \backslash J)$ is unbounded, and so is $U(J, I \backslash J)$ (since J is a jet in I, and I is an island). $\square$

Theorem 4.4 and Lemma 4.2 immediately imply the following corollary which is the deterministic counterpart of Theorem 4 of Touri and Nedić (2014).

Corollary 4.3. *Every weakly aperiodic chain in Class $\mathcal{P}^*$ is class-ergodic.*

Note that an equivalent definition of weak periodicity is as follows.

Definition 4.10. *A chain $\{A(t)\}$ of stochastic matrices is weakly aperiodic if there exists $\gamma > 0$ such that for every distinct $i, j \in \mathcal{V}$ and each $t \geq 0$, there exists $l \in \mathcal{V}$ such that*

$$\min\{a_{li}(t), a_{lj}(t)\} \geq \gamma a_{ij}(t). \tag{4.44}$$

To achieve class-ergodicity under the $\mathcal{P}^*$ class assumption, the number of times in which an agent moves from a jet to another must be finite. Indeed, let

$$\epsilon \triangleq \frac{1}{3} \min\{|x_s^* - x_k^*| \, |1 \leq k \neq s \leq c\}. \tag{4.45}$$

Then, there exists T_ϵ such that for every $t \geq T_\epsilon$,

$$|x_i(t) - x_k^*| < \epsilon, \forall i \in J^k(t). \tag{4.46}$$

If agent i moves from a jet, say J^1, to another jet, say J^2, at time t ($i \in J^1(t) \cap J^2(t+1)$), we must have:

$$|\sum_{j \notin J^1(t)} a_{ij}(t)(x_j(t) - x_i(t))| \geq \epsilon. \tag{4.47}$$

On the other hand,

$$\begin{aligned} &|\textstyle\sum_{j \notin J^1(t)} a_{ij}(t)(x_j(t) - x_i(t))| \\ &\quad \leq \textstyle\sum_{j \notin J^1(t)} a_{ij}(t)|x_j(t) - x_i(t)| \\ &\quad \leq L \textstyle\sum_{j \notin J^1(t)} a_{ij}(t), \end{aligned} \tag{4.48}$$

where L is defined in Eq. (4.34). Eqs. (4.47) and (112) imply:

$$\sum_{j \notin J^1(t)} a_{ij}(t) \geq \epsilon/L. \tag{4.49}$$

Thus, there exists $j \notin J^1(t)$ such that

$$a_{ij}(t) \geq \frac{\epsilon}{L(N-1)}. \tag{4.50}$$

Now, from the definition of weak aperiodicity, we know that there exists $l \in \mathcal{V}$ such that $\min\{a_{li}(t), a_{lj}(t)\} \geq \gamma a_{ij}(t) \geq \gamma\epsilon/L(N-1)$. Note that i and j are in different jets at time t. Thus, l cannot be in the same jet with both i and j at time t. Therefore, at least one of $a_{li}(t), a_{lj}(t)$ indicates an interaction between a jet and its complement. Since both values are bounded below by $\gamma\epsilon/L(N-1)$, the sum of interactions between jets J^k's and their complements is at least $\gamma\epsilon/L(N-1)$ at time t. On the other hand, from the D-S Theorem, we now that the total sum of flows between jets and their complements is finite over the infinite time interval. Since $\{A(t)\}$ is of Class $\mathcal{P}^*$, the total sum of interactions between the jets and their complements must be finite as well. Hence, the number of times that the sum of interactions is at least $\gamma\epsilon/L(N-1)$, must be finite. Therefore, there are finite times in which an agent moves from a jet to another, and the jets become time-invariant after a finite time. It is straightforward to see that the time-invariant jets are connected components of the infinite flow graph.

4.6.2 Self-Confident and Cut-Balanced Chains

Definition 4.11. *Bolouki and Malhamé (2012a) A chain $\{A(t)\}$ of stochastic matrices is self-confident with bound δ if $a_{ii}(t) \geq \delta$, $\forall i \in \mathcal{V}$, $\forall t \geq 0$.*

Definition 4.12. *Hendrickx and Tsitsiklis (2013) A chain $\{A(t)\}$ of stochastic matrices is cut-balanced with bound K if for every $\mathcal{V}_1 \subseteq \mathcal{V}$ and $t \geq 0$:*

$$\sum_{i \notin \mathcal{V}_1} \sum_{j \in \mathcal{V}_1} a_{ij}(t) \leq K \sum_{i \in \mathcal{V}_1} \sum_{j \notin \mathcal{V}_1} a_{ij}(t). \tag{4.51}$$

Proposition 4.2. *Bolouki and Malhamé (2012a); Touri and Nedić (2014) If chain $\{A(t)\}$ is self-confident and cut-balanced, then it is class-ergodic and the islands form the ergodic classes of $\{A(t)\}$.*

Proof. Assume that $\{A(t)\}$ has self-confidence and cut-balance properties with bounds δ and K respectively. The chain being self-confident and cut-balanced, it in Class $\mathcal{P}^*$ (see (Touri and

Nedić, 2014, Theorem 7) where self-confidence is referred to as strong aperiodicity). Thus, from Theorem 4.4, it is sufficient to show that for an arbitrary island I and an arbitrary proper jet J in I, we have $U(J, I\backslash J) = \infty$ (that is the infinite jet flow property holds island-wise). Indeed, if jet-limit J^* exists, unboundedness of $U(J, I\backslash J)$ is immediately implied from unboundedness of $U(J^*, I\backslash J^*)$ in view of the definition of islands. Otherwise, there are infinitely many instants t such that $J(t) \neq J(t+1)$. At every such t, there exists $i \in I$ such that $i \in (J(t)\backslash J(t+1)) \cup (J(t+1)\backslash J(t))$. Therefore, recalling (4.9), $U_t(J, I\backslash J) \geq a_{ii}(t) \geq \delta$. Since there are infinitely many such times, $U(J, I\backslash J)$ is unbounded. □

4.6.3 Balanced Asymmetric Chains

Definition 4.13. *Bolouki and Malhamé (2012a) A chain $\{A(t)\}$ of stochastic matrices is said be balanced asymmetric with bound M, if for every subsets $\mathcal{V}_1, \mathcal{V}_2 \subseteq \mathcal{V}$ of the same cardinality, and for every $t \geq 0$:*

$$\sum_{i \notin \mathcal{V}_1} \sum_{j \in \mathcal{V}_2} a_{ij}(t) \leq M \sum_{i \in \mathcal{V}_1} \sum_{j \notin \mathcal{V}_2} a_{ij}(t). \tag{4.52}$$

Proposition 4.3. *Every balanced asymmetric chain is in Class $\mathcal{P}^*$.*

To prove Proposition 4.3, we need the following lemma.

Lemma 4.3. *Let A be an $(N \times N)$ balanced asymmetric matrix with bound M. Then, there exists a permutation matrix $P_{N\times N}$ such that the product PA is self-confident with bound $\delta = 4/(MN^2 + 4N - 4)$.*

Proof. Form a bipartite-graph $\mathcal{H}(\mathcal{V}, \mathcal{E})$ from A with N nodes in each part. Let $\mathcal{V}_1$ and $\mathcal{V}_2$, each a copy of $\mathcal{V}$, be sets of nodes of the two parts of $\mathcal{H}$. For every $i \in \mathcal{V}_1$ and $j \in \mathcal{V}_2$, connect i to j if $a_{ij} \geq \delta = 4/(MN^2 + 4N - 4)$. We wish to show that $\mathcal{H}$ has a perfect matching. By Hall's Marriage Theorem (Bondy and Murty, 1976, Theorem 5.2), it suffices to show that for every subset $\mathcal{K} \subseteq \mathcal{V}_1$, we have $|D(\mathcal{K})| \geq |\mathcal{K}|$ where

$$D(\mathcal{K}) = \{j \in \mathcal{V}_2 | \exists i \in \mathcal{K} \text{ s.t. } (i,j) \in \mathcal{E}\}. \tag{4.53}$$

Indeed, assume that on the contrary, there exists $\mathcal{K} \subseteq \mathcal{V}_1$ such that $k' = |D(\mathcal{K})| < |\mathcal{K}| = k$. Let $\mathcal{K} = \{c_1, \ldots, c_k\}$ and $D(\mathcal{K}) = \{d_1, \ldots, d_{k'}\}$. Define $\mathcal{K}' \subsetneq \mathcal{K}$ by $\mathcal{K}' = \{c_1, \ldots, c_{k'}\}$. We now have:

$$\sum_{i \in \mathcal{K}'} \sum_{j \notin D(\mathcal{K})} a_{ij} < k'(N - k')\delta \leq \delta N^2/4. \tag{4.54}$$

On the other hand,

$$\begin{aligned}\textstyle\sum_{i\notin\mathcal{K}'}\sum_{j\in D(\mathcal{K})} a_{ij} &\geq \textstyle\sum_{i\in\mathcal{K}\setminus\mathcal{K}'}\sum_{j\in D(\mathcal{K})} a_{ij} \\ &= (k-k') - \textstyle\sum_{i\in\mathcal{K}\setminus\mathcal{K}'}\sum_{j\notin D(\mathcal{K})} a_{ij} \\ &\geq (k-k') - (k-k')(N-k')\delta \\ &\geq 1-(N-1)\delta.\end{aligned} \tag{4.55}$$

Since $\mathcal{K}', D(\mathcal{K}) \subsetneq \mathcal{V}$ are of identical cardinalities, the balanced asymmetry property of A together with (4.54) and (4.55) imply that

$$1-(N-1)\delta < \delta M N^2/4. \tag{4.56}$$

Thus, $\delta > 4/(MN^2+4N-4)$, which is a contradiction. Therefore, $\mathcal{H}$ has a perfect matching and consequently, there exists a permutation τ such that $a_{\tau(i),i} \geq \delta, \forall i$. Thus, the permutation matrix P with $e_{\tau(i)}$ as its ith row, where e_j denotes a row vector of length N with 1 in the jth position and 0 in every other position, is such that the product PA is self-confident with δ. □

Proof of Proposition 4.3: Let $\{A(t)\}$ be a balanced asymmetric chain with bound M. Set: $\delta = 4/(MN^2+4N-4)$. We recursively define sequence $\{P(t)\}$ of permutation matrices as follows: From Lemma 4.3, we know that there exists a permutation matrix $P(0)$ such that the product $P(0)A(0)$ is self-confident with δ. Find permutation matrix $P(t)$, $t \geq 1$, such that the product $P(t)A(t)P'(t-1)$ is self-confident with δ. Note that the existence of $P(t)$ is implied by Lemma 4.3, taking into account the fact that the product $A(t)P'(t-1)$ is balanced asymmetric with bound M, since the columns of the product are a permutation of the columns of $A(t)$, itself a balanced asymmetric matrix with bound M. Hence, if we define chain $\{B(t)\}$ by:

$$B(t) = P(t)A(t)P'(t-1), \tag{4.57}$$

then, $\{B(t)\}$ has both the self-confidence and balanced asymmetry properties. Since balanced asymmetry is stronger than cut-balance, chain $\{B(t)\}$ is both self-confidence and cut-balanced. Thus, from Touri and Nedić (2014), we conclude that chain $\{B(t)\}$ belongs to the set $\mathcal{P}^*$. Furthermore, it is straightforward to show that if $\{\pi(t)\}$ is an absolute probability sequence adapted to chain $\{B(t)\}$, then $\{\pi(t)P(t-1)\}$, where $P(-1) = I_{N\times N}$, is an absolute probability sequence adapted to chain $\{A(t)\}$. This immediately implies that $\{A(t)\} \in \mathcal{P}^*$. □

The class property $\mathcal{P}^*$ implies that absolute probabilities are uniformly bounded away from zero, and as a result, that there is no J^0 in the jet decomposition of the D-S Theorem.

Therefore, we again consider only $J^1, \ldots, J^c$ as the jet decomposition.

Proposition 4.4. *If $\{A(t)\}$ is balanced asymmetric, then the cardinality of each jet in the jet decomposition of the D-S Theorem, becomes time-invariant after a finite time.*

Proof. Let $\{A(t)\}$ be balanced asymmetric with bound M. It suffices to show that there are finite times in which cardinality of a jet, in the jet decomposition of the D-S Theorem, increases by at least 1. In the following, we see what happens when the cardinality of a jet, say J^k, increases. Assume that for a fixed $t \geq 0$, we have $|J^k(t+1)| > |J^k(t)|$. For an arbitrary $i \in J^k(t+1)$, let $\mathcal{T} \subsetneq J^k(t+1)$ be such that $i \notin \mathcal{T}$ and $|\mathcal{T}| = |J^k(t)|$. Thus by the balanced asymmetry property,

$$\begin{aligned} \textstyle\sum_{j\in J^k(t)} a_{ij}(t) &\leq \textstyle\sum_{l\notin\mathcal{T}} \sum_{j\in J^k(t)} a_{lj}(t) \\ &\leq M \textstyle\sum_{l\in\mathcal{T}} \sum_{j\notin J^k(t)} a_{lj}(t) \\ &\leq M \textstyle\sum_{l\in J^k(t+1)} \sum_{j\notin J^k(t)} a_{lj}(t). \end{aligned} \tag{4.58}$$

Therefore,

$$\begin{aligned} &\textstyle\sum_{i\in J^k(t+1)} \sum_{j\in J^k(t)} a_{ij}(t) \\ &\quad \leq |J^k(t+1)|.M \textstyle\sum_{i\in J^k(t+1)} \sum_{j\notin J^k(t)} a_{ij}(t). \end{aligned} \tag{4.59}$$

On the other hand,

$$\begin{aligned} &\textstyle\sum_{i\in J^k(t+1)} \sum_{j\in J^k(t)} a_{ij}(t) \\ &\quad = |J^k(t+1)| - \textstyle\sum_{i\in J^k(t+1)} \sum_{j\notin J^k(t)} a_{ij}(t). \end{aligned} \tag{4.60}$$

Eqs. (4.59) and (4.60) together imply:

$$\sum_{i\in J^k(t+1)} \sum_{j\notin J^k(t)} a_{ij}(t) \geq \frac{|J^k(t+1)|}{1+M|J^k(t+1)|} \geq \frac{1}{1+M}. \tag{4.61}$$

Once again, since the cumulative interaction between J^k and $\bar{J}^k$ must be finite over the infinite time interval because of the D-S Theorem and in view of the class property $\mathcal{P}^*$, (4.61) can occur only for finitely many times t, and this completes the proof. $\square$

An immediate corollary of Proposition 4.4 is as follows.

Corollary 4.4. *Bolouki and Malhamé (2012a) Consider a multi-agent system with dynamics (4.1), where $\{A(t)\}$ is balanced asymmetric. Then, $z_i(t)$ converges for every $i \in \mathcal{V}$, as t goes to infinity, where $z_i(t)$ is the ith least value among $x_1(t), \ldots, x_N(t)$.*

Definition 4.14. *Touri and Nedić (2012b) A chain $\{A(t)\}$ of stochastic matrices is said to have the absolute infinite flow property, if for every jet J in $\mathcal{V}$ with a time-invariant size, $U(J, \mathcal{V}\backslash J)$ is unbounded.*

Proposition 4.5. *Bolouki and Malhamé (2012a) If $\{A(t)\}$ is balanced asymmetric, then, $\{A(t)\}$ is class-ergodic if and only if the absolute infinity property holds over each island of $\{A(t)\}$. Furthermore, in case of class-ergodicity, islands are the ergodic classes of $\{A(t)\}$.*

Proof. From Proposition 4.3, we know that $\{A(t)\} \in \mathcal{P}^*$. Therefore, taking advantage of Theorem 4.4, it suffices to show that absolute infinite flow and infinite jet-flow properties are equivalent on each island. Obviously, the former is implied by the latter. We prove the converse as follows: Let the absolute infinite flow property hold over each island. Assume that I is an arbitrary island of $\{A(t)\}$ and J is an arbitrary jet in I. If the cardinality of jet J becomes time-invariant after a finite time, unboundedness of $U(J, I\backslash J)$ is immediately implied from the absolute infinite flow property over I. Otherwise, the cardinality of J increases infinitely many times by at least 1. In this case, from the proof of Proposition 4.4 (see (4.61)), we know that $V(J, \mathcal{V}\backslash J)$ is unbounded, and consequently $U(J, \mathcal{V}\backslash J)$ is unbounded following Lemma 4.1. Moreover,

$$U(J, \mathcal{V}\backslash J) + U(I\backslash J, \mathcal{V}\backslash I) = U(J, I\backslash J) + U(I, \mathcal{V}\backslash I), \tag{4.62}$$

and since $U(I, \mathcal{V}\backslash I)$ is bounded because I is an island, unboundedness of $U(J, \mathcal{V}\backslash J)$ implies that $U(J, I\backslash J) = \infty$. This completes the proof. □

Corollary 4.5. *Bolouki and Malhamé (2012a) If chain $\{A(t)\}$ is balanced asymmetric, then it is ergodic if and only if it has the absolute infinite flow property.*

4.7 A Geometric Approach towards Consensus Algorithms

In this section, we introduce a geometric framework for a general linear consensus algorithm, that not only interprets the notions of jets and the ocean as explained in the previous sections, but serves an alternative proof of our results stated in the previous sections, and furthermore, as will be shown in the next section, extends them naturally to the continuous time case.

Let $\Phi(t, \tau)$, $t, \tau \geq 0$, be the state transition matrix of discrete time model (4.1), i.e.,

$$\Phi(t, \tau) = A(t-1)A(t-2)\cdots A(\tau). \tag{4.63}$$

Therefore,

$$x(t) = \Phi(t, \tau)x(\tau), \forall t, \tau \geq 0. \tag{4.64}$$

For every $t \geq \tau \geq 0$, assume that $C_{t,\tau}$ is the convex hull of the columns of $\Phi'(t, \tau)$. Note that each column of $\Phi'(t, \tau)$ is a stochastic vector representing a point in $\mathbb{R}^N$, and $C_{t,\tau}$ is a polytope in $\mathbb{R}^N$ if we consider points and segments in $\mathbb{R}^N$ as polytopes with one and two vertices respectively.

Lemma 4.4. *For every $t_2 \geq t_1 \geq \tau$, we have: $C_{t_2,\tau} \subset C_{t_1,\tau}$, i.e., polytope $C_{t_1,\tau}$ contains polytope $C_{t_2,\tau}$.*

Proof. From (4.63), we have:

$$\Phi'(t_2, \tau) = \Phi'(t_1, \tau)\Phi'(t_2, t_1). \tag{4.65}$$

Since $\Phi(t_2, t_1)$ is a stochastic matrix, each column of $\Phi'(t_2, \tau)$ is a convex combination of columns of $\Phi'(t_1, \tau)$. Therefore, every column of $\Phi'(t_2, \tau)$ lies in $C_{t_1,\tau}$, and the lemma is proved. □

Lemma 4.4 shows that for a fixed $\tau \geq 0$, $C_{t,\tau}$ shrinks as t grows. A projection of nested polytopes $C_{t,\tau}$'s on a two-dimensional space is shown in Fig 4.3.

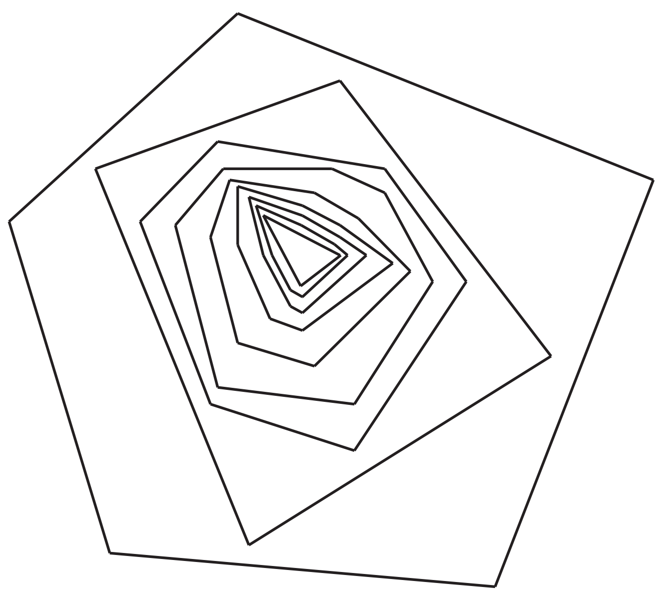

Figure 4.3 An example of nested polygons converging to a triangle.

It is to be noted that when underlying chain $\{A(t)\}$ of dynamics (4.1) is ergodic, the nested polygons will converge to a point in $\mathbb{R}^N$. In general, one concludes that for every $\tau \geq 0$, $\lim_{t\to\infty} C_{t,\tau}$ exists and is also a polytope in $\mathbb{R}^N$. Let C_τ be the limiting polytope with c_τ vertices. It is clear that $1 \leq c_\tau \leq N$. One can show that c_τ is a non-decreasing function of τ (see Section 5.9.2) and becomes constant after some finite time. We assume, without loss of generality, that c_τ is equal to constant c, $\forall \tau \geq 0$. It is worth mentioning that the choice of letter c here, that also represents the number of jets in the jet decomposition of the Sonin's D-S Theorem in this paper, for the number of vertices of limiting polytope C_0, is not accidental, as it will be shown, in the current section, that the two numbers are equal.

Let $v_1, \ldots, v_c$ be the c vertices of C_0. Assume that $\{0_t\}$ is a sequence of agents, i.e., $0_t \in \mathcal{V}$ for every $t \geq t_0$.

Theorem 4.5. *If sequence $\{0_t\}_{t\geq 0}$, $0_t \in \mathcal{V}$, $\forall t \geq 0$, is such that the distance between $\Phi'_{0_t}(t, 0)$ and set $\{v_1, \ldots, v_c\}$ does not converge to zero as t grows large, then:*

$$\inf\{\pi_{0_t}(t) \,|\, t \geq 0\} = 0. \tag{4.66}$$

Proof. We know that vector v_i, $1 \leq i \leq c$, lies outside of the convex hull of vectors v_j's, $j \neq i$. Let w_i be the nearest point to v_i, on the convex hull of v_j's, $j \neq i$. For a small $\epsilon' > 0$, draw a hyperplane, distant ϵ' from v_i, crossing segment $v_i w_i$ and orthogonal to it. Let $u(t) \triangleq \Phi'_{0_t}(t, 0)$. For a sufficiently small ϵ', there exists a subsequence of $\{u(t)\}$ such that v_i and the elements of the subsequence lie on opposite sides of the hyperplane for every i, $1 \leq i \leq c$. Otherwise, the distance between $\{u(t)\}$ and set $\{v_1, \ldots, v_c\}$ would converge to zero. Define:

$$\epsilon'_1 \triangleq \min\{|v_i - w_i| \,|\, 1 \leq i \leq c\}, \tag{4.67}$$

and:

$$\epsilon \triangleq \min\{\epsilon', \epsilon'_1/4\}, \tag{4.68}$$

and for an arbitrary constant δ, $0 < \delta < 1$, let:

$$\epsilon_1 \triangleq \delta\epsilon/(2N) \tag{4.69}$$

We summarize the rest of the proof, since it is very similar to the proof of Lemma 5.7, from (5.25) to (5.35). We know that for a sufficiently large time $T \geq 0$, if $t \geq T$, every vector in $C_{t,0}$ lies within an ϵ_1-distance of C_0. For every i, $1 \leq i \leq c$, draw a hyperplane l_i, parallel to the hyperplane drawn previously, distant ϵ from v_i, crossing segment $v_i w_i$. Draw also a hyperplane m_i, parallel to l_i, on the other side of v_i, distant ϵ_1 from v_i (see Fig. 4.4).

Define for every i, $1 \leq i \leq c$:

$$S^i = \{j \in \mathcal{V} \,|\, \Phi'_j(T, 0) \text{ lies in strip margined by } l_i, m_i\}. \tag{4.70}$$

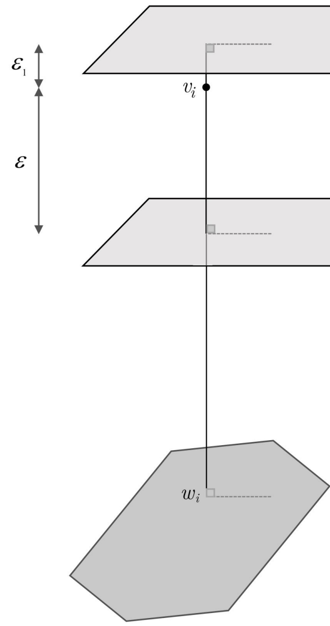

Figure 4.4 Planes l_i and m_i are orthogonal to segment $v_i w_i$.

One can show that, S^i's, $1 \leq i \leq c$, are disjoint non-empty sets. Define also:

$$S^0 = \mathcal{V} \setminus \bigcup_{j=1}^{c} S^j. \tag{4.71}$$

As mentioned above, there exists a subsequence of $\{u(t)\}$ such that v_i and the elements of the subsequence lie on the opposite sides of l_i (note that $\epsilon \leq \epsilon'$) for every $i = 1, \ldots, c$. Without loss of generality, assume that $\{u(T)\}$ belongs to that subsequence (otherwise, choose $T_1 > T$ such that $u(T_1)$ belongs to that subsequence, and replace T by T_1 in the argument). Hence, $S^0 \neq \emptyset$, and S^i 's partition agent set $\mathcal{V}$. Similar to the proof of Lemma 5.7, we have the following inequality (equivalence of (5.35)):

$$\sum_{j \notin S^i} (u_i)_j \leq 2\delta/(2N+1) < \delta/N. \tag{4.72}$$

Consequently,

$$\sum_{j \in S^0} (u_i)_j \leq \sum_{j \notin S^i} (u_i)_j < \delta/N. \tag{4.73}$$

Thus, for every $i \in \mathcal{V}$ and $j \in S^0$:

$$\inf\{\Phi_{i,j}(t,T) \,|\, t \geq T\} < \delta/N. \tag{4.74}$$

Consequently,

$$\inf\{\sum_{i\in\mathcal{V},j\in S^0}\Phi_{i,j}(t,T)\,|\,t\geq T\}<N\delta/N=\delta. \tag{4.75}$$

Since we have:

$$\pi_j(T)=\pi(t)\Phi^j(t,T)=\sum_{i\in\mathcal{V}}\pi_i(t)\Phi_{i,j}(t,T)\leq\sum_{i\in\mathcal{V}}\Phi_{i,j}(t,T), \tag{4.76}$$

from (4.75) we conclude that:

$$\pi_j(T)<\delta,\,\forall j\in S^0. \tag{4.77}$$

We recall that S^0 includes one of the elements of sequence $\{0_t\}$, i.e., 0_T. Hence, $\pi_{0_T}(T)<\delta$. Recall also that δ was chosen arbitrarily. By letting δ go to zero, we conclude that $\inf\{\pi_{0_t}(t)\,|\,t\geq 0\}=0$, and the theorem is proved. $\square$

Remark 4.2. *We explain, in the following, that there is a one-to-one correspondence between the vertices of limiting polytope C_0 and jets $J^1,\ldots,J^c$ of the Sonin's jet decomposition.*

Recall, from Section 4.4, that how we employed the absolute probability sequence of chain $\{A(t)\}$ to construct a forward propagating Markov chain from the given backward one. Now, let J^k be an arbitrary jet among $J^1,\ldots,J^c$. Let, also, $\{k_t\}$ be a sequence inside jet J^k, i.e., $k_t\in J^k(t)$, $\forall t\geq 0$. Since, due to the D-S Theorem, $\lim_{t\to\infty}x_{k_t}(t)$ exists irrespective of what $x(0)$ is, $\lim_{t\to\infty}\Phi'_{k_t}(t,0)$ exists as well, and is irrespective of how the sequence is chosen. We aim to show that $\lim_{t\to\infty}\Phi'_{k_t}(t,0)$ is one of $v_1,\ldots,v_c$. Since the volume of $J^k(t)$ converges to a non-zero constant, as $t\to\infty$, one can form a sequence $\{k_t\}$ inside jet J^k, i.e., $k_t\in J^k(t)$, $\forall t\geq 0$, such that:

$$\lim\inf\{\pi_{k_t}(t)\,|\,t\geq 0\}>0. \tag{4.78}$$

One way to form such a sequence is to pick, at each time instant, the cup in J^k that has the maximum volume. From Theorem 4.5 and inequality (4.78), we conclude that the distance between $\Phi'_{k_t}(t,0)$ and set $\{v_1,\ldots,v_c\}$, the vertices of limiting polytope C_0, must vanish as t grows large. Thus, $\lim_{t\to\infty}\Phi'_{k_t}(t,0)$ is belongs to set $\{v_1,\ldots,v_c\}$.

It is also clear that if sequences $\{s_t\}$ and $\{k_t\}$ are in two disjoint jets J^s and J^k respectively, $\lim_{t\to\infty}\Phi'_{s_t}(t,0)$ and $\lim_{t\to\infty}\Phi'_{k_t}(t,0)$ cannot converge to the same vertex of C_0, since otherwise, merging the two jets would violate the uniqueness of the Sonin's jet decomposition.

4.8 Consensus in the Continuous Time Case

One may define a general linear consensus algorithm in continuous time as follows:

$$\dot{x}=A(t)x(t),\,t\geq 0, \tag{4.79}$$

where $x(t)$ is the vector of opinions at each time instant $t \geq 0$ and $\{A(t)\}$ is the underlying chain of the system. It is assumed that each matrix of underlying chain $A(t)$ has zero row sum and non-negative off-diagonal elements and each element $a_{ij}(t)$ of $A(t)$ is a measurable function. These constraints on the underlying chain suggest a view of A(t) as the evolution of the intensity matrix of a time inhomogeneous Markov chain. We shall use in this section, a continuous time version of the geometric framework developed in Section 4.7, in convergence analysis of agents in a network with continuous time dynamics (4.79), particularly when underlying chain $\{A(t)\}$ is in a continuous time version of Class $\mathcal{P}^*$.

Let $\Phi(t, \tau)$, $t, \tau \geq 0$, represent the state transition matrix of system associated with (4.79), i.e.,

$$x(t) = \Phi(t, \tau)x(\tau), \ \forall t \geq \tau \geq 0. \tag{4.80}$$

Note that similar to the discrete time case, $\Phi(t, \tau)$ is a stochastic matrix for every $t \geq \tau \geq 0$. More specifically, $\Phi_{i,j}(t, \tau)$ can be considered as transition probability of a backward propagating inhomogeneous Markov chain. In particular, for every $t_2 \geq t_1 \geq \tau \geq 0$, we have:

$$\Phi_{i,j}(t_2, \tau) = \sum_k \Phi_{i,k}(t_2, t_1)\Phi_{k,j}(t_1, \tau), \tag{4.81}$$

with the conditions:

$$\Phi_{i,j}(t, \tau) \geq 0, \tag{4.82}$$

$$\sum_j \Phi_{i,j}(t, \tau) = 1, \tag{4.83}$$

$$\Phi_{i,j}(t, t) = \delta_{ij}, \tag{4.84}$$

where δ_{ij} is the Kronecker symbol.

Underlying chain $\{A(t)\}$ is said to be ergodic if for every $\tau \geq 0$, $\Phi(t, \tau)$ converges to a matrix with equal rows as $t \to \infty$. Similar to the discrete time case, ergodicity of $\{A(t)\}$ is equivalent to the occurrence of unconditional consensus in (4.79). Moreover, $\{A(t)\}$ is class-ergodic if for every $\tau \geq 0$, $\lim_{t\to\infty} \Phi(t, \tau)$ exists, but with possibly distinct rows. Chain $\{A(t)\}$ is class-ergodic if and only if multiple consensus occurs in (4.79) unconditionally. Recall that the associated state transition matrix associated with (4.79) can be expressed via the Peano-Baker series (see (Brockett, 1970, Section 1.3)):

$$\begin{aligned}\Phi(t, \tau) = I_{N\times N} &+ \textstyle\int_\tau^t A(\sigma_1)d\sigma_1 \\ &+ \textstyle\int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)d\sigma_2 d\sigma_1 \\ &+ \textstyle\int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)\int_\tau^{\sigma_2} A(\sigma_3)d\sigma_3 d\sigma_2 d\sigma_1 \\ &+ \cdots,\end{aligned} \tag{4.85}$$

where $I_{N\times N}$ denotes the $N \times N$ identity matrix. Remember that state transition matrix $\Phi(t,\tau)$ is invertible for every $t \geq \tau \geq 0$.

Furthermore, once again, based on Kolmogoroff (1936), we know that for every state transition matrix $\Phi(t,\tau)$, $t,\tau \geq 0$, there exists an absolute probability sequence $\{\pi(t)\}$, $t \geq 0$, such that:

$$\pi(\tau) = \pi(t)\Phi(t,\tau),\ \forall t,\tau \geq 0. \tag{4.86}$$

Having recalled the state transition matrix and the absolute probability sequence for the continuous time model (4.79), we can now carry out a continuous time version of the geometric framework developed in Section 4.7. Once again, for every $t \geq \tau \geq 0$, assume that $C_{t,\tau}$ is the convex hull of columns of $\Phi'(t,\tau)$, or equivalently transposed rows of $\Phi(t,\tau)$. Remember that each column of $\Phi'(t,\tau)$ is a stochastic vector as in the discrete time case. Now, note that Lemma 4.4 holds for the continuous time as well, since its proof remains valid assuming that the time indices refer to continuous time. Therefore, we again assume that limiting polytopes C_τ's, $\tau \geq 0$, exist. Let c_τ be the number of vertices of C_τ. We show in the following that, c_τ, $\tau \geq 0$, is constant (unlike the discrete time case in which c_τ was monotonic increasing with respect to τ). Assume that $\tau_2 \geq \tau_1 \geq 0$ are two arbitrary time instants. We wish to show that $c_{\tau_1} = c_{\tau_2}$. Define linear operator $\phi_{\tau_2,\tau_1} : \mathbb{R}^N \to \mathbb{R}^N$ by:

$$\phi_{\tau_2,\tau_1}(v) \triangleq \Phi'(\tau_2,\tau_1)v,\ \forall v \in \mathbb{R}^N. \tag{4.87}$$

Note now that from properties of state transition matrices, for $t \geq \tau_2 \geq \tau_1 \geq 0$, we have:

$$\Phi'(t,\tau_1) = \Phi'(\tau_2,\tau_1)\Phi'(t,\tau_2). \tag{4.88}$$

Therefore, in view of (4.88) by taking t to infinity, the vertices of C_{τ_2} are uniquely mapped to vectors in $\mathbb{R}^N$ which because of the linearity of map (4.87), will play the role of vertices for the generation of convex hull C_{τ_1}. Also, it is not difficult to show that the images of vertices of C_{τ_2} must remain vertices of C_{τ_1}, for if one of the images of a vertex of C_{τ_2}, say v, turned out to be a convex combination of other vertices of C_{τ_1}, this would also be true for the inverse images of these vertices (also vertices of C_{τ_2} due to invertibility of matrix $\Phi'(\tau_2,\tau_1)$), and v would then fail to be a vertex of C_{τ_2}. In conclusion, C_{τ_1} and C_{τ_2} will have the same number of vertices, and (4.87) constitutes a one to one map between corresponding pairs of vertices. One may now use the same argument to extend Theorem 4.5 to the continuous time case while t_0, the initial time in Theorem 4.5, can be chosen arbitrarily here (recall that for Theorem 4.5 to be true, C_{t_0} must have had the maximum number of vertices among all C_τ's, and since c_τ is constant for $\tau \geq 0$ in the continuous time case, t_0 can be chosen arbitrarily).

We now aim to take advantage of Theorem 4.5 to address the limiting behavior of system (4.79) when underlying chain $\{A(t)\}$ is in Class $\mathcal{P}^*$.

Lemma 4.5. *For every $j \in \mathcal{V}$,*

$$\pi_j(\tau) \leq \inf \left\{ \sum_{i \in \mathcal{V}} \Phi_{i,j}(t, \tau) \,|\, t \geq \tau \right\}. \tag{4.89}$$

Proof. Obvious, since for every $t \geq \tau$:

$$\pi_j(\tau) = \pi(t)\Phi^j(t, \tau) = \sum_{i \in \mathcal{V}} \pi_i(t)\Phi_{i,j}(t, \tau) \leq \sum_{i \in \mathcal{V}} \Phi_{i,j}(t, \tau). \tag{4.90}$$

□

Adopting the same definition of Class $\mathcal{P}^*$ as in the discrete time case (see Section 4.5), we state the following lemma.

Lemma 4.6. *A state transition matrix $\Phi(t, \tau)$, $t, \tau \geq 0$, associated with (4.79), is in Class $\mathcal{P}^*$ if and only if for every $j \in \mathcal{V}$:*

$$\inf \left\{ \sum_{i \in \mathcal{V}} \Phi_{i,j}(t, \tau) \,|\, t \geq \tau \right\} > 0. \tag{4.91}$$

Proof. The *only if* part is an immediate result of Lemma 4.5, and the *if* part is a result of the way the existence of the absolute probability sequence can be obtained in Kolmogoroff (1936) by always choosing to initialize agent probabilities on finite intervals with a uniform distribution. □

Let the infinite flow graph of a continuous time chain $\{A(t)\}$ is defined according to Definition 4.6 by replacing summation with integral. The following theorem describes the convergence properties of system (4.79) when the underlying chain is in Class $\mathcal{P}^*$.

Theorem 4.6. *If state transition matrix $\Phi(t, \tau)$, $t, \tau \geq 0$, is in Class $\mathcal{P}^*$, then multiple consensus occurs unconditionally in system (4.79). Moreover, the number of consensus clusters is equal to the number of the components of the infinite flow graph of the transition chain. In particular, consensus occurs unconditionally if and only if the infinite flow property holds.*

The following theorem clarifies that Theorem 4.6 generalizes continuous time consensus results of Hendrickx and Tsitsiklis (2013).

Theorem 4.7. *If transition chain $\{A(t)\}$ in (4.79) is cut-balanced, then state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, is in Class $\mathcal{P}^*$.*

Proof. Let $\{A(t)\}$ be cut-balanced with bound K. Assume that $\Phi(t,\tau)$, $t \geq \tau \geq 0$, is the state transition matrix associated with (4.79). In view of Lemma 4.6, our aim is to show that: $1/Ne'\Phi_A(t,\tau) \geq p^*e'$, for some $p^* > 0$, where $e' = \begin{bmatrix}1 & \cdots & 1\end{bmatrix}$, and the inequality is to be understood element-wise.

Assume that $\alpha = \sup\{-a_{ii}(t') \,|\, i \in \mathcal{V}, \tau \leq t' \leq t\}$. Let chain B be such that $B(t') = A(t') + 2\alpha I$, $\forall \tau \leq t' \leq t$, where I is the identity matrix. It is easy to verify that:

$$\Phi_B(t,\tau) = e^{2\alpha(t-\tau)}\Phi_A(t,\tau). \tag{4.92}$$

Moreover, by construction, on-diagonal elements of $B(t')$, $\tau \leq t' \leq t$, are greater than or equal to α. Note that $B(t')$ ($\tau \leq t' \leq t$) is not a stochastic matrix; instead each of its rows sums up to 2α. We calculate in the following, $1/Ne'\Phi_B(t,\tau)$. Therefore, from the Peano-Baker series (4.85), the expression:

$$\frac{1}{N}e'\int_\tau^t B(\sigma_1)\int_\tau^{\sigma_1} B(\sigma_2)\cdots\int_\tau^{\sigma_{k-1}} B(\sigma_k)d\sigma_k\cdots d\sigma_1 \tag{4.93}$$

is of interest. Expression (4.93) is equal to:

$$\frac{(2\alpha)^k}{N}e'\int_\tau^t \frac{B(\sigma_1)}{2\alpha}\int_\tau^{\sigma_1} \frac{B(\sigma_2)}{2\alpha}\cdots\int_\tau^{\sigma_{k-1}} \frac{B(\sigma_k)}{2\alpha}d\sigma_k\cdots d\sigma_1, \tag{4.94}$$

which is also equal to:

$$(2\alpha)^k\int_\tau^t\int_\tau^{\sigma_1}\int_\tau^{\sigma_{k-1}} \frac{1}{N}e'\frac{B(\sigma_1)}{2\alpha}\frac{B(\sigma_2)}{2\alpha}\cdots\frac{B(\sigma_k)}{2\alpha}d\sigma_k\cdots d\sigma_1. \tag{4.95}$$

Note that $B(t')/2\alpha$ is a sequence of transition matrices which generates a Markov chain which is both cut-balanced and self-confident, and hence in Class $\mathcal{P}^*$ ((Touri and Nedić, 2014, Theorem 7)). As a result, there exists a positive p^* such that:

$$\frac{1}{N}e'\frac{B(\sigma_1)}{2\alpha}\cdot\frac{B(\sigma_2)}{2\alpha}\cdot\cdots\cdot\frac{B(\sigma_k)}{2\alpha} \geq p^*e'. \tag{4.96}$$

Inequality (4.96) implies that expression (4.95), and consequently expression (4.93), is greater than or equal to $(2\alpha)^k p^*(t-\tau)^k/k!$. Now, if we write $1/Ne'\Phi_B(t,\tau)$ as sum of expressions like (4.93), we have:

$$1/Ne'\Phi_B(t,\tau) \geq \sum_{k=0}^{\infty}\frac{(2\alpha)^k p^*(t-\tau)^k}{k!} = p^*e^{2\alpha(t-\tau)}. \tag{4.97}$$

Thus,

$$1/Ne'\Phi_A(t,\tau) \geq p^* e^{2\alpha(t-\tau)}.e^{-2\alpha(t-\tau)} = p^*, \tag{4.98}$$

and from Lemma 4.6 the theorem is proved. □

4.9 Conclusion

We considered a general linear distributed averaging algorithm in both discrete time and continuous time. Following Touri and Nedić (2012b), and recalling the notion of jets from Blackwell (1945), we introduced a property of chains of stochastic matrices, more precisely, the infinite jet-flow property in the discrete time case. The latter property is shown to be a strong necessary condition for ergodicity of the chain. Moreover, for the chain to be class-ergodic, the infinite jet-flow property must hold over each connected component of the infinite flow graph, as defined in Touri and Nedić (2014).

We then illustrated the close relationship between Sonin's D-S Theorem and convergence properties of linear consensus algorithms. By employing the D-S Theorem, we showed in the discrete time case that the necessary conditions found earlier are also sufficient in case the chain is in Class $\mathcal{P}^*$ Touri and Nedić (2014). We argued that the obtained equivalent conditions for ergodicity and class-ergodicity of chains in Class $\mathcal{P}^*$ can subsume the previous related results in the literature, Bolouki and Malhamé (2012a); Hendrickx and Tsitsiklis (2013); Touri and Nedić (2014) in particular.

A geometric approach was then introduced to interpret the jets in the D-S Theorem. The approach turned out to be a powerful method to rediscover our aforementioned results, and also to extend them to the continuous time case. In future work, we shall attempt an extension of our results to the case when the number of agents increases to infinity, although the D-S Theorem holds only if N is finite.

CHAPTER 5

ARTICLE 3: EMINENCE GRISE COALITIONS: ON THE SHAPING OF PUBLIC OPINION

Sadegh Bolouki, Roland P. Malhamé, Milad Siami, and Nader Motee

Submitted to *IEEE Transactions on Control of Network Systems*, September 2014

5.1 Abstract

We consider a network of evolving opinions. It includes multiple individuals with first-order opinion dynamics defined in continuous time and evolving based on a general exogenously defined time-varying underlying graph. In such a network, for an arbitrary fixed initial time, a subset of individuals forms an *éminence grise coalition*, abbreviated as EGC, if the individuals in that subset are capable of leading the entire network to agreeing on any desired opinion, through a cooperative choice of their own initial opinions. In this endeavor, the coalition members are assumed to have access to full profile of the underlying graph of the network as well as the initial opinions of all other individuals. While the complete coalition of individuals always qualifies as an EGC, we establish the existence of a minimum size EGC for an arbitrary time-varying network; also, we develop a non-trivial set of upper and lower bounds on that size. As a result, we show that, even when the underlying graph does not guarantee convergence to a global or multiple consensus, a generally restricted coalition of agents can steer public opinion towards a desired global consensus without affecting any of the predefined graph interactions, provided they can cooperatively adjust their own initial opinions. Geometric insights into the structure of EGC's are given. The results are also extended to the discrete time case where the relation with Decomposition-Separation Theorem is also made explicit.

5.2 Introduction

In this paper, we are mainly concerned with the occurrence of consensus in networks of individuals with opinions updated via a class of continuous time weighted distributed averaging algorithms characterized in general by an exogenous underlying chain of opinion update matrices, which behave like intensity matrices of a continuous time Markov chain. In such networks, *consensus* is said to occur if all opinions converge to the same value as time grows large. Furthermore, *Multiple consensus* is said to occur if each individual's opinion asymptotically converges to an individual limit. It is well known that such asymptotic behaviors relate directly to the properties of the Markov chain which underlies the opinion update dynamics. More specifically, the underlying chain of an opinion network may be such that consensus or multiple consensus occurs *unconditionally*, i.e., irrespective of the values of initial opinions of the individuals in the network. The unconditional occurrence of consensus is proved to be equivalent to *ergodicity* of the underlying chain Chatterjee and Seneta (1977). There is a similar correspondence between the unconditional occurrence of multiple consensus and *class-ergodicity* of the underlying chain Bolouki and Malhamé (2013); Touri and Nedić (2012b).

Ergodic and class-ergodic chains, i.e., chains leading to unconditional consensus or multiple consensus, have attracted an increasing attention in the literature in the past decade. Researchers of many different fields including robotics, social networks, economics, biology, etc., have been particularly interested in conditions under which a consensus algorithm guarantees consensus or multiple consensus to occur for an arbitrary choice of initial opinions. It is generally accepted that the earliest work on this class of opinion formation models was done in DeGroot (1974). The model was defined in discrete time, and the considered underlying chain was time-invariant. Later, more general cases were considered in Chatterjee and Seneta (1977), where the authors also made explicit the relationship between consensus and ergodicity of the underlying chain. Some of the earliest significant results on consensus date back to Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986). Interest in distributed consensus for agents moving in space was triggered by the numerical experiments in Vicsek *et al.* (1995) where a nonlinear algorithm was proposed for modeling evolution of multi-agent systems in discrete time. In this model, agents are assumed to have the same speed but different headings, and states are headings of agents. Using simulations, convergence to some kind of consensus (emerging behavior) was displayed in Vicsek *et al.* (1995). A linearized version of the model in Vicsek *et al.* (1995) was considered in Jadbabaie *et al.* (2003), where sufficient conditions for consensus based on analyzing infinite products of stochastic matrices, consistent with those of Bertsekas and Tsitsiklis (1989); Tsitsiklis (1984); Tsitsiklis *et al.* (1986) are established. Following Jadbabaie *et al.* (2003), many works have focused on identifying the largest class of underlying update chains for which consensus occurs unconditionally. Because of their close relationship to our current work, we mention in particular Blondel *et al.* (2005); Bolouki and Malhamé (2011a,b, 2012a,b); Hendrickx and Blondel (2006); Hendrickx and Tsitsiklis (2013); Li *et al.* (2004); Lorenz (2005); Moreau (2005); Touri and Nedić (2011, 2012a,b, 2014). In addition, Bolouki and Malhamé (2011a,b, 2012a,b, 2013); Hendrickx and Tsitsiklis (2013); Lorenz (2005); Touri and Nedić (2012a,b, 2014) also addressed the unconditional multiple consensus problem, or equivalently class-ergodicity of the underlying chain. For the continuous time case, Hendrickx and Tsitsiklis (2013) appears to provide the most general results thus far on consensus and multiple consensus. On the other hand, in our recent work Bolouki and Malhamé (2013), inspired by Touri and Nedić (2014) and Sonin *et al.* (2008), and to the best of our knowledge, we have identified for the discrete time case, the largest class to date of ergodic and class-ergodic chains.

In contrast to the above papers, which are concerned with "unconditional" consensus, the current paper aims at providing some answers to the following questions: What if the underlying chain is not ergodic? How can consensus still be achieved in a network with absolutely no assumption on the underlying chain? In other words, for a network with a general time-varying underlying opinion update chain, having fixed the initial time, what can be said about particular (non-trivial) choices of initial opinions leading to a possible consensus? Geometric insights on the nature of the "march" towards consensus allow one to realize that such choices of initial opinion vectors form a vector space the dimension of which is related to the characteristics of the underlying chain. The fact that such initial opinion vectors form a vector space suggests the existence of a possibly small subgroup of individuals in the network who are *naturally* capable of leading the whole group to eventually agree on any desired value *only* by collectively adjusting their own initial opinions. The word "naturally" here refers to the fact that the subgroup does not need to manipulate the nature of the network, and particularly *leaves all the interactions between any two individuals including themselves untouched*. They act like hidden leaders, or "éminences grises", not identifiable by title or position, yet who can, given time, thoroughly shape the ultimate public opinion. A subgroup with such leadership property is referred to as an *Éminence Grise Coalition*, or simply EGC, in this work. The EGC's that a network admit are determined by the properties of the underlying chain of the network only. While it is trivial to establish the existence of at least one largest EGC, namely the universal coalition of individuals, one of our main points of interest in this work is to characterize the size and identity of the smallest coalition that can achieve public opinion shaping. Tight bounds on the size of that coalition are also of interest. The reasons why such individuals may want to act as a coalition can be multiple. Two such possibilities are: (i) They have been identified as key decision makers by a knowledgeable negotiator, have collectively agreed on a bargaining position, yet need to steer their peers towards the collective agreement, (ii) A shady opinion manipulator has identified them as key decision makers and has succeeded in "buying out" their collaboration.

The rest of the paper is organized in such a way that no confusion arises between the continuous time and the discrete time cases. We explicitly deal with the continuous time case in the largest part of the paper, that is Sections 5.3–5.8, and discuss the discrete time case in Section 5.9. More specifically, we explicitly state the problem setup in Section 5.3, where we introduce the notion of *rank* of a chain of matrices which is shown to be equal to the size of the smallest EGC of the network. In Section 5.4, a geometric framework is developed to interpret the notion of rank of a chain and also obtain an upper bound for the rank, or equivalently the size of the smallest EGC of a consensus algorithm. This geometric framework proves to be useful in dealing with both the continuous time and the discrete time cases. We establish in Section 5.5, lower bounds on the rank based on the existing notions in the literature, namely the so-called infinite flow graph and unbounded interactions graph of a chain. The rank of time-invariant chains is discussed in Section 5.6. We address a large class of time-varying chains, the so-called Class $\mathcal{P}^*$, and their rank in particular, in Section 5.7. It is shown that chains of the the two classes discussed in Sections 5.6 and 5.7, are examples of chains for which the bounds on rank obtained earlier in Sections 5.4 and 5.5 are actually attained. Full-rank chains, namely chains with rank equal to the size of the network are characterized in Section 5.8. In the process of characterizing full-rank chains, we also discover another upper bound on rank. In Section 5.9, we extend our analysis of the continuous time case to the discrete time case. As will be shown, the size of the smallest EGC is equal to the number of jets in the jet decomposition of the Sonin Decomposition Separation Theorem (see Bolouki and Malhamé (2013); Sonin *et al.* (2008)). Concluding remarks and suggestions of future work end the paper in Section 7.

5.3 Notions and Terminology

The notions, preliminaries, and notation described in this section are for the purposes of the continuous time part of this paper, i.e., Sections 5.3–5.8, although some may be consistent with the contents of Section 5.9, the discrete time analysis. Let N be the number of individuals and $\mathcal{V} = \{1, \ldots, N\}$ be the set of individuals. Assume that t stands for the continuous time index. Let a time-varying chain $\{A(t)\}_{t\geq 0}$ of square matrices of size N be such that each matrix $A(t)$, $t \geq 0$, has zero row sum and non-negative off-diagonal entries and each entry $a_{ij}(t)$ of $A(t)$, $i, j \in \mathcal{V}$, is a measurable function. Continuous time chains of matrices, that we deal with in this paper, are assumed to have these properties. According to these constraints, $A(t)$ can be viewed as the evolution of the intensity matrix of a time inhomogeneous Markov chain. Let dynamics of an opinion network be described by the following continuous time distributed averaging algorithm:

$$\dot{x}(t) = A(t)x(t),\ t \geq t_0, \tag{5.1}$$

where $t_0 \geq 0$ is the initial time and $x(t) \in \mathbb{R}^N$ is the vector of opinions at each time instant $t \geq t_0$. Thus, $x_i(t)$ is the scalar opinion of individual i at time $t \geq t_0$. Chain $\{A(t)\}_{t\geq 0}$, or simply $\{A(t)\}$, is referred to as the *underlying chain* of the network with dynamics (5.1).

Assume that $\Phi(t,\tau)$, $t \geq \tau \geq 0$ denotes the state transition matrix associated with chain $\{A(t)\}$. Therefore, for the network with dynamics (5.1), we must have:

$$x(t) = \Phi(t,\tau)x(\tau),\ \forall t \geq \tau \geq t_0. \tag{5.2}$$

From (Brockett, 1970, Section 1.3), the Peano-Baker series of state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, associated with chain $\{A(t)\}$ is expressed as:

$$\begin{aligned}\Phi(t,\tau) = I_{N\times N} &+ \textstyle\int_\tau^t A(\sigma_1)d\sigma_1 \\ &+ \textstyle\int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)d\sigma_2 d\sigma_1 \\ &+ \textstyle\int_\tau^t A(\sigma_1)\int_\tau^{\sigma_1} A(\sigma_2)\int_\tau^{\sigma_2} A(\sigma_3)d\sigma_3 d\sigma_2 d\sigma_1 \\ &+ \cdots,\end{aligned} \tag{5.3}$$

where $I_{N\times N}$ denotes the identity matrix of size N. Remember that state transition matrix $\Phi(t,\tau)$ is invertible for every $t \geq \tau \geq 0$.

We use the following notation throughout this paper: $\Phi_i(t,\tau)$ and $\Phi_{i,j}(t,\tau)$, $1 \leq i,j \leq N$, denote the ith column and the (i,j)th element of $\Phi(t,\tau)$ respectively. Moreover, the transposition of a matrix is indicated by the matrix followed by prime ($'$). We emphasize that $\Phi'_i(t,\tau)$ refers to the ith column of $\Phi'(t,\tau)$ (prime acts first). For an arbitrary vector $v \in \mathbb{R}^N$, and $1 \leq i \leq N$, v_i denotes the ith element of v. Vectors of all zeros and all ones in $\mathbb{R}^N$ are indicated by $\mathbf{0}_N$ and $\mathbf{1}_N$ respectively. For an arbitrary subset $\mathcal{S} \subset \mathcal{V}$, $\mathcal{V}\backslash\mathcal{S}$ denotes the complement of $\mathcal{S}$ in $\mathcal{V}$.

Remark 5.1. *Notice that $\Phi_{i,j}(t,\tau)$, $t \geq \tau \geq 0$, for a fixed τ, can be viewed as a transition probability in a backward propagating inhomogeneous Markov chain. In particular, for every $t_2 \geq t_1 \geq \tau \geq 0$, we have:*

$$\Phi_{i,j}(t_2,\tau) = \sum_k \Phi_{i,k}(t_2,t_1)\Phi_{k,j}(t_1,\tau), \tag{5.4}$$

with the conditions:

$$\Phi_{i,j}(t,\tau) \geq 0, \tag{5.5}$$

$$\sum_j \Phi_{i,j}(t,\tau) = 1, \tag{5.6}$$

$$\Phi_{i,j}(\tau,\tau) = \delta_{ij}, \tag{5.7}$$

where δ_{ij} is the Kronecker symbol.

5.3.1 Éminence Grise Coalition

Definition 5.1. *For an opinion network with dynamics (5.1), a subgroup of individuals $\mathcal{S} \subset \mathcal{V}$ is said to be an Éminence Grise Coalition if for any arbitrary $x^* \in \mathbb{R}$ and any initialization of opinions of individuals in $\mathcal{V} \backslash \mathcal{S}$, there exists an initialization of opinions of individuals in $\mathcal{S}$ such that $\lim_{t\to\infty} x(t) = x^*.\mathbf{1}_N$, i.e., all individuals asymptotically agree on x^*. The term Éminence Grise Coalition may also be referred to as acronym EGC.*

From another point of view that also justifies the selection of the term Éminence Grise Coalition, an EGC of a network with dynamics (5.1) is a subgroup of individuals who are capable of leading the whole group towards a global agreement on any desired ultimate opinion only by properly initializing their own opinions, with the assumption that they are aware of the underlying chain of the network and initial opinions of the rest of individuals.

Lemma 5.1. *In an opinion network with dynamics (5.1), a subset $\mathcal{S} \subset \mathcal{V}$ is an EGC if and only if for any initialization of opinions of individuals in $\mathcal{V} \backslash \mathcal{S}$, there exists an initialization of opinions of individuals in $\mathcal{S}$ such that $\lim_{t\to\infty} x(t) = \mathbf{0}_N$.*

Proof. The "only if" part is obvious by setting $x^* = 0$ in Definition 5.1. Conversely, assume that $\mathcal{S} \subset \mathcal{V}$ has the property that for any initialization of individuals in $\mathcal{V} \backslash \mathcal{S}$, there exists an initialization of individuals in $\mathcal{S}$ such that all opinions asymptotically converge to zero. To show that $\mathcal{S}$ is an EGC according to Definition 5.1, let arbitrary $x^* \in \mathbb{R}$ be the desired value of agreement and assume that for every $i \in \mathcal{V} \backslash \mathcal{S}$, the opinion of individual i is initialized at $\hat{x}_i \in \mathbb{R}$, where $\hat{x}_i$ is arbitrary. We seek an initialization of opinions of individuals in $\mathcal{S}$ leading to an asymptotic agreement of all individuals on x^*. For a moment, let us assume that for every $i \in \mathcal{V} \backslash \mathcal{S}$, the opinion of individual i was initialized at $\hat{x}_i - x^*$. For such an initialization, by the assumption on $\mathcal{S}$, there would be an initialization of opinions of individuals in $\mathcal{S}$, say at $\hat{x}_i$ for each individual $i \in \mathcal{S}$, such that all opinions would asymptotically converge to zero. In other words, if the individual opinions in the network with dynamics (5.1) were initialized as:

$$x_i(0) = \begin{cases} \hat{x}_i - x^* & \text{if } i \in \mathcal{V} \backslash \mathcal{S} \\ \hat{x}_i & \text{if } i \in \mathcal{S} \end{cases} \tag{5.8}$$

then, $\lim_{t\to\infty} x(t) = \mathbf{0}_N$. Now, the following initialization, which is basically a translation of the previous initialization by x^*, will lead to an agreement on x^*:

$$x_i(0) = \begin{cases} \hat{x}_i & \text{if } i \in \mathcal{V} \backslash \mathcal{S} \\ \hat{x}_i + x^* & \text{if } i \in \mathcal{S} \end{cases} \tag{5.9}$$

Agreement on x^* is easily proved from the previous agreement on zero and noticing that translations are preserved in consensus dynamics (5.1) since $\Phi(t, t_0)$, for every $t \geq t_0$, has an eigenvector $\mathbf{1}_N$ corresponding to eigenvalue 1. Thus, for an arbitrary initialization of individuals in $\mathcal{V} \backslash \mathcal{S}$, we found an initialization of individuals in $\mathcal{S}$ such that all opinions asymptotically converge to the desired value x^*, which completes the proof. $\square$

Our primary objective in this work is characterizing the smallest EGC in an opinion network with dynamics described by (5.1). In particular, the size of the smallest EGC is of interest.

5.3.2 Rank of a Chain

We now define several operators for chains of matrices. **Bold** style is used for chain operators in this paper to distinguish them from matrix operators that are in roman style. Let $\{A(t)\}$ be a chain of matrices and $\Phi(t, \tau)$, $t \geq \tau \geq 0$ be its associated state transition matrix.

Definition 5.2. *The null space of chain $\{A(t)\}$ at time $\tau \geq 0$, denoted by $\mathbf{null}_\tau(A)$, is defined by:*

$$\mathbf{null}_\tau(A) \triangleq \Big\{ v \in \mathbb{R}^N | \lim_{t \to \infty} \Big(\Phi(t, \tau) v \Big) = \boldsymbol{0}_N \Big\}. \tag{5.10}$$

It is straightforward to show that $\mathbf{null}_\tau(A)$ is a vector space for every $\tau \geq 0$.

Lemma 5.2. *The dimension of vector space $\mathbf{null}_\tau(A)$, $\tau \geq 0$, is independent of τ.*

Proof. Let $\tau_2 > \tau_1 \geq 0$ be two arbitrary time instants. Define linear operator $\phi_{\tau_2, \tau_1} : \mathbb{R}^N \to \mathbb{R}^N$ by:

$$\phi_{\tau_2, \tau_1}(v) \triangleq \Phi(\tau_2, \tau_1) v, \ \forall v \in \mathbb{R}^N. \tag{5.11}$$

Noticing that $\Phi(\tau_2, \tau_1)$ is invertible, it is not difficult to see that operator ϕ_{τ_2, τ_1} creates a one-to-one correspondence between the two vector spaces $\mathbf{null}_{\tau_1}(A)$ and $\mathbf{null}_{\tau_2}(A)$. As a result, the two vector spaces are of equal dimensions. $\square$

Definition 5.3. *The constant dimension of $\mathbf{null}_\tau(A)$, $\tau \geq 0$, which is independent of τ, is called nullity of chain $\{A(t)\}$ and is denoted by $\mathbf{nullity}(A)$. Moreover, the rank of chain $\{A(t)\}$ is defined by:*

$$\mathbf{rank}(A) \triangleq N - \mathbf{nullity}(A). \tag{5.12}$$

The following theorem suggests that one can investigate the size of the smallest EGC via the notion of rank.

Theorem 5.1. *For an opinion network with dynamics described by (5.1), the size of the smallest EGC is* $\mathbf{rank}(A)$.

Proof. To simplify the proof, let $r \triangleq \mathbf{rank}(A)$ and h be the size of the smallest EGC. Our aim is to show that $r = h$. Equivalently, we prove, in the following, that $h \leq r$ and $r \leq h$.

($h \leq r$): We show that there is an EGC of size r. From Lemma 5.1, it suffices to show that there exists a subset $\mathcal{S} \subset \mathcal{V}$ of size r with the property that for any initialization of the opinions of individuals in $\mathcal{V} \backslash \mathcal{S}$, there exists an initialization of the opinions of individuals in $\mathcal{S}$ such that all opinions asymptotically converge to zero. Note that $\mathbf{null}_{t_0}(A)$ is a vector space with dimension $\mathbf{nullity}(A) = N - r$. Let $\beta_1, \ldots, \beta_{N-r}$ be a basis of $\mathbf{null}_{t_0}(A)$. Notice that the column-rank of matrix

$$\left[\beta_1 | \cdots | \beta_{N-r}\right] \tag{5.13}$$

is $N-r$, and so is its row-rank. Thus, matrix (5.13) has $N-r$ linearly independent rows. Note that the choice of the $N-r$ linearly independent rows is not necessarily unique. Assume that $i_1, \ldots, i_{N-r}$ are the indices of $N-r$ independent rows of matrix (5.13). It is straightforward to show that subset $\mathcal{S} \subset \mathcal{V}$ defined by:

$$\mathcal{S} = \mathcal{V} \backslash \{i_1, \ldots, i_{N-r}\}, \tag{5.14}$$

has the desired property.

($r \leq h$): Since there exists an EGC of size h, there are $N-h$ individuals such that no matter what their initial opinions are, there is an initial opinion vector that results in all opinions asymptotically going to zero, or equivalently, an initial opinion vector that belongs to $\mathbf{null}_{t_0}(A)$. Thus, vector space $\mathbf{null}_{t_0}(A)$ has dimension greater than or equal to $N-h$, i.e., $N - r \geq N - h$. □

Remark 5.2. *Another point of interest regarding the issue of consensus, that we will not further discuss in this work, is that of the nature of the set of initial opinion vectors leading to consensus in the network with dynamics (5.1); more precisely:*

$$\{x(t_0) |\, \exists x^* \in \mathbb{R} : \lim_{t \to \infty} x(t) = x^*.\boldsymbol{1}_N\}, \tag{5.15}$$

It is straightforward to see that set (5.15) is the vector space generated by $\mathbf{null}_{t_0}(A)$ *and* $\boldsymbol{1}_N$*. Consequently, vector space (5.15) has dimension* $\mathbf{nullity}(A) + 1$.

Keeping Theorem 5.1 in mind, we focus on the notion of rank in the rest of the paper. In the following, we give the continuous time version of the definition of l_1-approximation initially introduced in Touri and Nedić (2012a) for discrete time chains.

Definition 5.4. *Chain $\{A(t)\}$ is said to be an l_1-approximation of chain $\{B(t)\}$ if:*

$$\int_0^\infty \|A(t) - B(t)\| dt < \infty, \tag{5.16}$$

where for convenience only, the norm refers to the max norm, i.e., the maximum of the absolute values of the matrix elements.

It is not difficult to show that l_1-approximation is an equivalence relation in the set of chains that are candidates of the underlying chain of an opinion network. The importance of the l_1-approximation notion in this work comes from the following lemma. The proof is eliminated due to its similarity to the proof of (Touri and Nedić, 2012a, Lemma 1).

Lemma 5.3. *The rank of a chain is invariant under an l_1-approximation.*

5.3.3 Ergodicity and Class-Ergodicity

Several other definition related to chains of matrices will be needed and are given as follows.

Definition 5.5. *Chain $\{A(t)\}$ is said to be ergodic if for every $\tau \geq 0$, its associated state transition matrix $\Phi(t, \tau)$ converges to a matrix with equal rows as $t \to \infty$.*

From Chatterjee and Seneta (1977), we know that ergodicity of $\{A(t)\}$ is equivalent to the occurrence of unconditional consensus in (5.1).

Definition 5.6. *Chain $\{A(t)\}$ is class-ergodic if for every $\tau \geq 0$, $\lim_{t\to\infty} \Phi(t, \tau)$ exists but has possibly distinct rows.*

It is known that chain $\{A(t)\}$ is class-ergodic if and only if multiple consensus occurs in (5.1) unconditionally (see Bolouki and Malhamé (2013); Touri and Nedić (2012b)). We define, in what follows, the ergodicity classes of a chain according to Touri and Nedić (2012a).

Definition 5.7. *For an opinion network with state transition matrix $\Phi(t, \tau)$, $t \geq \tau \geq 0$, two individuals $i, j \in \mathcal{V}$ are said to be mutually weakly ergodic if and only if for every $\tau \geq 0$:*

$$\lim_{t\to\infty} \|\Phi'_i(t, \tau) - \Phi'_j(t, \tau)\| = 0. \tag{5.17}$$

It is easy to see that the relation of being mutually weakly ergodic is an equivalence relation on $\mathcal{V}$. The equivalence classes of this relation are referred to as *ergodicity classes* in this paper. Indeed, these equivalence classes form a partitioning of $\mathcal{V}$, and while in some cases they may simply be singletons, they can always be defined for an arbitrary chain {A(t)}.

If chain $\{A(t)\}$ is class-ergodic, i.e,. $\lim_{t\to\infty}\Phi_i'(t,\tau)$ exists for every $i \in \mathcal{V}$ and $\tau \geq 0$, then $i, j \in \mathcal{V}$ are in the same ergodicity class if $\lim_{t\to\infty}\Phi_i'(t,\tau) = \lim_{t\to\infty}\Phi_i'(t,\tau)$, for every $\tau \geq 0$. We refer to the ergodicity classes of a class-ergodic chain as *ergodic classes*.

5.4 A Geometric Interpretation of the Rank

In this Section, we employ a geometric approach to analyze the asymptotic properties of a chain of matrices . This approach, which can be used for both the continuous and discrete time cases, will help us to (i) geometrically interpret the rank of a general time-varying chain, (ii) identify an upper bound for the rank, and (iii) investigate the limiting behavior of a large class of time-varying chains, namely Class $\mathcal{P}^*$ as discussed in Section 5.7.

For time-varying chain $\{A(t)\}_{t\geq 0}$, define $C_{t,\tau}$, $t \geq \tau \geq 0$ as the convex hull of points in $\mathbb{R}^N$ corresponding to the columns of the transpose of associated state transition matrix $\Phi(t,\tau)$. Note that $C_{t,\tau}$ is a polytope, with no more than N vertices, in $\mathbb{R}^N$. We recall that each column of $\Phi'(t,\tau)$ is a stochastic vector, i.e., its elements are non-negative and add up to 1. We now have the following lemma regarding convex hull $C_{t,\tau}$.

Lemma 5.4. *For every $t_2 \geq t_1 \geq \tau$, we have: $C_{t_2,\tau} \subset C_{t_1,\tau}$, i.e., polytopes $C_{t,\tau}$, for an arbitrary fixed τ, form a monotone decreasing sequence of polytopes in $\mathbb{R}^N$.*

Proof. Note that:

$$\Phi(t_2,\tau) = \Phi(t_2,t_1)\Phi(t_1,\tau), \tag{5.18}$$

or equivalently,

$$\Phi'(t_2,\tau) = \Phi'(t_1,\tau)\Phi'(t_2,t_1) \tag{5.19}$$

Since $\Phi'(t_2,t_1)$ is a column-stochastic matrix, relation (5.19) implies that each column of $\Phi'(t_2,\tau)$ is a convex combination of the columns of $\Phi'(t_1,\tau)$. Therefore, each column of $\Phi'(t_2,\tau)$ lies in or on $C_{t_1,\tau}$, and the lemma is proved. □

Lemma 5.4 shows that for a fixed $\tau \geq 0$, polytopes $C_{t,\tau}$'s, $t \geq \tau$, are nested in $\mathbb{R}^N$. An example of these nested polytopes projected on a two-dimensional subspace of $\mathbb{R}^N$ is depicted in Fig. 5.1.

Note that for every $\tau \geq 0$, $\lim_{t\to\infty} C_{t,\tau}$ exists and is also a polytope in $\mathbb{R}^N$ due to the existence of a uniform upper bound, namely N, on the number of vertices of the nested polytopes. Let C_τ denote the limiting polytope and c_τ be the number of its vertices.

Lemma 5.5. *c_τ, $\tau \geq 0$, is independent of τ.*

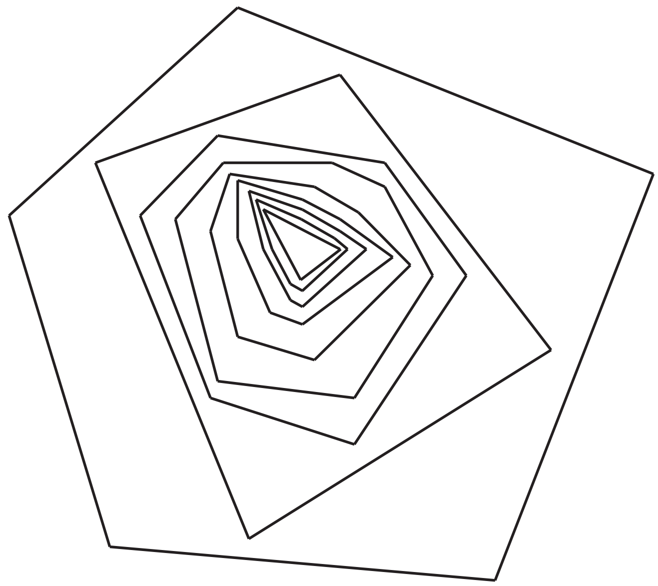

Figure 5.1 Nested polygons converging to a triangle.

Proof. Assume that $\tau_2 \geq \tau_1 \geq 0$ are two arbitrary time instants. Define linear operator $\phi'_{\tau_2,\tau_1} : \mathbb{R}^N \to \mathbb{R}^N$ by:

$$\phi'_{\tau_2,\tau_1}(v) \triangleq \Phi'(\tau_2, \tau_1)v, \ \forall v \in \mathbb{R}^N. \tag{5.20}$$

Note now that from (5.19), for $t \geq \tau_2 \geq \tau_1 \geq 0$ we have:

$$\Phi'(t, \tau_1) = \Phi'(\tau_2, \tau_1)\Phi'(t, \tau_2). \tag{5.21}$$

Therefore, in view of (5.21) by taking t to infinity, the vertices of C_{τ_2} are uniquely mapped to vectors in $\mathbb{R}^N$ which because of the linearity of map (5.20), will play the role of vertices for the generation of convex hull C_{τ_1}. Also, it is not difficult to show that the images of vertices of C_{τ_2} must remain vertices of C_{τ_1}, for if one of the images of a vertex of C_{τ_2}, say v, turned out to be a convex combination of other vertices of C_{τ_1}, this would also be true for the inverse images of these vertices (also vertices of C_{τ_2} due to invertibility of matrix $\Phi'(\tau_2, \tau_1)$), and v would then fail to be a vertex of C_{τ_2}. In conclusion, C_{τ_1} and C_{τ_2} will have the same number of vertices, and (5.20) constitutes a one to one map between corresponding pairs of vertices. $\square$

Let integer c be the constant value of c_τ, $\tau \geq 0$. We will show later in this section that c is equal to $\mathbf{rank}(A)$. To prove this, we first state the following two lemmas.

Lemma 5.6. $\mathbf{rank}(A)$ *is equal to the dimension of the vector space generated by the vectors* *corresponding to the vertices of C_τ, for every $\tau \geq 0$.*

Proof. It suffices to prove Lemma 5.6 for $\tau = 0$. Let $v_1, \ldots, v_c \in \mathbb{R}^N$ be the c vertices of C_0. It is easy to see that for any $u \in \mathbb{R}^N$:

$$u \in \mathcal{N}_0(A) \iff v_i' u = 0, \forall i, 1 \leq i \leq c. \tag{5.22}$$

It implies that the dimension of the vector space generated by $v_1, \ldots, v_c$ is $N - \textbf{nullity}(A)$, which proves the lemma. $\square$

Lemma 5.7. *For every $\tau \geq 0$, the vectors corresponding to the vertices of C_τ are linearly independent.*

Proof. It is sufficient to prove the lemma for $\tau = 0$, i.e., to show that the vertices of C_0, namely $v_1, \ldots, v_c$, are linearly independent. Assume that $\alpha_1, \ldots, \alpha_c \in \mathbb{R}$ are such that:

$$\sum_{i=1}^{c} \alpha_i v_i = 0. \tag{5.23}$$

We note that vector v_i, $1 \leq i \leq c$, must lie outside of the convex hull of vectors v_j's, $j \neq i$, for otherwise it would not qualify as a vertex. For every i, $1 \leq i \leq c$, let w_i be the projection of v_i on the convex hull of v_j's, $j \neq i$. Define the following positive numbers:

$$\epsilon \triangleq \frac{1}{4} \min\{\|v_i - w_i\| \,|\, 1 \leq i \leq c\}, \tag{5.24}$$

and:

$$\epsilon_1 \triangleq \epsilon/(2N). \tag{5.25}$$

Because C_0 is the limit of $C_{t,0}$ as t goes to infinity, there must exist a sufficiently large time $T \geq 0$, such that for $t \geq T$, every point in $C_{t,0}$ lies within an ϵ_1-distance of C_0. As depicted in Fig. 5.2, for every i, $1 \leq i \leq c$, let l_i be the hyperplane in $\mathbb{R}^N$ distant ϵ from v_i, crossing segment $v_i w_i$ and orthogonal to it. Let also m_i be the hyperplane which is parallel to l_i, on the other side of v_i, distant ϵ_1 from v_i.

Define for every i, $1 \leq i \leq c$:

$$S^i = \{j \in \mathcal{V} \,|\, \Phi_j'(T, 0) \text{ lies in the strip margined by } l_i, m_i\}. \tag{5.26}$$

Note that by the assumption, every point in $C_{T,0}$, including $\Phi_j'(T, 0)$, lies within an ϵ_1-distance of C_0. Therefore, $\Phi_j'(T, 0)$ must lie on the same side of m_i as v_i does. In other words, $\Phi_j'(T, 0)$ either lies in the strip margined by l_i and m_i or lies on the side of l_i opposite to v_i (below l_i

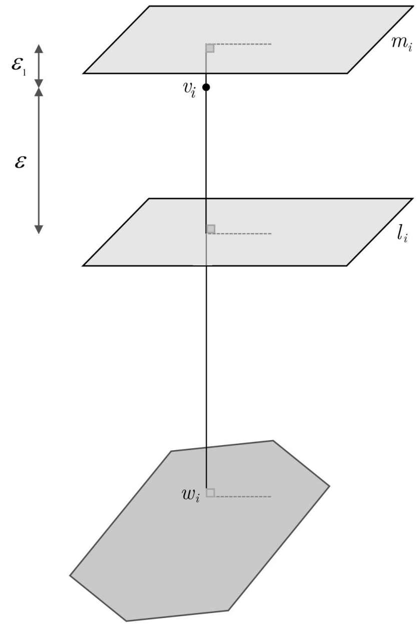

Figure 5.2 Planes l_i and m_i are orthogonal to segment $v_i w_i$.

in Fig. 5.2). This implies that S_i, $1 \le i \le c$, is non-empty. Indeed otherwise, $\Phi'_j(T,0)$ would lie below l_i in Fig. 5.2 for every j resulting in $C_{T,0}$ also lying below l_i, which would be a contradiction since $C_{T,0}$ must contain C_0 and v_i in particular. One can also show that S^i's, $1 \le i \le c$, are pairwise disjoint sets. More specifically, one can show that any point of $C_{T,0}$ that lies in the intersection of any two of sets S^i's cannot be within ϵ-distance of C_0, and since $\epsilon > \epsilon_1$, this would violate the defining property of T. C_0 being the limit of shrinking convex hulls $C_{t,0}$'s, it follows that for $i = 1, \ldots, c$, there exists sequences $\{i_t\}$ of individuals such that $\Phi'_{i_t}(t,0)$ converges to v_i. Therefore, after some finite time, we have the following inequality:

$$\|\Phi'_{i_t}(t,0) - v_i\| < \epsilon_1. \tag{5.27}$$

Without loss of generality, we can assume that the inequality (5.27) holds for every $t \ge T$ (otherwise, we would proceed by replacing T with T', $T' > T$, such that inequality (5.27) holds for every $t \ge T'$). We have for every $t \ge T$:

$$\begin{aligned}
\Phi'_{i_t}(t,0) &= \Phi'(T,0)\Phi'_{i_t}(t,T) \\
&= \textstyle\sum_{j \in \mathcal{V}} \Phi_{i_t,j}(t,T)\Phi'_j(T,0) \\
&= \textstyle\sum_{j \notin S^i} \Phi_{i_t,j}(t,T)\Phi'_j(T,0) + \sum_{j \in S^i} \Phi_{i_t,j}(t,T)\Phi'_j(T,0).
\end{aligned} \tag{5.28}$$

We now show that for every i, $1 \leq i \leq c$, the following two inequalities must hold:

$$\sum_{j \notin S^i} \Phi_{i_t,j}(t,T) < 2/(2N+1), \tag{5.29}$$

$$\sum_{j \in S^i} \Phi_{i_t,j}(t,T) > 1 - 2/(2N+1). \tag{5.30}$$

To prove (5.29) and (5.30), we use (5.28) to find a lower bound for the distance from $\Phi'_{i_t}(t,0)$, $t \geq T$, to hyperplane m_i as drawn in Fig. 5.2. Remember that if $j \in S^i$, then, $\Phi'_j(T,0)$ lies in the strip margined by m_i and l_i, while if $j \notin S^i$, then, $\Phi'_j(T,0)$ lies below l_i in Fig. 5.2. For a fixed i, $1 \leq i \leq c$, let $\eta \triangleq \sum_{j \notin S^i} \Phi_{i_t,j}(t,T)$. $\Phi(t,T)$ being row-stochastic, it immediately follows that $\sum_{j \in S^i} \Phi_{i_t,j}(t,T) = 1 - \eta$. Using (5.28), we now conclude that:

$$\eta(\epsilon_1 + \epsilon) + (1-\eta).0 \tag{5.31}$$

is a lower bound for the distance from $\Phi'_{i_t}(t,0)$, $t \geq T$, to hyperplane m_i. This distance, on the other hand, is upper bounded by $2\epsilon_1$ since inequality (5.27) is satisfied for every $t \geq T$. Thus, we must have:

$$\eta(\epsilon_1 + \epsilon) + (1-\eta).0 < 2\epsilon_1, \tag{5.32}$$

which immediately results in $\eta < 2/(2N+1)$ (remember that $\epsilon = 2N\epsilon_1$), and inequalities (5.29) and (5.30) follow. Now remember by construction that $\lim_{t\to\infty} \Phi'_{i_t}(t,0) = v_i$ where v_i is a given vertex of C_0. Furthermore, noting that:

$$\Phi'_{i_t}(t,0) = \Phi'(T,0)\Phi'_{i_t}(t,T), \tag{5.33}$$

and taking limits on both sides as t goes to infinity, it follows that $\lim_{t\to\infty} \Phi'_{i_t}(t,T)$ is the image of a vertex of C_0 and therefore (following the proof of Lemma 5.5) is itself a vertex of C_T, say u_i. Considering (5.30) again, and taking limits as $t \to \infty$, one can conclude:

$$\sum_{j \in S^i} (u_i)_j \geq 1 - 2/(2N+1), \tag{5.34}$$

and consequently:

$$\sum_{j \notin S^i} (u_i)_j \leq 2/(2N+1). \tag{5.35}$$

Inequality (5.34) can be established for $i = 1, \ldots, c$, where u_i, $i = 1, \ldots, c$ are the vertices of C_T. Recalling linear operator ϕ_{τ_2,τ_1} from (5.20) one can write for some permutation σ over set $\{1,\ldots,c\}$:

$$u_i = \Phi'(T,0)v_{\sigma(i)},\ \forall i,\ 1 \le i \le c, \tag{5.36}$$

Combining relations (5.23) and (5.36) yields:

$$\sum_{i=1}^{c} \alpha_{\sigma(i)} u_i = 0, \tag{5.37}$$

If we now assume that k, $1 \le k \le c$, is such that:

$$|\alpha_{\sigma(k)}| = \max_{1\le i\le c}\{|\alpha_i|\} \triangleq \alpha, \tag{5.38}$$

Now noting that (5.34) and (5.35) hold only for the vertex u_i which is the image of v_i, and that the S^i's are disjoint sets of agents, one can write the following:

$$\begin{aligned}
0 &= |\textstyle\sum_{j\in S^k}\sum_{i=1}^{c}\alpha_{\sigma(i)}(u_i)_j| \\
&= |\textstyle\sum_{j\in S^k}\alpha_{\sigma(k)}(u_k)_j + \sum_{j\in S^k}\sum_{i\neq k}\alpha_{\sigma(i)}(u_i)_j| \\
&\ge |\alpha_{\sigma(k)}|.|\textstyle\sum_{j\in S^k}(u_k)_j| - \sum_{i\neq k}\left(|\alpha_{\sigma(i)}|.\sum_{j\in S^k}(u_i)_j\right) \\
&\ge |\alpha_{\sigma(k)}|.|\textstyle\sum_{j\in S^k}(u_k)_j| - \sum_{i\neq k}\left(|\alpha_{\sigma(i)}|.\sum_{j\notin S^i}(u_i)_j\right) \\
&\ge \alpha(1-2/(2N+1)) - \alpha(c-1).2/(2N+1) = \alpha(2(N-c)+1)/(2N+1) \\
&> 0,
\end{aligned} \tag{5.39}$$

which is a contradiction. Thus, we must have $\alpha = 0$, which means $\alpha_i = 0$, $\forall i$, $1 \le i \le c$. This proves the lemma. □

Theorem 5.2. **rank**(A) *is equal to c, i.e, the constant value of c_τ, $\tau \ge 0$, where c_τ is the number of vertices of limiting polytope C_τ.*

Proof. Theorem 5.2 is an immediate result of Lemmas 5.6 and 5.7. □

Combining Theorems 5.1 and 5.2 result in the following corollary.

Corollary 5.1. *The size of the smallest EGC of a network with dynamics (5.1) is c.*

Lemma 5.8. *c is less than or equal to the number of ergodicity classes.*

Proof. Recall limiting polytope C_0 with vertices $v_1,\ldots,v_c$ from earlier in the section. Remember, from the proof of Lemma 5.7, that for $i = 1,\ldots,c$, there exists sequences $\{i_t\}$ of individuals such that $\Phi'_{i_t}(t,0)$ converges to v_i. Let:

$$\epsilon_2 = \frac{1}{3}\min\{\|v_i - v_j\|\,|i,j \in \mathcal{V},\ i \neq j\}. \tag{5.40}$$

By definition of ergodicity classes, there exists $T \geq 0$ such that for every $t \geq T$, for a fixed τ, and for every i, j in the same ergodicity class, we have:

$$\|\Phi'_i(t,\tau) - \Phi'_j(t,\tau)\| < \epsilon_2. \tag{5.41}$$

On the other hand, there exists $T' > 0$ such that for every $t \geq T'$, and $i = 1, \ldots, c$, we have:

$$\|\Phi'_{i_t}(t,0) - v_i\| < \epsilon_2. \tag{5.42}$$

Therefore, for every $t \geq T'$, and $i \neq j$, $1 \leq i, j \leq c$, we must have:

$$\begin{aligned} 3\epsilon_2 \leq \|v_i - v_j\| &\leq \|v_i - \Phi'_{i_t}(t,0)\| + \|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\| + \|\Phi'_{j_t}(t,0) - v_j\| \\ &< \epsilon_2 + \|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\| + \epsilon_2, \end{aligned} \tag{5.43}$$

where the first inequality above is a result of (5.40), the second inequality is the triangle inequality, and the third inequality is a consequence of (5.42). From (5.43), we now have:

$$\|\Phi'_{i_t}(t,0) - \Phi'_{j_t}(t,0)\| > \epsilon_2, \, \forall t \geq T'. \tag{5.44}$$

Taking (5.41) into account, from (5.44) we conclude that i_t and j_t cannot be in the same ergodicity class for every $t \geq \max\{T, T'\}$. Thus, there are at least c distinct ergodicity classes, and the lemma is proved. $\square$

Corollary 5.2. *For an arbitrary chain $\{A(t)\}$, $\mathbf{rank}(A)$ is less than or equal to the number of ergodicity classes of $\{A(t)\}$.*

Corollary 5.3. *For an opinion network with dynamics (5.1), the size of the smallest EGC is upper bounded by the number of ergodicity classes of $\{A(t)\}$.*

Remark 5.3. *In case $\{A(t)\}$, the underlying chain of a network with dynamics (5.1), is class-ergodic, the occurrence of multiple consensus in the network is guaranteed, and the number of ergodic classes becomes equal to the number of consensus clusters. Yet this number may be larger than the size of the smallest EGC of the network. In other words, there may exist an EGC in which some of the consensus clusters have no representative. As a simple illustrative example, consider system (5.1) of three individuals with a fixed underlying chain:*

$$A(t) = \begin{bmatrix} 0 & 0 & 0 \\ 1/3 & -1 & 2/3 \\ 0 & 0 & 0 \end{bmatrix}, \, \forall t \geq 0. \tag{5.45}$$

We then have:

$$\lim_{t\to\infty} x(t) = \begin{bmatrix} x_1(t_0) \\ (x_1(t_0)+2x_3(t_0))/3 \\ x_3(t_0) \end{bmatrix}. \tag{5.46}$$

Notice also that for the corresponding state transition matrix we have:

$$\lim_{t\to\infty} \Phi(t,\tau) = \begin{bmatrix} 1 & 0 & 0 \\ 1/3 & 0 & 2/3 \\ 0 & 0 & 1 \end{bmatrix}, \forall \tau \geq 0. \tag{5.47}$$

Therefore, each individual forms a consensus cluster, i.e., there are three consensus clusters. However, subgroup $\{1,3\}$ with size two, is an EGC of the network. In other words, starting at an arbitrary initial time $t_0 \geq 0$, irrespective of the initial opinion of individual 2, an agreement on value x^* is achieved if individuals 1 and 3 initialize their opinions at x^*.

5.5 Lower Bounds on the Rank of chains

In this section, we clarify how the underlying chain of a network with dynamics (5.1) imposes lower bounds on the size of its smallest EGC, which is equal to $\mathbf{rank}(A)$. We recall the following definition from Bolouki and Malhamé (2012a); Hendrickx and Tsitsiklis (2013).

Definition 5.8. *The unbounded interactions graph of a chain $\{A(t)\}$, $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$, is a fixed directed graph such that for every distinct nodes $i,j \in \mathcal{V}$, $(i,j) \in \mathcal{E}_1$ if and only if:*

$$\int_0^\infty a_{ji}(t)dt = \infty. \tag{5.48}$$

In other words, a link is drawn from i to j if the total influence of individual i on individual j is unbounded over the infinite time interval.

Definition 5.9. *A subset $\mathcal{S}' \subset \mathcal{V}$ is called a s-root of $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$ if for every node $i \in \mathcal{V}$, we have $i \in \mathcal{S}'$ or there exists $j \in \mathcal{S}'$ such that i is reachable from j.*

Theorem 5.3. *Let $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$ be the unbounded interaction graph associated with chain $\{A(t)\}$. Then, $\mathbf{rank}(A)$ is greater than or equal to the size of the smallest s-root of $\mathcal{H}_1(\mathcal{V},\mathcal{E}_1)$.*

Proof. Form a chain $\{B(t)\}$ from chain $\{A(t)\}$ by eliminating all influences that individual $i \in \mathcal{V}$ gets from individual $j \in \mathcal{V}$ if $(j,i) \notin \mathcal{E}_1$. More specifically, for every $i \neq j \in \mathcal{V}$ and $t \geq 0$, we have:

$$b_{ij}(t) = \begin{cases} a_{ij}(t) & \text{if } (j,i) \in \mathcal{E}_1 \\ 0 & \text{if } (j,i) \notin \mathcal{E}_1 \end{cases} \tag{5.49}$$

and $b_{ii}(t) = -\sum_{j \neq i} b_{ij}(t)$, for every $i \in \mathcal{V}$ and $t \geq 0$. Since chain $\{B(t)\}$ is an l_1-approximation of chain $\{A(t)\}$, from Lemma 5.3, the two chains share the same rank. Notice also that the two chains share the same unbounded interactions graph. Thus, it suffices to prove Theorem 5.3 for chain $\{B(t)\}$. Consider an opinion network with underlying chain $\{B(t)\}$:

$$\dot{y}(t) = B(t)y(t),\ t \geq t_0, \tag{5.50}$$

where $y(t) \in \mathbb{R}^N$ is the vector of opinions. Since $\mathbf{rank}(B)$ is the size of the smallest EGC of the network with dynamics (5.50), it is sufficient to show that every EGC of the network with dynamics (5.50) is a s-root of $\mathcal{H}_1$. Assume, on the contrary, that subset $\mathcal{S} \subset \mathcal{V}$ is an EGC which is not a s-root of $\mathcal{H}_1$. Define:

$$n(\mathcal{S})\ \triangleq\ \mathcal{S}\ \cup\ \{i \mid i \in \mathcal{V},\ \exists j \in \mathcal{S} : i \text{ is reachable from } j \text{ in } \mathcal{H}_1\} \tag{5.51}$$

Since $\mathcal{S}$ is not a s-root, $n(\mathcal{S}) \subsetneq \mathcal{V}$. From the definition of $n(\mathcal{S})$, it is easy to see that there is no link from $n(\mathcal{S})$ to $\mathcal{V} \backslash n(\mathcal{S})$ in $\mathcal{H}_1$. According to the way that chain $\{B(t)\}$ was constructed, this means that $n(\mathcal{S})$ has zero influence on $\mathcal{V} \backslash n(\mathcal{S})$ at any time instant. Thus, since $\mathcal{S} \subset n(\mathcal{S})$, individuals in $\mathcal{S}$ cannot, in general, lead individuals in $\mathcal{V} \backslash n(\mathcal{S})$ to agreeing on an arbitrary value x^*. For instance, given a desired consensus value x^*, if the opinions of individuals in $\mathcal{V} \backslash n(\mathcal{S})$ are all initialized at value $x^* + 1$, they will never change, and consequently, they will never converge to x^*. Thus, $\mathcal{S}$ is not an EGC, which completes the proof. $\square$

An important special case of Theorem 5.3 is described in the following. Let us first define the continuous time counterpart of the *infinite flow graph* of a chain according to Touri and Nedić (2011).

Definition 5.10. *The infinite flow graph $\mathcal{H}_2(\mathcal{V}, \mathcal{E}_2)$ of a given chain $\{A(t)\}$, is an undirected graph formed as follows: for two distinct nodes $i, j \in \mathcal{V}$, draw a link between i and j in $\mathcal{H}_2$, if and only if:*

$$\int_0^\infty (a_{ij}(t) + a_{ji}(t))dt = \infty \tag{5.52}$$

We now have the following lower bound on the rank of a chain which is a special case of Theorem 5.3.

Corollary 5.4. $\mathbf{rank}(A)$ *is greater than or equal to the number of connected components of the infinite flow graph associated with* $\{A(t)\}$.

5.6 Rank of Time-Invariant (TI) chains

Let $\{A(t)\}$ be a TI chain, i.e., $A(t) = \hat{A}$, $\forall t \geq 0$, where $\hat{A}$ is a fixed matrix with the property that each of its rows adds up to zero and its off-diagonal elements are non-negative. Assume that $\mathrm{rank}(\hat{A})$ and $\mathrm{nullity}(\hat{A})$ represent the rank and the nullity of $\hat{A}$. Notice that roman style is used for matrix operators as opposed to the chain operators so as to avoid any ambiguity. For state transition matrix $\Phi(t,\tau)$ associated with TI chain $\{\hat{A}\}$, we have:

$$\Phi(t,\tau) = e^{\hat{A}(t-\tau)},\ t \geq \tau \geq 0. \tag{5.53}$$

Note that $\hat{A}$ is marginally stable and has all negative eigenvalues but one eigenvalue zero with algebraic multiplicity $\mathrm{nullity}(\hat{A})$. Thus, $\lim_{t-\tau\to\infty} \Phi(t,\tau)$ exists, and the limit has eigenvalue zero with algebraic multiplicity $\mathrm{rank}(\hat{A})$ and eigenvalue one with algebraic multiplicity $\mathrm{nullity}(\hat{A})$. Hence:

$$\mathbf{rank}(A) = \mathrm{nullity}(\hat{A}). \tag{5.54}$$

Employing a graph theoretic approach, treating $\hat{A}$ as the Laplacian of its associated *weighted directed* graph, $\mathrm{nullity}(\hat{A})$ represents the size of the smallest s-root of the graph (see Fig. 5.3).

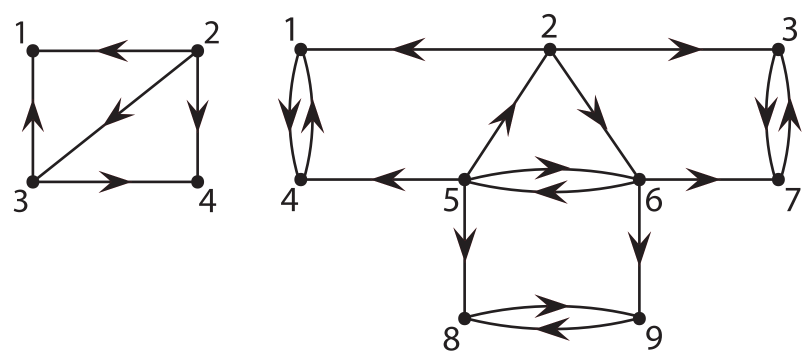

Figure 5.3 Unweighted underlying graph of two TI linear algorithms. $\{2\}$ (left) and either of $\{2\}$, $\{5\}$, and $\{6\}$ (right) are the smallest s-roots.

Since an unweighted version of the graph described above serves as the unbounded interactions graph associated with TI chain $\{A(t)\}$, $A(t) = \hat{A}$, $\forall t \geq 0$, we have the following corollary.

Corollary 5.5. *For a TI chain $\{A(t)\}$, the lower bound provided in Theorem 5.3 is achieved. More specifically,* $\mathbf{rank}(A)$ *is size of the smallest s-root of the unbounded interactions graph associated with $\{A(t)\}$.*

Remember that any TI chain $\{A(t)\}$ is class-ergodic and the number of ergodic classes provides an upper bound for $\mathbf{rank}(A)$ according to Corollary 5.2. For example, for the underlying graphs depicted in Fig. 5.3, the number of ergodic classes are 4 (left) and 6 (right).

The graph interpretation of the notion of rank explains the following two properties:

(i) For any TI chain $\{A(t)\}$ and $\alpha > 0$:

$$\mathbf{rank}(\{\alpha A(t)\}) = \mathbf{rank}(\{A(t)\}). \tag{5.55}$$

(ii) For any two TI chains $\{A(t)\}$ and $\{B(t)\}$,

$$\mathbf{rank}(\{A(t) + B(t)\}) \leq \min\Big\{ \mathbf{rank}(\{A(t)\}), \mathbf{rank}(\{B(t)\}) \Big\}. \tag{5.56}$$

Remark 5.4. *While Statement (i) seems to hold for any time-varying chain $\{A(t)\}$ as well, there exist time-varying chains $\{A(t)\}$ and $\{B(t)\}$ that do not satisfy Statement (ii). This means that more interactions between agents may surprisingly increase the size of the smallest EGC of a network. The following is an example; let:*

$$A(t) = \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \ \textit{if } t \in [2^{2k} - 1, 2^{2k}),\ k \in \mathbb{N}, \tag{5.57}$$

and,

$$A(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \ \textit{if } t \in [2^{2k}, 2^{2k+1} - 1),\ k \in \mathbb{N}, \tag{5.58}$$

and $A(t) = \boldsymbol{0}_{3\times 3}$ elsewhere. Let also:

$$B(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix} \ \textit{if } t \in [2^{2k+1} - 1, 2^{2k+1}),\ k \in \mathbb{N}, \tag{5.59}$$

and,

$$B(t) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \ \textit{if } t \in [2^{2k+1}, 2^{2k+2} - 1),\ k \in \mathbb{N}, \tag{5.60}$$

and $B(t) = \boldsymbol{0}_{3\times 3}$ elsewhere. Note that at every time instant either $A(t)$ or $B(t)$ is $\boldsymbol{0}_{3\times 3}$. It *is easy to see that both $\{A(t)\}$ and $\{B(t)\}$ are ergodic chains. More specifically, for every $\tau \geq 0$, we have:*

$$\lim_{t\to\infty} \Phi_A(t,\tau) = \begin{bmatrix}0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}', \tag{5.61}$$

and,

$$\lim_{t\to\infty} \Phi_B(t,\tau) = \begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}'. \tag{5.62}$$

Therefore, $\mathbf{rank}(A) = \mathbf{rank}(B) = 1$. *However, one can show that* $\mathbf{rank}(\{A(t)+B(t)\}) = 2$. *More precisely, subgroup $\{1,3\}$ forms the smallest EGC of the network with underlying chain $\{A(t)+B(t)\}$.*

5.7 Rank of chains in Class $\mathcal{P}^*$

From the fundamental work Kolmogoroff (1936), it is known that for every state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, associated with a chain $\{A(t)\}$, there exists a sequence of stochastic row vectors $\{\pi(t)\}$, called an *absolute probability sequence*, such that:

$$\pi(\tau) = \pi(t)\Phi(t,\tau),\ \forall t,\tau,\ t \geq \tau \geq 0. \tag{5.63}$$

Remember that by a stochastic vector, we mean a vector with elements adding up to 1. We may now extend (Touri and Nedić, 2014, Definition 3) to the continuous time case in the following.

Definition 5.11. *A chain $\{A(t)\}$ is said to be in Class $\mathcal{P}^*$ if its associated state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$ admits an absolute probability sequence $\{\pi(t)\}$ such that for some constant $p^* > 0$:*

$$\pi(t) > p^*,\ \forall t \geq 0. \tag{5.64}$$

It is possible to characterize chains of Class $\mathcal{P}^*$ more concretely. To do so, we first state the following lemma.

Lemma 5.9. *For every $j \in \mathcal{V}$,*

$$\pi_j(\tau) \leq \inf\left\{\sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau)\,|\, t \geq \tau\right\}. \tag{5.65}$$

Proof. Obvious, since for every $t \geq \tau$:

$$\pi_j(\tau) = \pi(t)\Phi^j(t,\tau) = \sum_{i\in\mathcal{V}} \pi_i(t)\Phi_{i,j}(t,\tau) \leq \sum_{i\in\mathcal{V}} \Phi_{i,j}(t,\tau). \tag{5.66}$$

□

We now have the following lemma that provides an alternative definition of chains in Class $\mathcal{P}^*$.

Lemma 5.10. *A chain $\{A(t)\}$ is in Class $\mathcal{P}^*$ if and only if for its state transition matrix $\Phi(t,\tau)$, $t \geq \tau \geq 0$, we have:*

$$\inf_{t,\tau} \left\{ \sum_{i \in \mathcal{V}} \Phi_{i,j}(t,\tau) \,|\, t \geq \tau \geq 0 \right\} > 0, \, \forall j \in \mathcal{V}. \tag{5.67}$$

Proof. The "only if" part is an immediate result of Lemma 5.9, and the "if" part is a result of the way an absolute probability sequence can be obtained in Kolmogoroff (1936) by always choosing to initialize agent probabilities on finite intervals with a uniform distribution. □

Lemma 5.10 roughly implies that the underlying chain of a system is in Class $\mathcal{P}^*$, if and only if the opinion of any individual, at any time, continues to have influence on the formation of individuals' opinions at all future times. We now state a theorem on the class-ergodicity of chains in Class $\mathcal{P}^*$ (see Theorem 4.6 of Chapter 4).

Theorem 5.4. *Every chain $\{A(t)\}$ in Class $\mathcal{P}^*$ is class-ergodic. Furthermore, the number of ergodic classes is equal to the number of connected components of the infinite flow graph of chain $\{A(t)\}$.*

Theorem 5.4 implies that if chain $\{A(t)\}$ is in Class $\mathcal{P}^*$, the upper bound provided for its rank in Corollary 5.2 is equal the lower bound provided in Corollary 5.4. Therefore, both bounds become equal to $\mathbf{rank}(A)$.

Corollary 5.6. *The rank of a chain in Class $\mathcal{P}^*$ is determined by the number of connected components of the infinite flow graph associated with the chain.*

5.8 Full-Rank chains

One can characterize chains with maximum possible rank as the following.

Theorem 5.5. *A chain $\{A(t)\}$ is full-rank, i.e., $\mathbf{rank}(A) = N$ if and only if $\{A(t)\}$ is an l_1-approximation of the neutral chain, i.e., the chain of matrix $\boldsymbol{0}_{N\times N}$.*

Proof. The sufficiency is immediately implied using Lemma 5.3 and taking into account that the neutral chain is full-rank. To prove the necessity, assume that $\{A(t)\}$ is full-rank. We may now once again take advantage of our geometric framework developed in Section 5.4 based on the associated state transition matrix. Recall that c is defined by the number of vertices of limiting polytope C_τ for an arbitrary $\tau \geq 0$. Since $\mathbf{rank}(A) = c$, we conclude that $c = N$. Letting $v_1, \ldots, v_N$ be the N vertices of C_0, for a permutation σ over $\{1, \ldots, N\}$, we must have:

$$\lim_{t \to \infty} \Phi'(t, 0) = \left[v_{\sigma(1)} | \cdots | v_{\sigma(N)} \right], \tag{5.68}$$

since each column of $\Phi'(t, 0)$ is a continuous function of t such that its distance from $\{v_1, \ldots, v_N\}$ vanishes as t grows large. Recalling:

$$\Phi(t, 0) = \Phi(t, \tau)\Phi(\tau, 0), \ \forall t \geq \tau \geq 0, \tag{5.69}$$

and taking into account that, based on Lemma 5.7, the columns of the RHS of relation (5.68) are linearly independent stochastic vectors, for a sufficiently large $T \geq 0$, $\Phi(t, \tau)$ is arbitrarily close to the $N \times N$ identity matrix for every $t \geq \tau \geq T$. In particular, $\Phi(t, \tau)$ has positive diagonal elements (well away from zero) for every $t \geq \tau \geq T$. Form chain $\{B(t)\}$ from $\{A(t)\}$ by eliminating all interactions between individuals over time interval $[0, T)$. Then, the state transition matrix associated with chain $\{B(t)\}$ has positive diagonal elements all the times. Recalling Lemma 5.10, we conclude that chain $\{B(t)\}$ is in Class $\mathcal{P}^*$. On the other hand, chain $\{B(t)\}$ is an l_1-approximation of chain $\{A(t)\}$ due to boundedness of interactions over time interval $[0, T)$. Consequently, $\mathbf{rank}(B) = \mathbf{rank}(A) = N$. Theorem 5.4 now implies that $\mathbf{rank}(B) = N$ is the number of connected components of the infinite flow graph associated with chain $\{B(t)\}$. This completes the proof since the two chains share the same infinite flow graph. $\square$

Assume that the infinite flow graph of chain $\{A(t)\}$, i.e., $\mathcal{H}_2(\mathcal{V}, \mathcal{E}_2)$, has h_2 connected components. Form chain $\{B(t)\}$, which is an l_1-approximation of $\{A(t)\}$ by eliminating all interactions between distinct connected components. Since the subchain corresponding to each connected component is full-rank if and only if it contains a single node, the following proposition follows from Lemma 5.3, that provides an upper bound for $\mathbf{rank}(A)$.

Proposition 5.1. *Let $\{A(t)\}$ be a time-varying chain with infinite flow graph $\mathcal{H}_2$. Then:*

$$\mathbf{rank}(A) \leq N - h_2', \tag{5.70}$$

where h_2' is the number of connected components of $\mathcal{H}_2$ containing two or more nodes.

5.9 Discrete Time Analysis

In this section, we turn our attention to the case in which the opinions of the individuals are updated at discrete time instants. Our aim is to characterize EGC's in a network for the discrete time case. To this aim, we adopt, with a slight modification, the same approach followed in the continuous time case, i.e., an approach based on the notion of rank. After we define the rank of a discrete time chain, we carry out the discrete time counterpart of our statements in Sections 5.3–5.8.

Remember that in this section, time variables t, τ, t_0, etc. refer to the discrete time indices. Let $\{A(t)\}_{t\geq 0}$ be a time-varying chain of row-stochastic square matrices of size N. A row-stochastic matrix, or simply stochastic matrix, is a matrix with non-negative elements and the property that its each row elements sum up to 1 . Discrete time chains of matrices, that we deal with in this paper, are assumed to be chains of stochastic matrices. Indeed, $A(t)$ can be viewed as the transition matrices of a time inhomogeneous Markov chain. Let dynamics of an opinion network be described by the following discrete time distributed averaging algorithm:

$$x(t+1) = A(t)x(t),\, t \geq t_0, \tag{5.71}$$

where $t_0 \geq 0$ is the initial time, $x(t) \in \mathbb{R}^N$ is the vector of opinions at each time instant $t \geq t_0$, and chain $\{A(t)\}_{t\geq 0}$, or simply $\{A(t)\}$, is the underlying chain of the network.

The notion of EGC in a network of individuals with discrete time dynamics (5.71) is defined consistently with Definition 5.1. More specifically, for an opinion network with dynamics (5.71), an EGC refers to a subgroup of individuals who are able to lead the whole group to asymptotically agreement on any desired value by cooperatively and properly choosing their own initial opinions, based on an awareness of underlying chain $\{A(t)\}$ as well as the initial opinions of the rest of individuals. Notice that Lemma 5.1, with a similar proof, also holds for a network with dynamics (5.71). In the following, by extending the notions of null space, nullity, and rank to discrete time chains, we exploit the relationship between the characterization of an EGC in a network, size of the smallest EGC, and properties of the underlying chain of the network.

For the sake of notational consistency, let $\Phi(t, \tau)$, $t \geq \tau \geq 0$, be the state transition matrix associated with discrete time chain $\{A(t)\}$. State transition matrix $\Phi(t, \tau)$ satisfies relation (5.2). we also have:

$$\Phi(t, \tau) = A(t-1) \cdots A(\tau),\, \forall t > \tau \geq 0, \tag{5.72}$$

and $\Phi(t, t) = I_{N\times N}$, $\forall t \geq 0$. Define the null space of discrete time chain $\{A(t)\}$ at an arbitrary time instant $\tau \geq 0$, $\mathbf{null}_\tau(A)$, consistently with its continuous time version, i.e., Definition 5.2. $\mathbf{null}_\tau(A)$, $\tau \geq 0$, is again a vector space. However, since the state transition matrix in the discrete time case may be singular at times, unlike the continuous time case, the dimension of $\mathbf{null}_\tau(A)$, denoted by $\dim(\mathbf{null}_\tau(A))$, can vary as τ grows. However, it is not difficult to show that $\dim(\mathbf{null}_\tau(A))$ is non-increasing with respect to τ. We now have the following theorem on the size of the smallest EGC of a network with dynamics (5.71). The proof is eliminated as it is similar to the proof of Theorem 5.1.

Theorem 5.6. *For an opinion network with dynamics (5.71), the size of the smallest EGC is* $N - \dim(\mathbf{null}_{t_0}(A))$.

Since $\dim(\mathbf{null}_\tau(A))$ is non-increasing with respect to τ, from Theorem 5.6, we conclude that initializing the network with dynamics (5.71) at a later time results in a greater or equal size of its smallest EGC. Notice now that $\dim(\mathbf{null}_{t_0}(A))$ is an integer-valued operator bounded below by zero. Thus, $\dim(\mathbf{null}_\tau(A))$ becomes constant after a *finite* time. Define the nullity of chain $\{A(t)\}$, $\mathbf{nullity}(A)$, by that constant:

$$\mathbf{nullity}(A) \triangleq \lim_{\tau\to\infty} \dim(\mathbf{null}_\tau(A)). \tag{5.73}$$

Define now the rank of chain $\{A(t)\}$, $\mathbf{rank}(A)$, as in continuous time, by $\mathbf{rank}(A) = N - \mathrm{null}(A)$. The following corollary, to be viewed as the discrete time counterpart of Theorem 5.1, is an immediate result of Theorem 5.6 and the definition of $\mathbf{rank}(A)$.

Corollary 5.7. *If a network with dynamics (5.71) is initialized at a sufficiently large time, the size of its smallest EGC is* $\mathbf{rank}(A)$*, where a sufficiently large time refers to some time after the RHS of (5.73) has converged.*

In the rest of this section, we focus on the notion of rank of a chain. We recall the definition of l_1-approximation of a discrete time chain from Touri and Nedić (2012a).

Definition 5.12. *Chain* $\{A(t)\}$ *is said to be an* l_1*-approximation of chain* $\{B(t)\}$ *if:*

$$\sum_{t=0}^{\infty} \|A(t) - B(t)\| < \infty, \tag{5.74}$$

where for convenience only, the norm refers to the max norm, i.e., the maximum of the absolute values of the matrix elements.

It can be shown that, rank, as we defined it for the discrete time case, is invariant under an l_1-approximation, i.e., Lemma 5.3 holds for the discrete time case as well.

5.9.1 Rank via Sonin Decomposition-Separation Theorem

We aim to address in this subsection, the rank of a discrete time chain of stochastic matrices via an approach based on the Sonin D-S Theorem Bolouki and Malhamé (2013); Sonin *et al.* (2008). Some preliminaries are required first. According to Blackwell (1945) as reported in Sonin *et al.* (2008), the definition of jet will be recalled. It plays a crucial role in our discrete time arguments.

Definition 5.13. *Given the set of individuals $\mathcal{V} = \{1, \ldots, N\}$, a jet J in $\mathcal{V}$ is a sequence $\{J(t)\}$ of subsets of $\mathcal{V}$. A jet J in $\mathcal{V}$ is called a proper jet if $\emptyset \neq J(t) \subsetneq \mathcal{V}$, $\forall t \geq 0$. Complement of jet $J = \{J(t)\}$ in $\mathcal{V}$, denoted by $\bar{J}$ is also a jet in $\mathcal{V}$ expressed by sequence $\{\mathcal{V} \backslash J(t)\}$. For a fixed subset $S \subset \mathcal{V}$, jet S refers to a jet which is equal to S at all time instants.*

Definition 5.14. *A tuple of jets $(J^1, \ldots, J^c)$ is a jet-partition of $\mathcal{V}$, if $(J^1(t), \ldots, J^c(t))$ forms a partition of $\mathcal{V}$ for every $t \geq 0$.*

Consider a multi-agent system with states evolving according to linear algorithm (5.71). Based on the work Kolmogoroff (1936), we know that discrete time chain $\{A(t)\}$ admits an absolute probability sequence $\{\pi(t)\}$ which propagates backwards in time:

$$\pi'(t+1)A(t) = \pi'(t), \forall t \geq 0. \tag{5.75}$$

From chain $\{A(t)\}$, construct chain $\{P(t)\}$ of stochastic matrices satisfying:

$$\pi_i(t)p_{ij}(t) = \pi_j(t+1)a_{ji}(t), \forall i, j \in \mathcal{V}, \forall t \geq 0. \tag{5.76}$$

More specifically, if $\pi_i(t) \neq 0$, then set:

$$p_{ij}(t) = \pi_j(t+1)a_{ji}(t)/\pi_i(t), \tag{5.77}$$

while if $\pi_i(t) = 0$ for some $i \in \mathcal{V}$ and $t \geq 0$, choose non-negative $p_{ij}(t)$'s arbitrarily such that:

$$\sum_{j=1}^{N} p_{ij}(t) = 1. \tag{5.78}$$

Note that in the former case ($\pi_i(t) \neq 0$), (5.78) is automatically satisfied, implying that $P(t)$ is a stochastic matrix for every $t \geq 0$. It is easy to see that:

$$\pi'(t)P(t) = \pi'(t+1), \forall t \geq 0, \tag{5.79}$$

indicating that $\{\pi(t)\}$ can now be viewed as a non homogeneous *forward* propagating Markov chain.

Definition 5.15. *Let the total flow between two arbitrary jets J^s and J^k in $\mathcal{V}$ over the infinite time interval, denoted by $V(J^s, J^k)$, be defined as:*

$$V(J^s, J^k) \triangleq \sum_{t=0}^{\infty} \left[\sum_{i \in J^k(t)} \sum_{j \in J^s(t+1)} r_{ij}(t) + \sum_{i \in J^s(t)} \sum_{j \in J^k(t+1)} r_{ij}(t) \right], \tag{5.80}$$

where

$$r_{ij}(t) = \pi_i(t) p_{ij}(t) = \pi_j(t+1) a_{ji}(t). \tag{5.81}$$

From a Markov chain point of view, value $r_{ij}(t)$ can be interpreted as the absolute joint probability of being in state i at time t and state j at time $t+1$.

Theorem 5.7. *(Sonin D-S Theorem) There exists an integer c, $1 \leq c \leq N$, and a decomposition of $\mathcal{V}$ into jet-partition $(J^0, J^1, \ldots, J^c)$, $J^k = \{J^k(t)\}$, such that irrespective of the particular time or values at which x_i's are initialized,*

(i) For every k, $1 \leq k \leq c$, there exist real constants π_k^* and x_k^*, such that:

$$\lim_{t \to \infty} \sum_{i \in J^k(t)} \pi_i(t) = \pi_k^*, \tag{5.82}$$

and:

$$\lim_{t \to \infty} x_{i_t}(t) = x_k^*, \tag{5.83}$$

for every sequence $\{i_t\}$, $i_t \in J^k(t)$. Furthermore, $\lim_{t \to \infty} \sum_{i \in J^0(t)} \pi_i(t) = 0$.

(ii) For every distinct k, s, $0 \leq k, s \leq c$: $V(J^k, J^s) < \infty$.

(iii) This decomposition is unique up to jets $\{J(t)\}$ such that for any $\{\pi(t)\}$ we have:

$$\lim_{t \to \infty} \sum_{i \in J(t)} \pi_i(t) = 0, \tag{5.84}$$

and:

$$V(J, \mathcal{V} \backslash J) < \infty. \tag{5.85}$$

Theorem 5.8. *The unique jet decomposition of $\mathcal{V}$ with respect to chain $\{A(t)\}$ in the Sonin D-S Theorem, consists of jet J^0 and* **rank**(A) *other jets.*

Proof. Theorem 5.8 is an immediate result of Remark 4.2 of Chapter 4, combined with Theorem 5.9, that will be stated later in the paper. $\square$

5.9.2 A Geometric Interpretation

We developed, in Section 5.4, a geometric framework, that interprets the rank of the underlying chain of a network, based on the state transition matrix of the network, i.e, $\Phi(t,\tau)$. A similar argument can be made for the discrete time case, with the state transition matrix expressed as (5.72). The only difference here is that c_τ, which is the number of vertices of limiting polytope C_τ, is not invariant as τ grows. As a matter of fact, it can be shown that:

$$c_\tau = N - \dim(\mathcal{N}_\tau(A)). \tag{5.86}$$

Therefore, c_τ is a non-decreasing function of τ and becomes constant after a finite time since it is bounded above by N. In correspondence to Theorem 5.2, we have the following theorem:

Theorem 5.9. *For the number of the vertices of limiting polytope C_τ, $\tau \geq 0$, i.e., c_τ:*

$$\lim_{\tau\to\infty} c_\tau = \mathbf{rank}(A). \tag{5.87}$$

Consequently, there exist $t_0 \geq 0$ such that c_τ is equal to $\mathbf{rank}(A)$ for every $\tau \geq t_0$.

Proof. (5.87) is easily obtained by taking the limit of both sides of (5.86) as $t \to \infty$. □

Similar to the continuous time case, we define ergodicity classes of a discrete time chain as equivalence classes resulted by the relation of being weakly mutually ergodic (see Definition 5.7). It can be shown, similar to the proof of Lemma 5.8, that c_τ for every $\tau \geq 0$ is less than or equal to the number of ergodicity classes (note that ergodicity classes are defined irrespective of the initial time). This, together with Theorem 5.9, implies that the number of ergodicity classes being an upper bound for the rank, i.e., Corollary 5.2, also holds in the discrete time case.

5.9.3 Lower Bounds

We stated, in Theorem 5.3 and Corollary 5.4, lower bounds on the rank of a continuous time chain. The discrete time counterparts of these theorems are subsumed through an approach employing the notion of jets.

Definition 5.16. *For a jet J in $\mathcal{V}$, let $U_{in}(J)$ denote the total influence of $\bar{J}$ on J over the infinite time interval:*

$$U_{in}(J) = \sum_{t=0}^{\infty} \sum_{i\in J(t+1)} \sum_{j\notin J(t)} a_{ij}(t). \tag{5.88}$$

Theorem 5.10. *For a discrete time chain $\{A(t)\}$,* $\mathbf{rank}(A)$ *is greater than or equal to the maximum number of disjoint jets, say J, each of which satisfying:*

$$U_{in}(J) < \infty. \tag{5.89}$$

Proof. The proof of Theorem 5.10 is similar to that of Theorem 5.3. For chain $\{A(t)\}$, let $J^1, \ldots, J^d$ be d disjoint jets. Form a chain $\{B(t)\}$ from chain $\{A(t)\}$ by eliminating all interactions between any two distinct jets among $J^1, \ldots, J^d$ over the infinite time interval. Since $\{B(t)\}$ is an l_1-approximation of $\{A(t)\}$, the two chains share the same rank, as well as the same collections of disjoint jets. Therefore, it is sufficient to prove Theorem 5.10 for chain $\{B(t)\}$. Note that for chain $\{B(t)\}$, for every $s \neq k$, $1 \leq s, k \leq d$, we have:

$$\sum_{t=0}^{\infty} \left[\sum_{i \in J^s(t+1)} \sum_{j \in J^k(t)} b_{ij}(t) + \sum_{i \in J^k(t+1)} \sum_{j \in J^s(t)} b_{ij}(t) \right] = 0. \tag{5.90}$$

We now consider an opinion network with underlying chain $\{B(t)\}$. Keeping Theorem 5.1 in mind, it suffices to show that the size of the smallest EGC of the opinion network defined over chain $\{B(t)\}$ is at least d. Consider a particular EGC of the opinion network defined over chain $\{B(t)\}$. By definition, that particular EGC is able to create global consensus under certain circumstances for infinitely many choices of initial time. Let $t_0 \geq 0$ be one of those infinitely many possible choices of initial time. Relation (5.90) means that for any jet among $J^1, \ldots, J^d$, say J^s, the opinions of individuals in $J^s(t)$, $\forall t \geq t_0$, only depend on the opinion of individuals in $J^s(t_0)$. Therefore, that particular EGC must contain at least one of the individuals in $J^s(t_0)$ or else it would have no control on the opinion of individuals in jet J^s at any future time. Thus, the size of that particular EGC is greater than or equal to d, which is the number of disjoint jets $J^1, \ldots, J^d$. This proves the theorem. $\square$

Theorem 5.10 would serve as the discrete time counterpart of Theorem 5.3, if the choice of jets were limited to the time-invariant jets.

We skip the analysis of time-invariant discrete time chains, since it is no different from its continuous time counterpart.

5.9.4 Rank of Discrete Time Chains in Class $\mathcal{P}^*$

We, first, briefly discuss the limiting behavior of a discrete time chain $\{A(t)\}$ in Class $\mathcal{P}^*$ from two viewpoints: (i) The Sonin D-S theorem; (ii) The geometric viewpoint. Given that $\{A(t)\}$ belongs to Class $\mathcal{P}^*$, there is a representation of Sonin's jet decomposition without a J^0 jet. Therefore, each individual lies within $\cup_{k=1}^{c} J^k(t)$ for any $t \geq 0$, with c being equal to

$\mathbf{rank}(A)$. Thus, the opinion of each individual stays arbitrarily close to set $\{x_k^* \,|\, 1 \leq k \leq c\}$, with size $\mathbf{rank}(A)$, as t grows large. Considering now the geometric viewpoint, we focus on limiting polytopes C_τ as discussed in Section 5.9.2. For the discrete time case, it was pointed out that the number of vertices of C_τ is non-decreasing and becomes constant past a finite time $t_0 \geq 0$, with $\mathbf{rank}(A)$ being that constant. As proved in Chapter 4, if $\{A(t)\}$ is in Class $\mathcal{P}^*$, for every arbitrary fixed $\tau \geq t_0$, every column of $\Phi'(t,\tau)$ stays arbitrarily close to the $\mathbf{rank}(A)$ vertices of C_τ as t grows large. Since $x(t) = \Phi(t,\tau)x(\tau)$, each column i of $\Phi'(t,\tau)$ (row i of $\Phi(t,\tau)$) is in correspondence with the opinion of an individual i. Thus, columns of $\Phi'(t,\tau)$ staying arbitrary close to the $\mathbf{rank}(A)$ vertices of C_τ as $t \to \infty$, leads to the same conclusion from the other point of view, that is the opinions staying arbitrary close to a set of $\mathbf{rank}(A)$ (generally distinct) values. Thus, to sum up, although convergence of each individual's opinion is not guaranteed here unlike the continuous time case, there is a *finite number* of accumulation points for the opinions over the infinite time interval, and that finite number is $\mathbf{rank}(A)$.

Now reconsider jet-partition $(J^1, \ldots, J^c)$ in the Class $\mathcal{P}^*$ based jet-decomposition of the Sonin D-S Theorem. According to the Sonin D-S Theorem, for every two jets J^k and J^s, we have:

$$V(J^k, J^s) < \infty. \tag{5.91}$$

Recalling (5.81) and taking into account that $\pi_j(t+1)$ in (5.81) is greater than or equal to some $p^* > 0$ since chain $\{A(t)\}$ has been assumed to be in Class $\mathcal{P}^*$, inequality (5.91) implies that the total interaction between any two jets J^k and J^s is finite over the infinite time interval, i.e.,

$$\sum_{t=0}^{\infty} \left[\sum_{i \in J^{k+1}(t)} \sum_{j \in J^s(t)} a_{ij}(t) + \sum_{i \in J^{s+1}(t)} \sum_{j \in J^k(t)} a_{ij}(t) \right] < \infty. \tag{5.92}$$

Fix an arbitrary k, and consider the set of inequalities obtained as $s \neq k$ goes from 1 to c in (5.92). Adding the $c-1$ obtained inequalities of type (5.92), and noting that $J^1, \ldots, J^c$ is a jet partition of $\mathcal{V}$, we conclude that the total interaction between J^k and $\bar{J}^k$, and in particular the total influence of $\bar{J}^k$ over J^k, is also finite over the infinite time interval. Therefore, for each of disjoint jets $J^1, \ldots, J^c$, say J^k, $V_{in}(J^k) < \infty$ (see (5.89)). Thus, recalling $\mathbf{rank}(A) = c$, we conclude that the lower bound provided in Theorem 5.10 is achieved for discrete time chains in Class $\mathcal{P}^*$.

5.9.5 Full-Rank Chains

One characterizes full-rank discrete time chains according to the following theorem.

Theorem 5.11. *A discrete time chain $\{A(t)\}$ is full-rank, i.e.,* $\mathbf{rank}(A) = N$ *if and only if $\{A(t)\}$ is an l_1-approximation of a permutation chain, i.e., a chain of permutation matrices.*

Proof. The proof of Theorem 5.11, which is the discrete time version of Theorem 5.5, is omitted since the proofs of the two theorems are very similar. □

5.10 Conclusion

We considered a network of multiple individuals with opinions updated via a general time-varying continuous or discrete time linear algorithm. The notion of EGC, an acronym associated with Éminence Grise Coalition, in the network was defined as follows. Given the time that network starts to update, an EGC is a subgroup of individuals who, cooperatively, can manage to create a global consensus on any desired opinion in the network only by adequately setting their initial opinions assuming that they are aware of the underlying chain of the network as well as the rest of individuals initial opinions. The size of the smallest EGC can be treated as a characteristic of the underlying update chain of the network. We then introduced an extension of the notion of rank, from an individual matrix related notion to one related to a Markov chain in continuous or discrete time. A key result is that the rank of the underlying chain of a network is also the size of its smallest EGC in the continuous time case. The same holds in the discrete time case provided the initial time is "sufficiently large" in a sense made precise in the paper. Geometrically, and associated with the chain, one can define a monotone decreasing convex hulls (polytopes) generated by an underlying sequence of vertices. The rank of the chain is the limiting number of linearly independent vertices in the sequence of polytopes, which is reached in finite time.

The continuous time case is peculiar in the sense that the rank (number of linearly independent vertices) of the elements of the polytopic sequence remains constant, while it is monotonically increasing in the discrete time case. This, in turn, makes consensus behavior somewhat simpler in continuous time than in discrete time. A collection of upper and lower bounds on the rank was also established. These two bounds are shown to be equal to the rank for both time invariant chains (possibly not in Class $\mathcal{P}^*$), as well as for Class $\mathcal{P}^*$ chains in the time inhomogeneous case.

From a practical standpoint, this work establishes the rather intuitive result that the less "natural" dissension exists in an opinion network, the easier it is to steer the network towards global consensus. In cases where an "average" amount of natural dissonance exists, then the theory points at the need to minimally "infiltrate" identifiable dissenting clusters and work from the inside so to speak to steer the global opinion to a consensus. Success in doing so hinges on an ability to enlist key agents cooperation given that they must act as a “grand coalition” of key agents. This in turn opens the door to games over opinion networks whereby key agents might choose to break up into smaller coalitions and work towards conflicting goals. This will be the subject of future research. Another direction for future research is that of developing simple algorithms to identify key agents in the opinion network. Finally, a question of mathematical interest is the following:

Given an arbitrary non-ergodic time-varying chain, what is the sparsest time-invariant chain such that sum of the two chains becomes ergodic? There seems to be a relationship between the sparsity index of the corresponding graph of the sparsest time-invariant chain and the rank of the time-varying chain.

CHAPTER 6

GENERAL DISCUSSION

Consensus problems for distributed time-varying averaging algorithms have gained increasing attention in various research communities. One of the fundamental problems related to consensus is the *unconditional* occurrence of consensus or multiple consensus via distributed time-varying averaging algorithms, where by "unconditional", we mean irrespective of time or values at which states are initialized. Such problem turn out to be equivalent to ergodicity or class-ergodicity of the underlying chain of the system (see Bolouki and Malhamé (2013); Chatterjee and Seneta (1977); Touri and Nedić (2012b)). Discovering necessary and/or sufficient conditions for ergodicity and class-ergodicity of a time-varying chain of matrices has been the aim of a significant body of literature in the past decade (see Blondel *et al.* (2005); Hendrickx and Blondel (2006); Hendrickx and Tsitsiklis (2013); Jadbabaie *et al.* (2003); Li *et al.* (2004); Lorenz (2005); Moreau (2005); Touri and Nedić (2011, 2012a,b, 2014)). One of our two main objectives in this thesis has been to extend, as far as possible, the existing results regarding the proposed problem. Our contributions to this problem can be summarized as the following.

Balanced asymmetric chains. Balanced asymmetry is a property of a chain of stochastic matrices defined in discrete time. Balanced asymmetry is a hybrid of notions of subsymmetry Bolouki and Malhamé (2011a), and cut-balance Hendrickx and Tsitsiklis (2011) which were already very much discussed in the literature, and which are essentially point-wise verifiable properties of the underlying chain. We found, for balanced asymmetric chains, necessary and sufficient conditions for ergodicity and class-ergodicity based on a dynamic notion proposed by Touri and Nedić, that of absolute infinite flow, which is a property that can be verified only when considering the chain as a whole. The notion of balanced asymmetry, on its own, helped us subsume and generalize virtually all known convergence results thus far, albeit not convergence rate issues which is thoroughly a different concern.

Applications to known models. We showed that our techniques, which are employed to derive the convergence results on the exogenous averaging algorithms, together with the results themselves, can also be applied to some well-known nonlinear models, such as the Cucker-Smale model Cucker and Smale (2007) and the Hegselmann-Krause model Hegselmann and Krause (2002). These nonlinear models can be viewed as endogenous averaging algorithms, i.e., averaging algorithms with coefficients dynamically changing according to the evolution of states in the network.

Connection to Sonin's Decomposition-Separation Theorem. Our basic conviction that the theory of inhomogeneous Markov chains could help understand the convergence properties of consensus algorithms, which essentially depended on the properties of the underlying chain of the system, led us to employ the Sonin's Decomposition-Separation Theorem Sonin *et al.* (2008). The D-S Theorem together with the intuitions of Touri and Nedić (2014) about the importance of Kolmogorov's notion of absolute probability sequence, helped us obtain a meaningful generalization of the notion of absolute infinite flow to so-called *infinite jet-flow*.

A geometric framework. Attempts to understand the convergence mechanisms of inhomogeneous Markov chains led us to our first geometric insights of the Markov chain convergence as the intersection of decreasing convex hulls of appropriate sets of vertices. The vertices of each set correspond to the rows of the state transition matrix of the system at a certain time. This geometric interpretation was employed to extend our theorems, obtained based on the D-S Theorem, to the continuous time case.

Centered on these geometric insights, we then explored a question which is often raised for endogenously evolving consensus algorithms, such as the celebrated Cucker-Smale model Cucker and Smale (2007): Are there particular sets of initial conditions which will guarantee that the resulting consensus algorithm will converge unconditionally? Instead, the question is raised here for an exogenously generated sequence of update matrices. The geometric insights and the proposed question led us to defining the following notions and addressing their related issues which has been our second main objective in this thesis:

Éminence Grise Coalitions. It turns out that there exists a minimal subset of agents, which by mere setting of their initial conditions (under the rather idealized condition that they know where everyone else stands initially and the evolution of the network update chain), can steer the complete set of agents towards a global consensus. Such agents may be viewed as the consummate negotiators in a polarized environment, and we believe that such results are important for the study of opinion dynamics. A subset of agents with that property, even if it is not minimal, is called an *éminence grise coalition*, or simply EGC, in this thesis. We extensively investigated the size of the minimal EGC in a system.

Rank of a chain. We extended the notion of rank, as defined for a matrix, to a chain of matrices in both continuous and discrete time. We proved that the rank of the underlying chain of a multi-agent system is equal to the size of the minimal EGC that the system admits.

CHAPTER 7

CONCLUSION

In this thesis, we have proposed novel approaches towards analysis of the limiting behavior of the state vector in a network of multiple agents with dynamics updated via a predefined distributed averaging algorithm. To determine the asymptotic behavior of agents in such a network, three items must be taken into account:

(i) the initial time at which the agents start to update their states.

(ii) the initial conditions, i.e., the state values at the initial time.

(iii) the predefined (exogenously given) update algorithm.

Our first interest has been to identify the largest class of discrete and continuous time update algorithms, for which global consensus or multiple consensus within the network is achieved "unconditionally". Unconditional convergence refers to convergence irrespective of the values or time at which states of agents are initialized (items (i) and (ii)). Since each distributed averaging algorithm is uniquely defined by a chain of row-stochastic matrices (the so-called underlying chain of the network), guaranteeing the occurrence of unconditional global or multiple consensus can be considered as a property of the underlying chain. More specifically, (class-) ergodicity of the underlying chain is equivalent to the occurrence of global (multiple) consensus in the network (see Bolouki and Malhamé (2013); Chatterjee and Seneta (1977); Touri and Nedić (2012b)).

In the first attempt to characterize the largest class of discrete time (class-) ergodic chains, using only elementary methods, and by developing the notion of "balanced asymmetry" in chains of stochastic matrices, we rediscovered and, although not significantly, generalized the previous classes of (class-) ergodic chains in the literature (most notably Hendrickx and Tsitsiklis (2013); Touri and Nedić (2014)).

We then exploited the relationship between ergodicity (class-) ergodicity of chains of stochastic matrices and the Sonin's Decomposition-Separation Theorem Sonin *et al.* (2008). The D-S Theorem and the work Kolmogoroff (1936) on the existence of an absolute probability sequence for an arbitrary Markov chain, helped us define the notion of "infinite jet-flow", as a generalization of (absolute) infinite flow (see Touri and Nedić (2011, 2012a)), and obtain a larger class of (class-) ergodic chains.

Developing a geometric framework, we verified (class-) ergodicity of the obtained class of chains. This geometric framework, which is based on the associated state transition matrix of the network, helped us extend our previous results to the continuous time case. En route to find the the correct extension, we again employed Kolmogorov's work Kolmogoroff (1936) that proves the existence of an absolute probability sequence for both discrete and continuous time Markov chains.

One can think of our second interest in this thesis as the complement of our first interest. We assume now that, items (i) and (ii) above are to be tuned with the goal of achieving consensus, while there is absolutely no assumption on the update algorithm, i.e., item (iii). More precisely, given a chain of stochastic matrices as the underlying chain of the network, starting at an arbitrary initial time t_0, what is the set of initial conditions resulting in global consensus? This set proves to be a vector space with a dimension that is obtained by our geometric approach and (in the discrete time case) the D-S Theorem. The set of initial conditions resulting in global consensus led us to define the notion of "éminence grise coalition" (EGC) which resembles as a group of hidden leaders within the network. Using our previous arguments and results in (class-) ergodicity analysis of chains, the geometric framework in particular, we have addressed the size of the smallest EGC of a network.

REFERENCES

ACEMOGLU, D., NEDIC, A. and OZDAGLAR, A. (2008). Convergence of rule-of-thumb learning rules in social networks. *Proceedings of 47th IEEE Conference on Decision and Control (CDC 2008)*. 1714–1720.

AEYELS, D. and ROGGE, J. A. (2004). Existence of partial entrainment and stability of phase locking behavior of coupled oscillators. *Progress of Theoretical Physics*, 112, 921–942.

AYSAL, T. C. and BARNER, K. E. (2010). Convergence of consensus models with stochastic disturbances. *Information Theory, IEEE Transactions on*, 56, 4101–4113.

AYSAL, T. C., COATES, M. J. and RABBAT, M. G. (2008). Distributed average consensus with dithered quantization. *IEEE Transactions on Signal Processing*, 56, 4905–4918.

BERTSEKAS, D. P. and TSITSIKLIS, J. N. (1989). *Parallel and distributed computation: numerical methods*. Prentice-Hall, Inc.

BERTSEKAS, D. P. and TSITSIKLIS, J. N. (2007). Comments on "coordination of groups of mobile autonomous agents using nearest neighbor rules". *IEEE Transactions on Automatic Control*, 52, 968.

BLACKWELL, D. (1945). Finite non-homogeneous chains. *Annals of Mathematics*, 594–599.

BLONDEL, V., HENDRICKX, J., OLSHEVSKY, A. and TSITSIKLIS, J. (2005). Convergence in multiagent coordination, consensus, and flocking. *Proceedings of 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 2005)*. 2996–3000.

BLONDEL, V. D., HENDRICKX, J. M. and TSITSIKLIS, J. N. (2007). On the 2r conjecture for multi-agent systems. *Proceedings of the European Control Conference (ECC 2007)*. 874–881.

BLONDEL, V. D., HENDRICKX, J. M. and TSITSIKLIS, J. N. (2009). On krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transactions on Automatic Control*, 54, 2586–2597.

BOLOUKI, S. and MALHAMÉ, R. P. (2011a). On consensus with a general discrete time convex combination based algorithm for multi-agent systems. *Proceedings of 19th Mediterranean Conference on Control & Automation (MED 2011)*. 668–673.

BOLOUKI, S. and MALHAMÉ, R. P. (2011b). On the limiting behavior of linear or convex combination based updates of multi-agent systems. *Proceedings of 18th IFAC World Congress*. 8819–8823.

BOLOUKI, S. and MALHAMÉ, R. P. (2012a). Ergodicity and class-ergodicity of balanced asymmetric stochastic chains. *GERAD Technical Report G-2012-93.*

BOLOUKI, S. and MALHAMÉ, R. P. (2012b). Theorems about ergodicity and class-ergodicity of chains with applications in known consensus models. *Proceedings of 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton).* 1425–1431.

BOLOUKI, S. and MALHAMÉ, R. P. (2013). Consensus algorithms and the decomposition-separation theorem. *Proceedings of 52th IEEE Conference on Decision and Control Conference (CDC 2013).* 1490–1495.

BONDY, J. A. and MURTY, U. S. R. (1976). *Graph theory with applications*, vol. 6. Macmillan London.

BOYD, S., GHOSH, A., PRABHAKAR, B. and SHAH, D. (2006). Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52, 2508–2530.

BROCKETT, R. W. (1970). *Finite dimensional linear systems.* New York Wiley.

CARLI, R., FAGNANI, F., FRASCA, P., TAYLOR, T. and ZAMPIERI, S. (2007). Average consensus on networks with transmission noise or quantization. *Proceedings of European Control Conference (ECC 2007).* 1852–1857.

CARLI, R., FAGNANI, F., SPERANZON, A. and ZAMPIERI, S. (2008). Communication constraints in the average consensus problem. *Automatica*, 44, 671–684.

CHATTERJEE, S. and SENETA, E. (1977). Towards consensus: some convergence theorems on repeated averaging. *Journal of Applied Probability*, 89–97.

COUZIN, I. D., KRAUSE, J., FRANKS, N. R. and LEVIN, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature*, 433, 513–516.

CUCKER, F. and MORDECKI, E. (2008). Flocking in noisy environments. *Journal de mathématiques pures et appliquées*, 89, 278–296.

CUCKER, F. and SMALE, S. (2007). Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52, 852–862.

CUCKER, F., SMALE, S. and ZHOU, D.-X. (2004). Modeling language evolution. *Foundations of Computational Mathematics*, 4, 315–343.

DEFFUANT, G., NEAU, D., AMBLARD, F. and WEISBUCH, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3, 87–98.

DEGROOT, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69, 118–121.

DIMAKIS, A. G., KAR, S., MOURA, J. M., RABBAT, M. G. and SCAGLIONE, A. (2010). Gossip algorithms for distributed signal processing. *Proceedings of the IEEE*, 98, 1847–1864.

DOOB, J. L. (1953). *Stochastic processes*, vol. 101. New York Wiley.

ERMENTROUT, G. B. (1992). Stable periodic solutions to discrete and continuum arrays of weakly coupled nonlinear oscillators. *SIAM Journal on Applied Mathematics*, 52, 1665–1687.

FAX, J. A. and MURRAY, R. M. (2004). Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49, 1465–1476.

FLIERL, G., GRÜNBAUM, D., LEVINS, S. and OLSON, D. (1999). From individuals to aggregations: the interplay between behavior and physics. *Journal of Theoretical Biology*, 196, 397–454.

GRAVER, J., SERVATIUS, H. and SERVATIUS, B. (1984). *Chemical Oscillators, Waves, and Turbulance*. Berlin, Germany: Springer-Verlag.

HEGSELMANN, R. and KRAUSE, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5.

HENDRICKX, J. M. (2008). *Graphs and networks for the analysis of autonomous agent systems*. Thèse de doctorat, Ecole Polytechnique.

HENDRICKX, J. M. and BLONDEL, V. (2006). Convergence of different linear and non-linear vicsek models. *Proceedings of 17th International Symposium on Mathematical Theory of Networks and Systems (MTNS2006)*. 1229–1240.

HENDRICKX, J. M. and TSITSIKLIS, J. N. (2011). A new condition for convergence in continuous-time consensus seeking systems. *Proceedings of 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC 2011)*. 5070–5075.

HENDRICKX, J. M. and TSITSIKLIS, J. N. (2013). Convergence of type-symmetric and cut-balanced consensus seeking systems. *IEEE Transactions on Automatic Control*, 58, 214–218.

HUANG, M. and MANTON, J. H. (2007). Stochastic approximation for consensus seeking: Mean square and almost sure convergence. *Proceedings of 46th IEEE Conference on Decision and Control (CDC 2007)*. 306–311.

HUANG, M. and MANTON, J. H. (2009). Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior. *SIAM Journal on Control and Optimization*, 48, 134–161.

JADBABAIE, A., LIN, J. and MORSE, A. S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48, 988–1001.

JADBABAIE, A., MOTEE, N. and BARAHONA, M. (2004). On the stability of the kuramoto model of coupled nonlinear oscillators. *Proceedings of American Control Conference (ACC 2004).* vol. 5, 4296–4301.

KARP, R., SCHINDELHAUER, C., SHENKER, S. and VOCKING, B. (2000). Randomized rumor spreading. *Proceedings of 41st Annual Symposium on Foundations of Computer Science, 2000*. 565–574.

KASHYAP, A., BAŞAR, T. and SRIKANT, R. (2007). Quantized consensus. *Automatica*, 43, 1192–1203.

KEMPE, D. and KLEINBERG, J. (2002). Protocols and impossibility results for gossip-based communication mechanisms. *Proceedings of 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002*. 471–480.

KOLMOGOROFF, A. (1936). Zur theorie der markoffschen ketten. *Mathematische Annalen*, 112, 155–160.

KRAUSE, U. (1997). Soziale dynamiken mit vielen interakteuren. eine problemskizze. *Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren*, 37–51.

KRAUSE, U. (2000). A discrete nonlinear and non-autonomous model of consensus formation. *Communications in difference equations*, 227–236.

LI, S., WANG, H. and WANG, M. (2004). Multi-agent coordination using nearest neighbor rules: a revisit to vicsek model. *arXiv preprint cs/0407021.*

LIN, P. and JIA, Y. (2009a). Consensus of second-order discrete-time multi-agent systems with nonuniform time-delays and dynamically changing topologies. *Automatica*, 45, 2154–2158.

LIN, P. and JIA, Y. (2009b). Further results on decentralised coordination in networks of agents with second-order dynamics. *Control Theory & Applications, IET*, 3, 957–970.

LIN, Z., BROUCKE, M. and FRANCIS, B. (2004). Local control strategies for groups of mobile autonomous agents. *IEEE Transactions on Automatic Control*, 49, 622–629.

LORENZ, J. (2005). A stabilization theorem for dynamics of continuous opinions. *Physica A: Statistical Mechanics and its Applications*, 355, 217–223.

LORENZ, J. (2007). Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18, 1819–1838.

LYNCH, N. A. (1996). *Distributed algorithms.* Morgan Kaufmann.

MOREAU, L. (2003). Leaderless coordination via bidirectional and unidirectional time-dependent communication. *Proceedings of 42nd IEEE Conference on Decision and Control (CDC 2003).* vol. 3, 3070–3075.

MOREAU, L. (2004). Stability of continuous-time distributed consensus algorithms. *Proceedings of 43rd IEEE Conference on Decision and Control (CDC 2004)*. vol. 4, 3998–4003.

MOREAU, L. (2005). Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50, 169–182.

MOREAU, L., BACHMAYER, R. and LEONARD, N. E. (2003). Coordinated gradient descent: A case study of lagrangian dynamics with projected gradient information. *IFAC Lagrangian and Hamiltonian Methods in Nonlinear Control*, 57–62.

NEDIC, A., OLSHEVSKY, A., OZDAGLAR, A. and TSITSIKLIS, J. N. (2009). On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54, 2506–2517.

OLFATI-SABER, R. and MURRAY, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49, 1520–1533.

REN, W. and BEARD, R. W. (2004). Consensus of information under dynamically changing interaction topologies. *Proceedings of American Control Conference (ACC 2004)*. vol. 6, 4939–4944.

REN, W., BEARD, R. W. and MCLAIN, T. W. (2005a). Coordination variables and consensus building in multiple vehicle systems. *Cooperative Control*, Springer. 171–188.

REN, W., BEARD, R. W. *ET AL.* (2005b). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50, 655–661.

SONIN, I. M. *ET AL.* (2008). The decomposition-separation theorem for finite nonhomogeneous markov chains and related problems. *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*, Institute of Mathematical Statistics. 1–15.

STROGATZ, S. H. (2000). From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143, 1–20.

STROGATZ, S. H. (2001). Exploring complex networks. *Nature*, 410, 268–276.

SUN, Y. G. and WANG, L. (2009a). Consensus of multi-agent systems in directed networks with nonuniform time-varying delays. *IEEE Transactions on Automatic Control*, 54, 1607–1613.

SUN, Y. G. and WANG, L. (2009b). Consensus problems in networks of agents with double-integrator dynamics and time-varying delays. *International Journal of Control*, 82, 1937–1945.

TOURI, B. and NEDIĆ, A. (2011). On ergodicity, infinite flow, and consensus in random models. *IEEE Transactions on Automatic Control*, 56, 1593–1605.

TOURI, B. and NEDIĆ, A. (2012a). On approximations and ergodicity classes in random chains. *IEEE Transactions on Automatic Control*, 57, 2718–2730.

TOURI, B. and NEDIĆ, A. (2012b). On backward product of stochastic matrices. *Automatica*, 48, 1477–1488.

TOURI, B. and NEDIĆ, A. (2014). Product of random stochastic matrices. *IEEE Transactions on Automatic Control*, 59, 437–448.

TSITSIKLIS, J. N. (1984). Problems in decentralized decision making and computation. Rapport technique, DTIC Document.

TSITSIKLIS, J. N., BERTSEKAS, D. P., ATHANS, M. *ET AL.* (1986). Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE transactions on automatic control*, 31, 803–812.

VICSEK, T., CZIRÓK, A., BEN-JACOB, E., COHEN, I. and SHOCHET, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75, 1226–1229.

WEISBUCH, G., DEFFUANT, G., AMBLARD, F. and NADAL, J.-P. (2002). Meet, discuss, and segregate! *Complexity*, 7, 55–63.

WOLFOWITZ, J. (1963). Products of indecomposable, aperiodic, stochastic matrices. *Proceedings of the American Mathematical Society*, 14, 733–737.

APPENDIX A

APPLICATION TO SOME KNOWN MODELS

In this chapter, we apply our main theorems of Chapter 3 to chains corresponding to different types of models and consensus algorithms found in the literature in order to analyze when their transition chains become ergodic or class-ergodic.

A.1 JLM Model

The parameter considered in the JLM model Jadbabaie *et al.* (2003) is the heading of each agent. If we write $\theta_i(t)$ as the heading of agent i at moment t, the model describing evolution of headings is

$$\theta_i(t+1) = \frac{1}{1+d_i(t)}[\theta_i(t) + \sum_{j\in D_i(t)} \theta_j(t)], \tag{A.1}$$

where $D_i(t)$ and $d_i(t)$ denote respectively the set and the number of neighbors of agent i at time t. It is also assumed that for each $t \geq 0$, if $j \in D_i(t)$, then $i \in D_j(t)$ too (undirected communication graph). Assuming that $\{A(t)\}$ is the transition chain of the JLM model, one concludes:

$$a_{ij}(t) = \begin{cases} 1/(d_i(t)+1) & \text{if } j \in D_i(t) \text{ or } j = i, \\ 0 & \text{else.} \end{cases} \tag{A.2}$$

In Jadbabaie *et al.* (2003), the authors proved that a sufficient condition for consensus to occur is existence of an infinite sequence of contiguous, nonempty, bounded, time-intervals $[t_i, t_{i+1})$, $i = 0, 1, \ldots$, such that across each such interval, the N agents are linked together.

In the following, we wish to apply Theorems 3.2 and 3.4 to the transition chain of the JLM model. To take advantage of Theorem 3.4, we show that in the JLM model, the transition chain is both self-confident and cut-balanced. Note that:

$$a_{ii}(t) = 1/(1+d_i(t)) \geq 1/N. \tag{A.3}$$

This proves self-confidence of the chain. To prove cut-balancedness, it suffices to show that the chain is subsymmetric. If at time $t \geq 0$, we have $j \notin D_i(t)$, then $i \notin D_j(t)$ either. Therefore $a_{ij}(t) = a_{ji}(t) = 0$ (consistent with the subsymmetry requirement). If $j \in D_i(t)$

then $i \in D_j(t)$ also. In this case, since

$$1 \leq d_i \leq N-1 \text{ and } 1 \leq d_j \leq N-1, \tag{A.4}$$

it is easy to conclude from (A.1) that $a_{ij}(t)$ and $a_{ji}(t)$ both lie in the interval $[1/2, 1/N]$. Therefore, the subsymmetry condition holds by setting $M = N/2$. Thus, from Theorem 3.4, we conclude that the chain is class-ergodic. In other words, in the JLM model, unconditional multiple consensus occurs without any additional assumption.

We also note that the chain is balanced asymmetric as well, since self-confidence and cut-balancedness imply balanced asymmetry. Thus, from Theorem 3.2, we obtain that the absolute infinite flow property of the transition chain is necessary and sufficient for the ergodicity of the chain. On the other hand, since the chain is self-confident, the absolute infinite flow property is equivalent to the infinite flow property. Hence, in the JLM model, the infinite flow property is necessary and sufficient for ergodicity of the transition chain.

Graph interpretation of the infinite flow property in the JLM model is as follows. Due to the subsymmetry property of the JLM model, if agent i is influence by agent j infinitely many times, then j is also influence by i infinitely many times. Therefore, in this case, the strong interactions graph can be considered as an undirected graph. The infinite flow property is now equivalent to connectivity of the strong interactions graph.

Another equivalent condition to the infinite flow property, that is more similar to the condition derived in Jadbabaie *et al.* (2003), is existence of an infinite sequence of contiguous and non empty time-intervals $[t_i, t_{i+1})$, $i \geq 0$, with the property that across each such interval, the N agents are linked together. Note that the boundedness of the time-intervals is not required unlike the argument of Jadbabaie *et al.* (2003). More importantly, the condition derived here is not only sufficient, but also necessary for ergodicity of the chain (unconditional occurrence of consensus in the model). On the other hand, unlike Jadbabaie *et al.* (2003), without extra conditions, no statement can be made about the speed of convergence to consensus.

A.2 Models with Finite Range Interactions

The HK model Hegselmann and Krause (2002) is an example of endogenous models with finite range interactions. These models are special cases of first order models in which interaction rates depend directly on states. In these models, agent i receives information from agent j if and only if the distance between the two agents is less than some pre-specified level R_i, which is in general different for distinct agents. We define in the following, the interaction rates between agents. For every agent i, we set a decaying function $f_i : R^{\geq 0} \to R^{\geq 0}$ that vanishes at R_i, and define:

$$a_{ij} = \frac{f_i(\|x_i - x_j\|)}{\sum_{k=1}^{N} f_i(|x_i - x_k|)}. \tag{A.5}$$

Let us consider a particular case of models defined above. Assume that agents have the same range of connectivity, i.e., $R_i = R$ for every i, and use identical decaying functions, i.e., $f_i = f$. In the Krause model, $f(x) = 1$ for $0 \leq x < R$ and $f(x) = 0$ elsewhere.

It can be proved that in this case, the transition chain is self-confident with $\delta = 1/N$. It can also be shown as follows that the transition chain is subsymmetric. If at time $t \geq 0$, agents i and j do not communicate, then $a_{ij}(t) = a_{ji}(t) = 0$. If the two agents communicate, then $f(|x_i(t) - x_j(t)|) > 0$. Using (A.5), we have:

$$\frac{a_{ij}(t)}{a_{ji}(t)} = \frac{\sum_{k=1}^{N} f(|x_j(t) - x_k(t)|)}{\sum_{k=1}^{N} f(|x_i(t) - x_k(t)|)}. \tag{A.6}$$

Noting that f is non-increasing and $f(|x_i(t) - x_i(t)|) = f(|x_j(t) - x_j(t)|) = f(0)$, we conclude that the RHS of (A.6) lies in interval $[1/N, N]$. Hence, subsymmetry is established by setting $M = N^2$. The chain being both self-confident and subsymmetric, it is also cut-balanced. Thus, according to Theorem 3.4, the chain is class-ergodic, i.e., unconditional multiple consensus occurs.

A.3 The C-S model

The C-S (Cucker-Smale) model Cucker and Smale (2007) is an example of endogenous consensus models with interaction rates remaining strictly positive. We apply our results to a generalized version of the C-S model Cucker and Smale (2007) that describes evolution of positions x_i's and velocities v_i's in a bird flock, in a three dimensional Euclidian space:

$$\begin{cases} x_i(t+1) &= x_i(t) + hv_i(t), \\ v_i(t+1) &= v_i(t) + \sum_{j \neq i} f(\|x_i(t) - x_j(t)\|)(v_j(t) - v_i(t)), \end{cases} \tag{A.7}$$

where $f : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ is a non increasing function. Note that in this model, the limiting behavior of velocities is of interest. We have:

$$a_{ij}(t) = f(\|x_i(t) - x_j(t)\|), \ \forall i \neq j, \tag{A.8}$$

and:

$$a_{ii}(t) = 1 - \sum_{j \neq i} f(\|x_i(t) - x_j(t)\|), \ \forall i. \tag{A.9}$$

Clearly, the transition chain in this algorithm is symmetric. To establish self-confident, we assume that $f(x) < 1/N$ for any $x \geq 0$. By this assumption, we have $a_{ii}(t) > 1/N$ for every $i = 1, \ldots, N$ and $t \geq 0$. Therefore, the chain becomes self-confident. The combination of self-confidence and cut-balancedness of the chain allows an application of Theorem 3.4 to yield the following result.

Theorem A.1. *Consider the system with dynamics described by (A.7). If $f(x) < 1/N$ for any $x \geq 0$, then the transition chain is class-ergodic, and consequently, unconditional multiple consensus occurs.*

To state the consensus result for the generalized C-S model, we define parameters M_x and M_v calculated from initial positions and velocities:

$$M_x = \max_{i,j}\{\|x_i(0) - x_j(0)\| | 1 \leq i < j \leq N\}, \tag{A.10}$$

$$M_v = \max_{i,j}\{\|v_i(0) - v_j(0)\| | 1 \leq i < j \leq N\}. \tag{A.11}$$

Theorem A.2. *For the multi-agent system with dynamics described by Eq. (A.7), assume that $f(x)$ has the following property:*

$$f(x) < 1/N, \ \forall x \geq 0. \tag{A.12}$$

Assume also that

$$M_v < \frac{N}{3h} \int_{M_x}^{\infty} f(y) dy. \tag{A.13}$$

Then, all agents' velocities converge to a common value. Moreover, there exists a non negative number R such that for every i, j, $1 \leq i, j \leq N$,

$$\|x_i(t) - x_j(t)\| \leq R, \ \forall t \geq 0 \tag{A.14}$$

Unlike the models described previously, Theorem A.2 is not an immediate result of Theorems 3.2 and 3.4. However, to prove Theorem A.2, we employ a technique similar to that used in the proof of Theorem 3.2 in Bolouki and Malhamé (2012b).

Proof. For every $i = 1, \ldots, N$, let $v_{i1}(t), v_{i2}(t), v_{i3}(t)$ be components of $v_i(t)$, i.e.,

$$v_i(t) = [v_{i1}(t) \ v_{i2}(t) \ v_{i3}(t)]'. \tag{A.15}$$

It is straight forward to verify that v_{ir}'s ($r = 1, 2, 3$) satisfy for straightforwardly identifiable coefficients a_{ij}, the same update equation as v_i's do, i.e.,

$$v_{ir}(t+1) - v_{ir}(t) = \sum_{j \neq i} a_{ij}(t)(v_{jr}(t) - v_{ir}(t)). \tag{A.16}$$

One can rewrite (A.16) as the following:

$$v_{ir}(t+1) = \left(1 - \sum_{j \neq i} a_{ij}(t)\right) v_{ir}(t) + \sum_{j \neq i} a_{ij}(t) v_{jr}(t). \tag{A.17}$$

Let us define $z_{ir}(t) : \mathbb{R}^{\geq 0} \to \mathbb{R}$, $1 \leq i \leq N$, $r = 1, 2, 3$ from $v_{ir}(t)$'s as follows. At every time $t \geq 0$, $z_{ir}(t)$ is equal to the ith least number among $v_{1r}(t), \ldots, v_{Nr}(t)$.

Note that the coefficients in the RHS of (A.17) are all positive and add up to 1. This means that $v_{ir}(t+1)$ is a convex combination of the values $v_{1r}(t), \ldots, v_{Nr}(t)$. Thus, for every r, $1 \leq r \leq 3$, interval $[z_{1r}(t), z_{Nr}(t)]$, which is the smallest interval containing all the values $v_{ir}(t)$'s, shrinks during time. Particularly, this shows that $v_{ir}(t)$, and consequently $z_{ir}(t)$, are bounded for every $i = 1, \ldots, N$ and $r = 1, 2, 3$. For consensus to occur, we require that interval $[z_{1r}(t), z_{Nr}(t)]$ converges to a point for every r, $1 \leq r \leq 3$. In the following, we investigate how these intervals shrink with time. Let us define:

$$z(t) = \sum_{r=1}^{3} (z_{Nr}(t) - z_{1r}(t)). \tag{A.18}$$

We know that:

$$x_i(t) - x_j(t) = (x_i(0) - x_j(0)) + h \sum_{\tau=0}^{t-1} (v_i(\tau) - v_j(\tau)). \tag{A.19}$$

Thus,

$$\|x_i(t) - x_j(t)\| \leq \|x_i(0) - x_j(0)\| + h \sum_{\tau=0}^{t-1} \|v_i(\tau) - v_j(\tau)\|. \tag{A.20}$$

However,

$$\|v_i(\tau) - v_j(\tau)\| \leq \sum_{r=1}^{3} |v_{ir}(\tau) - v_{jr}(\tau)| \leq \sum_{r=1}^{3} (z_{Nr}(\tau) - z_{1r}(\tau)) \triangleq z(\tau). \tag{A.21}$$

Eqs. (A.10), (A.20), and (A.21) imply:

$$\|x_i(t) - x_j(t)\| \leq M_x + h \sum_{\tau=0}^{t-1} z(\tau). \tag{A.22}$$

From f being non increasing, we obtain:

$$f(\|x_i(t) - x_j(t)\|) \geq f(M_x + h \sum_{\tau=0}^{t-1} z(\tau)). \tag{A.23}$$

Note that the RHS of (A.23) is independent of i and j. Defining:

$$g(t) = M_x + h \sum_{\tau=0}^{t-1} z(\tau), \tag{A.24}$$

(A.23) implies

$$a_{ij}(t) \geq f(g(t)), \; \forall i, j, \; \forall t \geq 0. \tag{A.25}$$

Recalling (A.17), it is straightforward to verify that all the coefficients in the RHS of (A.17) lie between $f(g(t))$ and $1 - (N-1)f(g(t))$. Thus, as the sum of coefficients is 1, to find a lower bound for the value of $v_{ir}(t+1)$, we put higher weights on lower valued v_{jr}'s, in particular, $1 - (N-1)f(g(t))$ on the least one, which is $z_{1r}(t)$ and $f(g(t))$ on the rest of them. Hence, we conclude:

$$v_{ir}(t+1) \geq \Big(1 - (N-1)f(g(t))\Big) z_{1r}(t) + \sum_{j=2}^{N} f(g(t)) z_{jr}(t). \tag{A.26}$$

As a result,

$$z_{1r}(t) \Big(1 - (N-1)f(g(t))\Big) z_{1r}(t) + \sum_{j=2}^{N} f(g(t)) z_{jr}(t). \tag{A.27}$$

Using the opposite process to build an upper bound for the value of $v_{ir}(t+1)$, we obtain:

$$z_{Nr}(t) \leq \Big(1 - (N-1)f(g(t))\Big) z_{Nr}(t) + \sum_{j=1}^{N-1} f(g(t)) z_{jr}(t). \tag{A.28}$$

Subtracting (A.27) from (A.28) implies:

$$z_{Nr}(t+1) - z_{1r}(t+1) \leq \Big(1 - N f(g(t))\Big)\Big(z_{Nr}(t) - z_{1r}(t)\Big). \tag{A.29}$$

By adding up Eq. (A.29) for $r = 1, 2, 3$, we obtain

$$z(t+1) \leq \Big(1 - N f(g(t))\Big) z(t), \tag{A.30}$$

or equivalently,

$$z(t+1) - z(t) \leq -N f(g(t)) z(t). \tag{A.31}$$

We now note that $hz(t) = g(t+1) - g(t)$. Since f is non increasing and $g(t+1) - g(t) = hz(t) \geq 0$, we have:

$$f(g(t))z(t) = f(g(t)).\frac{g(t+1) - g(t)}{h} \geq \frac{1}{h}\int_{g(t)}^{g(t+1)} f(y)dy. \tag{A.32}$$

Eqs. (A.31) and (A.32) imply:

$$z(t+1) - z(t) \leq \frac{-N}{h}\int_{g(t)}^{g(t+1)} f(y)dy. \tag{A.33}$$

The above equation holds for every $t \geq 0$. If we substitute variable t in Eq. (A.33) with t' and sum it up for $t' = 1, \ldots, t-1$. We obtain

$$z(t) - z(0) \leq \frac{-N}{h}\int_{g(0)}^{g(t)} f(y)dy. \tag{A.34}$$

Recalling the definition of $g(t)$ we conclude:

$$z(t) - z(0) \leq \frac{-N}{h}\int_{M_x}^{M_x + h\sum_{\tau=0}^{t-1} z(\tau)} f(y)dy. \tag{A.35}$$

If consensus does not occur, then $\sum_{\tau=0}^{\infty} z(\tau)$ diverges. Thus by taking t to infinity and noting that $\lim_{t\to\infty} z(t)$ exists as $z(t)$ is non-increasing and non-negative, if consensus did not occur, (A.35) would be modified as:

$$\lim_{t\to\infty} z(t) - z(0) \leq \frac{-N}{h}\int_{M_x}^{\infty} f(y)dy. \tag{A.36}$$

On the other hand, according to our assumption (A.13), we know that:

$$\begin{aligned} z(0) &= \textstyle\sum_{r=1}^{3}(z_{Nr}(0) - z_{1r}(0)) = \sum_{r=1}^{3} \max_{i,j}\{(v_{ir}(0) - v_{jr}(0)\} \\ &\leq \textstyle\sum_{r=1}^{3} \max_{i,j}\{(\|v_i(0) - v_j(0)\|\} = 3M_v < \frac{N}{h}\int_{M_x}^{\infty} f(y)dy. \end{aligned} \tag{A.37}$$

Eqs. (A.36) and (A.37) together imply that:

$$\lim_{t\to\infty} z(t) < 0, \tag{A.38}$$

which is a contradiction, since $z(t)$ is non-negative. Hence, consensus must occur and moreover, $\sum_{t=1}^{\infty} z(t)$ converges. Recalling (A.22) we conclude that $\|x_i(t) - x_j(t)\|$ is bounded for every i, j, i.e., there is $R \geq 0$ such that:

$$\|x_i(t) - x_j(t)\| \leq R, \forall i, j, \forall t \geq 0. \tag{A.39}$$

□

Applying Theorem A.2 to the extended C-S model (A.7) with:

$$f(x) = \frac{K}{(\sigma^2 + x^2)^\beta}, \tag{A.40}$$

results in the following corollary.

Corollary A.1. *Let the dynamics of a multi-agent system be described by (A.7) with f defined by (A.40). Assume that $K/\sigma^{2\beta} < 1/N$. Then, under either of the following two conditions, agents velocities converge to a common value:*

1. $\beta \leq 1/2$,
2. $\beta > 1/2$ *and*

$$M_v < \frac{NK}{3h(2\beta - 1)(M_x + \sigma)^{2\beta-1}}, \tag{A.41}$$

where M_x and M_v are defined by (A.10) and (A.11) respectively.