ABSTRACT

Title of dissertation: PROBLEMS IN DISTRIBUTED CONTROL SYSTEMS, CONSENSUS AND FLOCKING NETWORKS

Somarakis Christoforos, Doctor of Philosophy, 2015

Dissertation directed by: Professor John S. Baras
Department of Electrical & Computer Engineering

The present thesis discusses the consensus problem from a unification perspective. A general stability theory is developed discussing the majority of linear and nonlinear consensus networks with emphasis on the rate of convergence as an explicit estimate of the systems' parameters.

The discussion begins from the classical deterministic linear consensus problem in discrete and continuous-time setting. Vital assumptions are dropped and new types of non-uniform convergence are proven. All the related past results turn out to be only special cases of the developed framework, the central contribution of which is the derivation of explicit estimates on the rate of convergence. We proceed with the study of communication regimes that are governed by stochastic measures and we show that this setup is general enough to include many proposed stochastic settings as special cases. We highlight the strong interdependence between stochastic and deterministic signals and comment on how the imposed probabilistic regularity simply recaptures the deterministic sufficient conditions for consensus.

An important variant of the linear model is the delayed one where it is discussed in great detail under two theoretical frameworks: a variational stability analysis based on fixed point theory arguments and a standard Lyapunov-based analysis. The investigation revisits scalar variation unifying the behavior of old biologically inspired model and extends to the multi-dimensional (consensus) alternatives. We compare the two methods and assess their applicability and the strength of the results they provide whenever this is possible.

The obtained results are applied to a number of nonlinear consensus networks. The first class of networks regards couplings of passive nature. The model is considered on its delayed form and the linear theory is directly applied to provide strong convergence results. The second class of networks is a generally nonlinear one and the study is carried through under a number of different conditions. In additions the non-linearity of the models in conjunction with delays, allows for new type of synchronized solutions. We prove the existence and uniqueness of non-trivial periodic solutions and state sufficient conditions for its local stability. The chapter concludes with a third class of nonlinear models. We introduce and study consensus networks of neutral type. We prove the existence and uniqueness of a consensus point and state sufficient conditions for exponential convergence to it.

The discussion continues with the study of a second order flocking network of Cucker-Smale or Motsch-Tadmor type. Based on the derived contraction rates in the linear framework, sufficient conditions are established for these systems' solutions to exhibit exponentially fast asymptotic velocity. The network couplings are essentially state-dependent and non-uniform and the model is studied in both the ordinary and the delayed version. The discussion in flocking models concludes with two "noisy" networks where convergence with probability one and in the r^{th} square mean is proved under certain smallness conditions.

The linear theory is, finally, applied on a classical problem in electrical power networks. This is the economic dispatch problem (EDP) and the tools of the linear theory are used to solve the problem in a distributed manner. Motivated by the emerging field of Smart Grid systems and the distributed control methods that are needed to be developed in order to fit their architecture we introduce a distributed optimization algorithm that calculates the optimal point for a network of power generators that are needed to operate at, in order to serve a given load. In particular, the power grid of interconnected generators and loads is to be served at an optimal point based on the cost of power production for every single power machine. The power grid is supervised by a set of controllers that exchange information on a different communication network that suffers from delays. We define a consensus based dynamic algorithm under which the controllers dynamically learn the overall load of the network and adjust the power generator with respect to the optimal operational point.

PROBLEMS IN DISTRIBUTED CONTROL SYSTEMS, CONSENSUS AND FLOCKING NETWORKS

by

Christoforos Somarakis

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor John S. Baras, Chair/Advisor
Professor Konstantina Triviza
Professor Stuart S. Antman
Professor Eitan Tadmor
Professor Perinkulam S. Krishnaprasad

Preface

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Maryland. The research described herein was conducted under the supervision of Professor J. S. Baras in the Department of Electrical & Computer Engineering, between September 2008 and April 2015.

This work is to the best of my knowledge original, except where acknowledgments and references are made to previous works. Neither this nor any other substantially similar dissertation has been or is being submitted for any other degree, diploma or other qualification at any other university. This dissertation contains less than 70,000 words.

Part of this work has been presented in the following publications: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

I would like to acknowledge financial support from US Army Research Office under MURI grant W911NF-08-1-0238, US Air Force Office of Scientific Research under MURI grant FA9550-09-1-0538, National Science Foundation grant CNS-1018346, National Institute for Standards and Technology grant 70NANB11H148, National Science Foundation grant CNS-1035655 and the Academy of Athens, Greece for the financial support from "Philonos Vasileiou" bequest.

Dedication

To Leia.

Every elegant argument of this thesis, signifies her Love and Affection.

Every weak argument, signifies my incompetence to match these Feelings.

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I'd like to thank my advisor, John S. Baras for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past six years. He has always made himself available for help and advice and there has never been an occasion when I've knocked on his door and he hasn't given me time. It has been a pleasure to work with and learn from such an extraordinary individual.

Thanks are due to Professor Konstantina Triviza, Professor Stuart Antman, Professor P.S. Krishnaprasad and Professor Eitan Tadmor for agreeing to serve on my thesis committee and for sparing their invaluable time reviewing the manuscript.

An overdue thanks to my teachers and mentors during my years at the National Technical University of Athens, Professor Th.M. Rassias, Professor G.P. Papavassilopoulos and Professor G. Cambourakis.

A wholehearted thanks to my first and everlasting mathematics tutor, Maria Sfakianaki. Her immersive teaching not only introduced me to the astonishing world of Mathematics, but also defined my conclusive bond with it.

I owe my deepest thanks to my family - my mother, my father and my sister who have always stood by me and guided me through my career, and have pulled me through against impossible odds at times.

Table of Contents

List of Figures

List of Abbreviations

EDP	Economic Dispatch Problem
FPT	Fixed Point Theory
LTI	Linear Time Invariant
LTV	Linear Time Varying

Chapter 1: Introduction

Self-organized dynamics lie in the core of modern complex dynamical systems, a most interesting branch of which is their application in the field of networked control systems.

Examples of networks that illustrate a collective behavior as a result of local interaction among the nodes of the network are ubiquitous both in nature and in human societies. Ants cooperate together to form a nest or to transfer provisions and birds form flocks and fly together enhancing their hunting abilities. Humans interact and socialize by exchanging opinions and sometimes may converge to a fairly common view (especially after choosing a leader). Engineers build mobile communication or robotic networks which coordinate their behavior by local exchange of information. These are all examples of what in the control community is called collaborative control of multi-agent systems. The self-organized aspect of the aforementioned examples is usually understood by a decentralized, local exchange of information. The central phenomenon is the manner with which agents, as individuals, exchange information on a state of interest and update this state, so that eventually all agents' states concentrate around a common value. These problems are known as consensus problems and enjoy a durable interdisciplinary interest in the applied sciences. As a result several mathematical models have been introduced to appraise the so-called *emergence of consensus* among agents.

The problem of consensus is very old in the literature and it has been proposed under numerous variations. From the classical linear model to complex nonlinear variations, all these systems basically sustain a common underlying mechanism that essentially characterizes the asymptotic convergence to a common value. Researchers use consensus-based algorithms to model the computational aspect of biological, robotic, social and other communication networks in the process of exchanging information in order to agree on a particular state of interest. From the seminal work of Tsitsiklis et al. on distributed computation [17] to the monograph of Smith on the Theory of Competitive and Cooperative Systems [18] and from the biological model of bird co-ordination, proposed by Vicsek et al. [19] to the mathematical proof for its asymptotic speed alignment proved by Jadbabaie et al [20] and from there to an ocean of related results, consensus systems carry the beacon of research in the networked control community and applied mathematics for over 20 years.

The related mathematics, however, go even more back in time. The term "consensus" was initially introduced in the work of DeGroot [21] in 1974 whereas the central mathematical concept was first appeared in a paper of Markov [22] in 1906. This tool, known as the contraction coefficient, measures the contraction effect of a stochastic matrix applied on a vector with respect to a particular sub-space.

Despite the enormous literature in consensus networks there is still a number of aspects of the problem yet to be explored. For instance, while the question of convergence and its sufficient conditions is pretty clear, the rate at which this happens remains open for the generic case of the continuous-time model, even in its linear version. This issue is not only important in real world applications but also vital in certain nonlinear 2^{nd} order consensus systems, known as flocking networks, for proving simple asymptotic convergence. The same question is yet to be answered in an important variation of the problem that involves delayed signals. All these questions can be repeated for nonlinear variations of the problem causing a multiplicative increase on the new research problems, a researcher is called to shed light upon. Furthermore, stochastic variations are to be considered within a universal framework the special cases of which are the very many probabilistic setups for the network communication topology proposed in the literature.

1.1 Contribution of the Thesis

The present thesis discusses the consensus problem from its most fundamental form and develops a general theory for the stability of the solutions of these systems. The analysis concentrates on the rate of convergence as an explicit estimate of the systems parameters.

The discussion begins from the classical deterministic linear consensus problem in discrete and continuous-time setting. Vital assumptions are dropped and new types of non-uniform convergence are proven. All the related past results turn out to be only special cases of the developed framework, the central contribution of which is the derivation of explicit estimates on the rate of convergence. We proceed with the study of communication regimes that are governed by stochastic measures and we show that this setup is general enough to include many proposed stochastic settings as special cases. We highlight the strong interdependence between stochastic and deterministic signals and comment on how the imposed probabilistic regularity simply recaptures the deterministic sufficient conditions for consensus.

An important variant of the linear model is the delayed one where it is discussed in great detail under two theoretical frameworks. The connection between the different types of delays that are present in linear system and their effect on the stability of the solutions is the main objective of the discussion. The study of stability with emphasis on the rate of convergence leads us to investigate a scalar functional equation and effectively a very old problem on the subject of delayed differential equations with asymptotically constant solutions emanating from mathematical biology and population growth models. The stability analysis is based on fixed point theory for the linear time invariant case and on a Lyapunov-Razumikhin argument for the general linear case. The results that were derived for the scalar case are extended to the multidimensional one, i.e. the consensus networks. We compare the two methods and assess their applicability and the strength of the results they provide whenever this is possible.

The obtained results are applied to a number of nonlinear consensus networks. The first class of networks regards couplings of passive nature. The model is considered on its delayed form and the linear theory is directly applied to provide strong convergence results. The second class of networks is a generally nonlinear one and the study is carried through under basically two sets of conditions. The first set will be called "non-monotonic" and it is considered as a small deviation of the linear consensus system. The analysis relies on stability in variation and fixed point theory. The second set of assumptions is imposed on the first derivative of the nonlinear coupling functions and it essentially enables certain crucial features of the behavior of solutions in the classical linear consensus systems. Non-linearity allows for new type of synchronized solutions. We prove the existence and uniqueness of non-trivial periodic solutions and state sufficient conditions for its local stability. The chapter concludes with a third class of nonlinear models. We introduce and study consensus networks of neutral type. We prove the existence and uniqueness of a consensus point and state sufficient conditions for exponential convergence to it.

The analysis continues with the study of a second order flocking network of Cucker-Smale or Motsch-Tadmor type. Based on the derived contraction rates in the linear framework, sufficient conditions are established for these systems' solutions to exhibit exponentially fast asymptotic velocity. The network couplings are essentially state-dependent and non-uniform and the model is studied in both the ordinary and the delayed version. The discussion in flocking models concludes with two "noisy" networks where convergence with probability one and in the r^{th} square mean is proved under certain smallness conditions.

The linear theory is, finally, applied on a classical problem in electrical power networks. This is the economic dispatch problem (EDP) and the tools of the linear theory are used to solve the problem in a distributed manner. Motivated by the emerging field of Smart Grid systems and the distributed control methods that are needed to be developed in order to fit their architecture we introduce a distributed optimization algorithm that calculates the optimal point for a network of power generators that are needed to operate at, in order to serve a given load. In particular, the power grid of interconnected generators and loads is to be served at an optimal point based on the cost of power production for every single power machine. The power grid is supervised by a set of controllers that exchange information on a different communication network that suffers from delays. We define a consensus based dynamic algorithm under which the controllers dynamically learn the overall load of the network and adjust the power generator with respect to the optimal operational point.

1.2 Organization of the Thesis

In the rest of this introductory note I briefly present the chapters the thesis consists of.

Chapter 2 develops the basic principles that will be used throughout this work. We introduce basic notations and definitions and we review the main concepts that will come at hand from various mathematical theories. The contribution is a number of non-trivial extensions of the central mathematical concept of the contraction coefficient. These results constitute the foundation of the analysis on which the main theorems of the subsequent sections rely.

Chapter 3 discusses the linear consensus model in the discrete and the continuous-time settings, under both deterministic and stochastic variants.

Chapter 4 introduces the problem of delays in consensus networks and studies its effect on the stability and rate of convergence.

Chapter 5, 6 and 7 are applications of the results obtained in Chapter 3 and Chapter 4.

Chapter 5 examines the first three types of nonlinear models, i.e. passive systems, general nonlinear systems and systems of neutral type.

Chapter 6 nonlinear second order consensus (flocking) networks.

Chapter 7 applies the results of Chapter 4 to solve the EDP in a distributed manner and a Smart-Grid compatible setup.

Finally, in Chapter 8 an overall discussion is held where the derived results are enveloped and open questions and prospects for future research are posed.

Chapter 2: Fundamentals

This thesis employs various mathematical theories in order to state its results in a rigorous way. Consequently, a large number of notations and symbols will be used, the majority of which will be defined locally. On condition that there is no direct notation conflict, each chapter will preserve its own nomenclature. Nevertheless, some symbols are to be reserved globally representing common ground concepts. Here, we introduce these universal symbols and we review elements of mathematical theories and fields that will come at hand throughout this work.

The purpose of this chapter is, therefore, primarily preparatory. It introduces the basic symbols and it serves as a quick reference of the results, the main chapters of the thesis rely on. The contribution of this chapter is limited. We provide a number of generalizations of the contraction coefficient. Additionally, a secondary result on the completeness property of a metric space will be provided.

The discussion is drawn from classical textbooks such as [23, 24, 25] for Graph Theory and Algebraic Methods in Multi-Agent Systems, [26, 27, 28] for the theory of Non-Negative Matrices and the contraction coefficient, [29, 30] for Dynamical Systems Theory, [31, 32, 33] for Stochastic Differential Equations, [34, 35, 36] for Fixed Point Theory and applications in the Stability of Differential Equations and

[37] for Convex Analysis and Linear Inequatilies.

2.1 Notations and Definitions

$\mathbb{Z}$ denotes the set of integers, $\mathbb{N}$ the set of natural numbers and $\mathbb{R}$ the set of the real numbers. For $N \in \mathbb{N}$

$$[N] := \{1, \ldots, N\}$$

denotes the set of the first N positive integers. $\mathbb{R}^N$ is the N-dimensional Euclidean space and $\mathbf{x} \in \mathbb{R}^N$ is considered as a column vector, unless otherwise stated. The *agreement* or *consensus* space $\Delta \subset \mathbb{R}^N$ is defined as

$$\Delta = \{\mathbf{x} \in \mathbb{R}^N : x_1 = x_2 = \cdots = x_N\}$$

A *rank-1 matrix* is the $N \times N$ matrix M that has identical rows. For a rank-1 matrix M, $M\mathbf{x} \in \Delta$, $\forall\ \mathbf{x} \in \mathbb{R}^N$. The *spread* of a vector $\mathbf{x} \in \mathbb{R}^N$ is

$$S(\mathbf{x}) = \max_i x_i - \min_i x_i. \tag{2.1}$$

This quantity will serve as a pseudo-norm for the stability analysis to follow. Indeed it is always non-negative and satisfies the triangle inequality, but

$$S(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} \in \Delta.$$

By $\mathbb{1}$ we understand the $N-$dimensional vector with all entries equal to 1 and obviously $S(\mathbb{1}c) = 0$ for any $c \in \mathbb{R}$. $I_{N\times N}$ stands for the $N \times N$ identity matrix, $||\cdot||_p$ stands for p-norm so that $\mathbf{x}^T\mathbf{x} = ||\mathbf{x}||_2^2$. $|\cdot|$ will usually denote the absolute value of a number and $||\cdot||$ an arbitrary norm of a linear space.

For a closed subset I of $\mathbb{R}$, $L^1(I, \mathbb{R}^N)$ denotes the space of integrable functions defined on I and taking values in $\mathbb{R}^N$. Similarly $C^l(I, \mathbb{R}^N)$ denotes the space of functions with $l \geq 0$ continuous derivatives defined accordingly. For $\mathbf{z} \in L^1$ or $\mathbf{z} \in C^l$ we define the set

$$W_{I,\mathbf{z}} = \big[\max_i \max_{s \in I} z_i(s), \min_i \min_{s \in I} z_i(s)\big]$$

in particular if $I = \{t\}$ is a singleton then we will use the notation

$$W_{t,\mathbf{z}} = \big[\min_i z_i(t), \max_i z_i(t)\big].$$

The length of $W_{I,\mathbf{z}}$ is the spread of $\mathbf{z}$

$$S_I(\mathbf{z}) = \max_i \max_{s \in I} z_i(s) - \min_i \min_{s \in I} z_i(s). \tag{2.2}$$

This is a natural generalization of (2.1). It serves as a pseudonorm with respect to the *agreement function space*

$$\Delta_I = \big\{\mathbf{z} \in L^1(I, \mathbb{R}^N) : z_i(t) = z_j(t), t \in I, i = 1, \ldots, N\big\}.$$

For function spaces the norm is defined to be the as $|\boldsymbol{z}| = \sup_{t \in I} |\mathbf{z}(t)|$.

Let for any $t \in \mathbb{R}$, $\mathbf{x}_t \in L^1(I, \mathbb{R}^N)$ or $C^1(I, \mathbb{R}^N)$ for some compact $I \subset \mathbb{R}$ defined as $\mathbf{x}_t(s) = \mathbf{x}(t+s), s \in I$. We will write

$$\mathbf{x}(t) \to \Delta \text{ as } t \to \infty \quad \Leftrightarrow \quad x_i(t) \to k,\ i = 1, \ldots, N \text{ as } t \to \infty$$

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

Due to potential discontinuities in the systems' parameters, the solutions of continuous-time dynamics are occasionally defined to be absolutely continuous functions of time. The appropriate time derivative $\frac{d}{dt}$ to be used is the upper righ Dini derivative. For a real-valued function $f(t)$ this is defined as

$$\frac{d}{dt} f(t) = \limsup_{h \to 0^+} \frac{f(t+h) - f(t)}{h}.$$

The thesis systematically discusses the asymptotic behavior of solutions of differential equations emphasizing on the rate of convergence. For fixed $t_0 \in \mathbb{R}$, a *rate function* $h : [t_0, \infty) \to [1, \infty)$ is an integrable function with $h(t_0) = 1$, monotonically increasing so that $h(t) \to \infty$ as $t \to \infty$. An example of rate function is the exponential $e^{\theta(t-t_0)}$ for some $\theta > 0$ that will be reserved to denote the explicit rate estimate.

2.2 Algebraic Graph Theory

By a *topological directed graph* $\mathbb{G}$ we define the pair $([N], E)$ where $[N]$ is the (static) set of nodes or nodes, $E = \{(i,j) : i, j \in [N]\}$ is the set of edges where $(i,j) \neq (j,i)$. The *degree* N_i of a node $i \in [N]$ is defined as the subset $N_i := \{j \in [N], (i,j) \in E\}$, all the nodes adjacent to i. The graph $\mathbb{G}$ is *routed-out branching* if there exists a node $i \in [N]$ (called the route of the graph) such that for any $j \neq i \in [N]$ there is a *path* of edges $(l_k, l_{k-1})|_{k=0}^{m}$ such that $l_0 = i$ and $l_m = j$. The graph $\mathbb{G}$ is *connected* if any node is a route. For two graphs $\mathbb{G}_1 = ([N], E_1)$ and $\mathbb{G}_2 = ([N], E_2)$, we say that $\mathbb{G}_1$ is a sub-graph of $\mathbb{G}_2$ if $E_1 \subset E_2$. The *adjacency matrix* A is a $0-1$, $N \times N$ matrix with elements $A_{ij} = 1 \Leftrightarrow (i,j) \in E$. The *degree matrix* $D := \text{Diag}\big[d_i\big]$. Finally, the *Laplacian* of $\mathbb{G}$ is the matrix $L := D - A$ with the sum of its rows be identically equal to zero. The spectral properties of L are vital in this work. In particular, 0 is a always an eigenvalue of L and for any other eigenvalue $\lambda \in \mathbb{C}$ of

L it holds that $\Re\{\lambda\} > 0$ if and only if $\mathbb{G}$ is connected (see also Proposition 2.2.1, below).

Let us now introduce a terminology that comes from the theory of non-negative matrices and describes similar ideas. We say that two nodes $i, j \in [N]$ *communicate* if there is a path from i to j and a path from j to i. A node is *essential* if whenever there is a path from i to j then there is a path from j to i. A node is called *inessential* if it is not essential. All essential nodes are divided into communication classes and all inessential nodes that communicate with at least one node may be divided into inessential classes such that all nodes within a class communicate. All such classes are *self-communicating*. Each remaining inessential node communicates with no nodes and individually forms an inessential class called *non self-communicating*. These definitions will also be of use in the present work.

By $\mathcal{S}$ we denote the family of graphs with fixed N nodes and self-edges on every node, and by $\mathcal{T} \subset \mathcal{S}$ the set of graphs each of which is routed-out branching.

A *weighted graph* is a graph where the edge between two nodes is associated with a positive number. This coupling weight, or coupling rate, is a positive number, usually denoted by a_{ij}, and it quantifies the effect of j on i. Unless otherwise specified $a_{ii} \equiv 0$. The corresponding degree, the adjacency and the Laplacian matrices are defined as $D = \mathrm{Diag}[\sum_j a_{ij}]$, $A = [a_{ij}]$, $L = D - A$ accordingly.

2.2.1 Agreement dynamics

The term *agreement dynamics* will be used to describe the elementary static time invariant linear consensus system:

$$\dot{\mathbf{x}} = -L\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{x}^0 \in \mathbb{R}^N \tag{2.3}$$

where

$$L = \begin{bmatrix} \sum_j a_{1j} & -a_{12} & -a_{13} & \dots & -a_{1N} \\ -a_{21} & \sum_j a_{2j} & -a_{23} & \dots & -a_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_j a_{Nj} & -a_{N2} & -a_{N3} & \dots & \sum_j a_{Nj} \end{bmatrix}$$

is the weighted Laplacian matrix and $a_{ij} \geq 0$ are the constant coupling weights. The dynamics of (2.3) can be analyzed by standard linear algebra tools (see for example [38]), much of which is summarized below.

Proposition 2.2.1. *Let $\mathbb{G}$ be a routed-out branching graph and denote by L its Laplacian matrix. The following properties hold:*

1. *$L\mathbf{p} = 0$ if and only if $\mathbf{p} \in \Delta$.*
2. *The spectrum of L, $\{\lambda_i\}|_{i=1}^N$, can be enumerated so that $\lambda_1 = 0$ and*

$$0 < \Re(\lambda_2) \leq \Re(\lambda_3) \leq \cdots \leq \Re(\lambda_N).$$

3. *The left eigenvector of L, $\mathbf{c}$ is unique, up to normalization and non-negative elementwise.*

4. There exists $J > 0 : ||e^{-Lt} - \mathbb{1}\mathbf{c}^T|| \leq Je^{-\Re(\lambda)t}$ where $\Re\{\lambda\} := \Re\{\lambda_2\}$, i.e. λ stands for the eigenvalue with the smallest non-zero real part.

5. $e^{-Lt}L$ is a Laplacian matrix that vanishes exponentially fast as $t \to \infty$, with rate no worse than $\Re(\lambda)$.

Proof. For *(1)*, *(2)* see Propositions 3.8 and 3.11 of [25] respectively.

For *(3)* we work as follows: *(2)* implies that the rank of L equals $N - 1$. Then if $\mathbf{c}_1, \mathbf{c}_2$ are two left eigenvectors of L associated with the zero eigenvalue then $\mathbf{c}_1 = \epsilon \mathbf{c}_2$ for some $\epsilon > 0$. From the normalization condition $\mathbf{c}_1^T \mathbb{1} = \epsilon \mathbf{c}_2 \mathbb{1} = 1$ so that in fact $\mathbf{c}_1 = \mathbf{c}_2$ and uniqueness follows. To conclude it suffices to show that the elements of $\mathbf{c}$ are of the same sign. Assume, for the sake of contradiction that the result does not hold and without loss of generality let $c_1, c_2, \ldots, c_r$ be the negative components of $\mathbf{c}$. Then $\mathbf{c}^T \mathbb{1} = 1$ implies $r \leq N - 1$. Next, $\mathbf{c}^T L = 0$ gives

$$\sum_{j=1}^{N} a_{ij} c_i = \sum_{j=1}^{r} a_{ji} c_j + \sum_{j=r}^{N} a_{ji} c_j, \quad i = 1, \ldots, N$$

Take the sum of the first r equations and after cancelling the common terms observe that the resulting equation has negative left hand-side and positive right hand side, a contradiction. The zero terms of $\mathbf{c}$ were neglected as they play no role in the proof.

For *(4)*, consider the Jordan canonical form of $L := PJ(L)P^{-1}$ so that $e^{-Lt} = Pe^{-J(L)t}P^{-1} =$. For the Jordan blocks, we know that $J(\lambda_1) = J(0) = 0$ so that both $LP = PJ$ and $P^{-1}L = JP^{-1}$ have a zero first row. So the first row of P^{-1} is the left eigenvector $\mathbf{c}$, canonicalized so that $\mathbf{c}^T \mathbb{1} = 1$ and the first column of L is the right eigenvector of L chosen to be $\mathbb{1}$. The projection of any vector $\boldsymbol{\zeta}$ onto Δ is $\mathbb{1}\mathbf{c}^T \boldsymbol{\zeta}$. Let L have $q \leq N$ distinct eigenvalues. Again $\lambda_1 = 0$ and $\Re\{\lambda_i\} > 0$ for

$i \geq 2$. Since any $\boldsymbol{\zeta} \in \mathbb{C}^N$ can be written as $\boldsymbol{\zeta} = \mathbf{w}_1 + \cdots + \mathbf{w}_q$ where $\mathbf{w}_j \in M(\lambda_j)$, the generalized eigenspace for λ_j and in particular, $\mathbf{w}_1 = \mathbb{1}\mathbf{c}^T\boldsymbol{\zeta}$. Then by standard calculations

$$e^{-Lt}\boldsymbol{\zeta} = \mathbb{1}\mathbf{c}^T\boldsymbol{\zeta} + \sum_{j=2}^{q} e^{-\lambda_j t} \sum_{k=0}^{\iota(-\lambda_j)-1} (-L + \lambda_j I)^k \frac{t^k}{k!}\mathbf{w}_j \tag{2.4}$$

where $\iota(\lambda_j)$ is the maximum degree of the generalized vectors for λ_j. The rest of the proof follows exactly this of Theorem 5.8 in [38].

For (5), observe that $e^{-Lt}L$ is a Laplacian by the definition of the exponential of a matrix and the fact that any power of a Laplacian matrix is Laplacian as well as the sum of two Laplacian matrices is a Laplacian matrix. Finally, since $e^{-Lt}L = (e^{-Lt} - \mathbb{1}\mathbf{c}^T)L$ we obtain

$$||Le^{-Lt}|| \leq ||L||Je^{-\Re(\lambda)t}.$$

□

2.2.1.1 On Laplacians of non-negative symmetric matrices

In the special case of an undirected network, the Laplacian is a symmetric positive semi-definite matrix with real spectrum: $\{0 < \lambda_2 \leq \cdots \leq \lambda_N\}$ and this makes the analysis significantly simpler[5, 7]. In particular,

$$\mathbf{c} \in \mathbb{R}^N : c_i \geq 0, \sum_i c_i \equiv 1, \mathbf{c}^T L = 0 \Rightarrow \mathbf{c} = \mathbb{1}\frac{1}{N}$$

and the convergence to $\mathbf{c}^T\mathbf{x}_0$ is exponential with rate $\lambda_2 = \lambda$. The second smallest eigenvalue of a symmetric Laplacian L is called the *Fiedler number* of the adjacency matrix A [39] and its variational definition is

$$\lambda = \inf_{\mathbf{x} \notin \Delta} \frac{\mathbf{x}^T L \mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Clearly the Fiedler number is positive if and only if the corresponding graph $\mathbb{G}$ is (simply) connected. See also [40, 39, 25]. For more on Graph Theory and relevant methods on multi-agent systems, the interested reader is referred to [23, 24, 25].

2.3 Non-Negative Matrix Theory

A *non-negative matrix* $P = \{p_{ij}\}$ is such that $p_{ij} \geq 0$ for all i, j.[1] The non-negative matrix P is *generalized stochastic*, or m-stochastic, if

$$\sum_j p_{ij} = m, \qquad \forall\, i = 1, \ldots, N.$$

A crucial property of an m-stochastic matrix is that m is always an eigenvalue of the matrix. For $m = 1$ we have the well-known stochastic matrix. Let $\mathcal{M}$ denote the collection of m-stochastic matrices for some $m \in \mathbb{R}$. We will now introduce and discuss the standard mathematical tool that handles infinite products of stochastic matrices.

Given an m-stochastic matrix $P = [p_{ij}]$, the quantity

$$\kappa(P) = \frac{1}{2} \max_{i,j} \sum_s |p_{is} - p_{js}| = m - \min_{i,j} \sum_s \min\{p_{is}, p_{js}\} \tag{2.5}$$

is *the coefficient of ergodicity* of P, a crucial set of properties of which is presented below:

[1] Unless otherwise specified each matrix is supposed to be square and of dimension $N \times N$.

Theorem 2.3.1. *For any m-stochastic matrix P and $\mathbf{z} \in \mathbb{R}^N$ the following properties hold:*

1. $||\boldsymbol{\delta} P|| \leq \kappa(P)||\boldsymbol{\delta}||$, *for all real row vectors $\boldsymbol{\delta}$ such that $\boldsymbol{\delta}\mathbb{1} = 0$.*

2. $S(P\mathbf{z}) \leq \kappa(P)S(\mathbf{z})$

3. $|\lambda| \leq \kappa(P)$, *for any (possibly complex) eigenvalue λ of P with the property that $\lambda \neq m$.*

Proof. See [26]. □

The coefficient of ergodicity measures the averaging effect of stochastic matrices and it will play an instrumental role in the present thesis. The majority of the convergence results are critically based on it. Its history dates back to one of Markov's first papers [22]. There exists an abundance of similar ideas in the literature: the coefficient of ergodicity is also known as contraction coefficient, Markov coefficient, Dobrushin coefficient, Birkhoff coefficient, Hajnal diameter, as each corresponding researcher has arrived at it independently and/or under different setups, [27, 28]. For a recent review on the coefficients of ergodicity we refer to [41].

Remark 2.3.2. Property 1 of Theorem 2.3.1 leads to the sub-multiplicative property: If P_1 and P_2 are m-stochastic matrices, their product P_1P_2 constitutes an m^2 stochastic matrix and it satisfies

$$\kappa(P_1P_2) \leq \kappa(P_1)\kappa(P_2).$$

The sub-multiplicative property becomes particularly useful when $m = 1$ exactly because the set of stochastic matrices becomes closed under matrix multiplication. As Theorem 2.3.1 suggests, the coefficient κ applies to dynamics of the type

$$\mathbf{w} = P\mathbf{z} \tag{2.6}$$

with P being m-stochastic.

We will prove two important extensions of Theorem 2.3.1. The first one explains that similar results can be applied when (2.6) is replaced by a double inequality.

Theorem 2.3.3. *Let $\underline{P} = [\underline{p}_{ij}]$ and $\overline{P} = [\overline{p}_{ij}]$ be two m-stochastic matrices and $\mathbf{z}, \mathbf{y} \in \mathbb{R}^N$. If $\underline{P}\mathbf{y} \leq \mathbf{z} \leq \overline{P}\mathbf{y}$, then*

$$S(\mathbf{z}) \leq \left(m - \min_{h,h' \in [N]} \sum_j \min\{\bar{p}_{hj}, \underline{p}_{h'j}\}\right) S(\mathbf{y}).$$

Proof. For fixed $h, h' \in [N]$, $z_h - z_{h'} \leq \sum_j p_j y_j$ with $p_j := \overline{p}_{hj} - \underline{p}_{h'j}$.

Let j',j'' denote the indices in $[N]$ such that $p_{j'} > 0$ and $p_{j''} < 0$ and note that $\sum_j p_j \equiv 0$.

$$\begin{aligned} 0 < \theta := \sum_{j'} p_{j'} = \sum_{j'} |p_{j'}| = -\sum_{j''} p_{j''} = \sum_{j''} |p_{j''}| = \\ = \frac{1}{2}\sum_j |p_j| = \frac{1}{2}\sum_j |p_j| = \frac{1}{2}\sum_j |\overline{p}_{h'j} - \underline{p}_{hj}| \end{aligned}$$

and see that for appropriate $h, h' \in [N]$

$$S(\mathbf{z}) = z_h - z_{h'} = \theta\left(\frac{\sum_{j'} p_{j'} y_{j'}}{\sum_{j'} p_{j'}} - \frac{\sum_{j''} p_{j''} y_{j''}}{\sum_{j''} p_{j''}}\right) \leq \theta S(\mathbf{y})$$

and the expression for $\theta = m - \min_{h,h' \in [N]} \sum_j \min\{\bar{p}_{hj}, \underline{p}_{h'j}\}$ can be obtained in view of the identity $|\alpha - \beta| = \alpha + \beta - 2\min\{\alpha, \beta\}$ and the fact that $\underline{P}$ and $\overline{P}$ are m-stochastic. $\square$

The proof of Theorem 2.3.3 follows the steps of the original proof of the contractive property of κ, provided by Markov in [22] (see also [26]).

Another extension of Theorem 2.3.1 considers the case when the P acts as an abstract linear operator an appropriate space of functions and it is summarized in the following result:

Theorem 2.3.4. *Let I be a compact subset of $\mathbb{R}$ and assume that for any compact $I' \subset I$, $W_{I'} = \int_{s\in I'} P(s)\,ds \in \mathcal{M}$ and W_I is m-stochastic. If $\mathbf{w} = \int_{s\in I} P(s)\mathbf{z}(s)\,ds$, then*

$$S(\mathbf{w}) \leq \kappa(W_I)S(\mathbf{z}^*)$$

for some $\mathbf{z}^* = \big(z_1(s_1), \ldots z_N(s_N)\big)$ for $s_i \in I$ and

$$\begin{aligned} \kappa(W_I) &= \frac{1}{2}\max_{h,h'}\sum_{k=1}^{N}\int_{s\in I}|p_{hk}(s) - p_{h'k}(s)|\,ds \\ &= m - \min_{h,h'}\sum_{k=1}^{N}\min\left\{\int_{s\in I} p_{hk}(s)\,ds, \int_{s\in I} p_{h'k}(s)\,ds\right\} \end{aligned} \tag{2.7}$$

The proof of this result relies on the first mean value theorem for integration and a technical lemma, both of which are cited below:

Lemma 2.3.5 (The first mean value theorem for integration)**.** *If $G \in C^0[J,\mathbb{R}]$ and ϕ is integrable that does not change sign on J then there exists $x \in J$ such that*

$$G(x)\int_J \phi(t)\,dt = \int_J G(t)\phi(t)\,dt.$$

We recall that two vectors $\mathbf{x}, \mathbf{y}$ are sign compatible if $x_i y_i \geq 0$ for all i.

Lemma 2.3.6. *Suppose $\boldsymbol{\delta} \in \mathbb{R}^N$ such that $\boldsymbol{\delta}^T\mathbb{1} = 0$ and $\boldsymbol{\delta} \neq 0$. Then there is an index $\mathcal{I} = \mathcal{I}(\delta)$ of ordered pairs (i,j) with $i, j \in [N]$ such that*

$$\boldsymbol{\delta}^T = \sum_{(i,j)\in\mathcal{I}} \frac{T_{ij}}{2}(\mathbf{e}_i - \mathbf{e}_j)$$

where $T_{ij} > 0$, $\mathbf{e}_i$ *is the row vector* $(0, \dots, 0, 1, 0, \dots, 0)$ *with* 1 *at the* i^{th} *position,* $\mathbf{e}_i - \mathbf{e}_j$ *is sign compatible to* $\boldsymbol{\delta}$ *for all* i, j*. Thus* $||\boldsymbol{\delta}||_1 = \sum_{(i,i)\in\mathcal{I}} T_{ij}$*.*

Proof. This is Lemma 1.1 of [26]. □

Proof of Theorem 2.7. Pick $h, h^{'} \in [N]$. Then for $\mathbf{p}_h, \mathbf{p}_{h'}$ the h^{th} and $h^{'th}$ rows of P respectively, we have

$$\int_{s\in I} \big(\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)\big)\mathbf{z}(s)\, ds$$

Since $N < \infty$, there is a partition $\{I_l\}_{l=1}^m$ of I which depends on h, h' such that for any I_l, $p_{hk}(s) - p_{h'k}(s)$ does not change sign in for $s \in I_l$, $k \in [N]$ and it is not identically zero. Then for fixed I_l we apply Lemma 2.3.5 to obtain

$$\sum_k \int_{s\in I_l} \big(p_{hk}(s) - p_{h'k}(s)\big) z_k(s)\, ds = \sum_k \int_{s\in I_l} \big(p_{hk}(s) - p_{h'k}(s)\big)\, ds z_k(s^*_k) = \boldsymbol{\delta}_l^T \mathbf{z}_l^*$$

for some $s^*_k = s(I_l, h, h')$ and

$$\boldsymbol{\delta}_l^T = \int_{I_l} \big(\mathbf{p}_h(s) - \mathbf{p}'_h(s)\big)\, ds \neq 0, \qquad \mathbf{z}_l^* = (z_1(s_1^*), \dots z_N(s_N^*))^T.$$

By Assumption $\int_{I_l} P(s)\, ds \in \mathcal{M}$ and therefore $\boldsymbol{\delta}_l^T \mathbb{1} = 0$. Hence, Lemma 2.3.6 is applied and together with the triangle inequality

$$|\boldsymbol{\delta}_l^T \mathbf{z}_l^*| \leq \frac{1}{2} ||\boldsymbol{\delta}_l||_1 S(\mathbf{z}_l^*)$$

(see also [26]). Then if we let $S(\mathbf{z}*) = \max_l S(\mathbf{z}_l^*)$, we obtain the bound

$$\begin{aligned} S(\mathbf{w}) &= \max_{h,h'} \left| \int_{s\in I} \big(\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)\big)\mathbf{z}(s)\, ds \right| \\ &= \sum_l |\boldsymbol{\delta}_l^T \mathbf{z}_l^*| \leq \max_{h,h'} \frac{1}{2} \int_I ||\mathbf{p}_h(s) - \mathbf{p}_{h'}(s)||_1 ds S(\mathbf{z}^*). \end{aligned}$$

Finally, from the identity

$$|x - y| = x + y - 2\min\{x, y\}, \qquad \forall x, y \in \mathbb{R}$$

and the fact that $\forall h, h' \in [N]$

$$\sum_k \int_{s\in I} p_{hk}(s) ds = \sum_k \int_{s\in I} p_{h'k}(s)\, ds = m$$

we get:

$$\frac{1}{2}\max_{h,h'} \sum_k \int_{s\in I} |p_{hk}(s) - p_{h'k}(s)|\, ds =$$

$$= m - \min_{h,h'} \sum_k \min\left\{ \int_{s\in I} p_{hk}(s)\, ds, \int_{s\in I} p_{h'k}(s)\, ds \right\}.$$

$\square$

Similarly, for

$$\mathbf{w} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)\mathbf{z}(q)\, dq ds$$

one can show following the proof of Theorem 2.3.4 that if

$$W_I^{(2)} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)\, dq ds$$

is stochastic, then

$$S(\mathbf{w}) \leq \kappa(W_I^{(2)}) S(\mathbf{z}^*)$$

for some $\mathbf{z}^* = \big(z_1(s_{(ij)}^{(1)}), z_2(s_{(ij)}^{(2)}), \ldots, z_N(s_{(ij)}^{(N)})\big)$ all $s_{(ij)}^{(l)}$ of which are in $I_1 \cup I_2$. Finally, the sub-multiplicativity property for pairs of stochastic matrices of the particular form discussed in this section, applies to expressions of the type

$$\mathbf{w} = \int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)\mathbf{z}(q)\, dq ds$$

so long as $\int_{s\in I_1} P_1(s) \int_{q\in I_2(s)} P_2(q)\, dq ds$ is stochastic.

Regardless if we are working with products of matrices in integrals or not, a crucial point is to ask for which values $\{p_{ij}\}$ does

$$\kappa < m.$$

It is this feature that characterizes the contractive (averaging) nature of the stochastic matrices and it is equivalent to

$$\min_{i,j} \sum_s \min\{p_{is}, p_{js}\} > \underline{\kappa} > 0$$

for some constant $\underline{\kappa}$ that will denote this lower bound or the corresponding extension in (2.7). The latter condition is true for any m-stochastic matrix P that possesses a strictly positive column. These matrices are called *scrambling* and they lie in the core of the analysis of non-homogeneous discrete Markov Chains [27, 28].

A note on the scrambling index The properties of stochastic matrices and their products play a crucial role in the stability analysis held in the subsequent chapters as they serve as an appropriate tool to model the various imposed communication schemes. A standard approach for the characterization of these properties is through graph theory: Any non-negative (and in particular stochastic) matrix P can be represented as a graph $\mathbb{G}_P$ with its adjacency matrix A_P the elements of which satisfy the property $A_{ij} = 1 \Leftrightarrow P_{ij} \neq 0$. For two stochastic matrices P_1 and P_2, we write $P_1 \sim P_2$ if $\mathbb{G}_{P_1} = \mathbb{G}_{P_2}$ (consequently $P_1 = P_2$). This way we can study P from the point of view of graph theory and use the terminology of § 2.2.

A non-negative matrix P is called *irreducible* if $\mathbb{G}_P$ consists of a single essential class and a stochastic matrix P is called *regular* if $\mathbb{G}_P$ is routed-out branching.

A classical result in the theory of products of stochastic matrices is that for a regular matrix P there is a power of it that makes it scrambling: i.e.

$$\exists\ \gamma \geq 1 : \kappa(P^\gamma) < 1$$

so that from the sub-multiplicative property $P^n \to \mathbb{1}\mathbb{1}^T c$ for some $c \in \mathbb{R}$, as $n \to \infty$. The power of P that makes it scrambling is known as the *scrambling index* and the aforementioned statement on the asymptotic behavior of P^t is the ergodic theorem of stochastic matrices [26]. As the product of stochastic matrices is stochastic as well, the preceding notions can be extended to study the behavior of the non-homogeneous products of stochastic matrices. We exclusively study *backward products* of stochastic matrices defined as

$$P_{p,h} := P_{p+h}P_{p+h-1}\cdots P_{p+1} = [p_{ij}^{p,h}]. \tag{2.8}$$

for $p \geq 0, h \geq 1$.

We recall now the set $\mathcal{S}$ and its subset $\mathcal{T}$. Let $R = R_N$ denote the cardinality of $\mathcal{T}$. Each member $\mathbb{G}_i$ of it has a scrambling index γ_i. In fact, $\mathcal{T}$ can be partitioned in such mutually disjoint subsets: $\mathcal{T} = \bigsqcup_v \mathcal{Y}_v$ so that for $\mathbb{G}_1 \in \mathcal{Y}_{z_1}$, $\mathbb{G}_2 \in \mathcal{Y}_{z_2}$, $z_1 \neq z_2$ if and only if $\gamma_{z_1} \neq \gamma_{z_2}$. Consequently, we can enumerate

$$1 = \gamma_0 < \gamma_1 < \cdots < \gamma_{\max} \leq \left\lceil \frac{N}{2} \right\rceil$$

For instance, $\mathcal{Y}_0$ is the subclass of routed-out branching graphs, each member $\mathbb{G}_{\mathcal{Y}_0}$ of which has scrambling index, $\gamma_0 = 1$, i.e. there exist i such that $[\mathbb{G}_{\mathcal{Y}_0}]_{ji} \in E_{\mathbb{G}_{\mathcal{Y}_0}}$. Next we note that for any $\mathbb{G}_1, \mathbb{G}_2 \in \mathcal{T}$ with $\mathbb{G}_2$ being a sub-graph of $\mathbb{G}_1$, it holds that $\gamma_1 \leq \gamma_2$, and thus we understand that by adding an edge to any graph, the scrambling index will certainly not increase. In particular, there exists a sufficient number of new edges that will decrease the scrambling index. Fix $j < i$. Then for any $\mathbb{G}_i \in \mathcal{Y}_i$ there exists a positive number $l_{i,j}$ such that the graph $\mathbb{G}_j$ formed out of $\mathbb{G}_i$ with $l_{i,j}$ additional edges will be a member of $\bigcup_{v=0}^{j} \mathcal{Y}_v$, in which case $\gamma_j \leq \gamma_i - 1$.

Remark 2.3.7. The minimum number of edges needed to be added on an arbitrary member of $\mathcal{Y}_i$ so that the resulting graph is a member of $\bigcup_{v=0}^{i-1} \mathcal{Y}_v$, denoted by $l^* := \max_i\{l_{i,i-1}\}$.

The latter quantity is very important for the stability analysis of network consensus systems, and characterizes the convergence rate estimates to be established. For more on the dynamics of products of non-negative matrices the reader is referred to [26, 27, 28].

2.4 Dynamical Systems Theory

Let $(\mathbb{X}, \mathcal{B}, \mu)$ be a finite measure space (that is $\mu(\mathbb{X}) < \infty$) and actually, without loss of generality, we will take $\mu(\mathbb{X}) = 1$. We define a measurable transformation $T : \mathbb{X} \to \mathbb{X}$, as a map with the property that $T^{-1}(\mathcal{B}) \subset \mathcal{B}$. $T : \mathbb{X} \to \mathbb{X}$ is *measure preserving* if $\mu(T^{-1}B) = \mu(B)$ for any $B \in \mathcal{B}$. A measure preserving transformation is called *ergodic* if for any $B \in \mathcal{B}$ with the property that $T^{-1}B = B$ either $\mu(B) = 0$ or $\mu(B) = 1$.

For a collection of probability spaces, $\left\{(\mathbb{X}_n, \mathcal{B}_n, \mu_n)\right\}_{n\geq 0}$, we define the *product* probability space in the natural way: $\mathbb{X} = \prod_{n\geq 0} \mathbb{X}_n$ and a point $\chi \in \mathbb{X}$ is considered to be the sequence $\chi = \chi_0\chi_1\chi_2\ldots$ where $\chi_n \in \mathbb{X}_n$. The σ-algebra $\mathcal{B}(\mathbb{X})$ generated by subsets of $\mathbb{X}$ is the product of σ-algebras $\mathcal{B}_i$ and it is defined as the intersection of all σ-algebras that contain the collection of subsets of $\mathbb{X}$:

$$\begin{aligned} \mathcal{J} &= \left\{ \prod_{j\leq n_1-1} \mathbb{X}_j \times \prod_{n_1\leq j\leq n_2} A_j \times \prod_{j\geq n_2+1} \mathbb{X}_j \right\} \\ &= \left\{\chi \in \mathbb{X} : \chi_j \in A_j, j \in [n_1, n_2]\right\}_{0\leq n_1\leq n_2}^{A_j\in\mathcal{B}_j} \end{aligned}$$

each of which is a *measurable rectangle* (or a *cylinder*). On each of the above rectangles we attach the value $\prod_{n=n_1}^{n_2} \mu_t(A_t)$ and this can be extended to a probability measure μ on $(\mathbb{X}, \mathcal{B})$ in the standard way [30], concluding the definition of the product probability space $(\mathbb{X}, \mathcal{B}, \mu)$. A measurable transformation $T : \mathbb{X} \to \mathbb{X}$ on the product space, known as *shift*, is defined by

$$T(\chi_0\chi_1\chi_2 \ldots) = \chi_1\chi_2 \ldots$$

and it may attain all the desired properties of measure preserving and ergodicity. By $T^n\chi$ we mean the element $\chi_n\chi_{n+1}\ldots$ and we will also use the projection map $\{T^n\chi\} = \chi_n$, $\chi_n \in \mathbb{X}_n$.

For more on dynamical systems and ergodic theory we refer to [29, 30].

2.5 Fixed Point Theory

An arbitrary metric space is defined axiomatically.

Definition 2.5.1. A pair $(\mathbb{M}, \rho)$ is a *metric space* if $\mathbb{M}$ is a set and $\rho : \mathbb{M} \times \mathbb{M} \to [0, \infty)$ such that when x, y, z are in $\mathbb{M}$ then

(i.) $\rho(x, z) \geq 0$, $\rho(y, y) = 0$ and $\rho(y, z) = 0$ implies $y = z$.

(ii.) $\rho(y, z) = \rho(z, y)$ and

(iii.) $\rho(y, z) \leq \rho(y, x) + \rho(x, z)$

A metric space is *complete* if every Cauchy sequence in $\mathbb{M}$ converges in $\mathbb{M}$. A set L in a metric space $(\mathbb{M}, \rho)$ is *compact* if each sequence $\{x_n\} \subset L$ has a subsequence with limit in L.

Definition 2.5.2. Let U be an interval on $\mathbb{R}$ and $\{f_n\}$ be a sequence of functions with $f_n : U \to \mathbb{R}^N$.

(a) $\{f_n\}$ is *uniformly bounded on* U if there exists $M > 0$ such that $|f_n(t)| \leq M$ for all n and all $t \in U$.

(b) $\{f_n\}$ is *equi-continuous* if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $t_1, t_2 \in U$ and $|t_1 - t_2| < \delta$ imply $|f_n(t_1) - f_n(t_2)| < \varepsilon$ for all n.

The definition provides the standard method for proving compactness according to the following result that can be found in an text on real analysis.

Theorem 2.5.3. *If $\{f_n(t)\}$ is a uniformly bounded and equi-continuous sequence of real functions on an interval $[a, b]$ then there is a sub-sequence which converges uniformly on $[a, b]$ to a continuous function.*

Proof. See [35]. □

Since our t-intervals are possibly infinite, we need an extension of that result. We can extend it by using a weighted norm as follows:

Theorem 2.5.4. *Let $\mathbb{R}_+ = [0, \infty)$ and let $q : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous function such that $q(t) \to 0$ as $t \to \infty$. If $\{\phi_k(t)\}$ is an equi-continuous sequence of $\mathbb{R}^N$-valued functions on $\mathbb{R}_+$ with $||\phi_k(t)|| \leq q(t)$ for $t \in \mathbb{R}_+$, then there is a sub-sequence that converges uniformly on $\mathbb{R}_+$ to a continuous function $\phi(t)$ with $||\phi(t)|| \leq q(t)$ for $t \in \mathbb{R}_+$.*

Proof. See [34]. □

Definition 2.5.5. An operator $\mathcal{Q} : \mathbb{M} \to \mathbb{M}$ is a contraction in $(\mathbb{M}, \rho)$ if for any $y_1, y_2 \in \mathbb{M}$

$$\rho(\mathcal{Q}y_1, \mathcal{Q}y_2) \leq \alpha\rho(y_1, y_2)$$

for $\alpha \in [0, 1)$.

Banach's fixed point theorem The major result of Fixed Point Theory is the following result also known as the Contraction Mapping Principle:

Theorem 2.5.6. *Let $(\mathbb{M}, \rho)$ be a complete metric space and let $\mathcal{Q} : \mathbb{M} \to \mathbb{M}$ to be a contraction. Then there exists a unique $y \in \mathbb{M}$ such that $\mathcal{Q}y = y$.*

In the present work we are interested in special types of spaces which we will now prove that they attain this desired property by a verbatim use of the definition of completeness.

Proposition 2.5.7. *For any fixed $\gamma, p, \tau > 0, k \in \mathbb{R}$ and $\boldsymbol{\psi} \in C^0([-p, \tau], \mathbb{R}^N)$, the pair $(\mathbb{M}, \rho)$*

$$\mathbb{M} = \left\{ \mathbf{y} \in \mathbb{B} : \mathbf{y} = \boldsymbol{\psi}|_{[-p,\tau]}, \ \sup_{t \geq \tau} e^{\gamma t} ||\mathbf{y}(t) - \mathbb{1}k|| < \infty \right\} \tag{2.9}$$

with

$$\rho(\mathbf{y}_1, \mathbf{y}_2) = \sup_{t \geq -p} e^{\gamma t} ||\mathbf{y}_1 - \mathbf{y}_2||, \tag{2.10}$$

constitutes a complete metric space.

Proof. It is trivial to show that ρ as defined in (2.10) is a metric function according to Definition 2.5.1. It then suffices to show that a Cauchy sequence in $\mathbb{M}$ has a limit in $\mathbb{M}$. Let $\{\phi_i\}$ be such a sequence. Then

$$||\phi_j(t) - \phi_i(t)|| \leq e^{\gamma t}||\phi_j(t) - \phi_i(t)|| \leq \rho(\phi_i, \phi_j)$$

implies that $\phi_j(t)$ is a Cauchy sequence in $(\mathbb{R}^N, ||\cdot||)$ for all t, so $\phi_j(t) \to \phi(t)$. We will show that $\phi \in \mathbb{M}$ and this is the result of the following claims. The first claim is to prove that ϕ is continuous: Given $\varepsilon > 0$ there exists Q such that $\sup_{t \geq t_0} ||\phi_i(t) - \phi_j(t)|| < \varepsilon$ for $j, i > Q$. Fix such $j > Q$ and let $i \to \infty$ from above we get $||\phi(t) - \phi_j(t)|| < \varepsilon$ for all t. So $\phi_k \rightrightarrows \phi$ and thus ϕ is continuous.[2] The second claim is that ϕ is bounded. Indeed,

$$||\phi(t)|| \leq \sup_t ||\phi(t) - \phi_j(t)|| + \sup_t ||\phi_j(t)|| < \infty$$

The third claim is that $\phi \to \Delta$: For any $\varepsilon > 0$ take $Q > 0$ such that $j > Q$ implies $||\phi(t) - \phi_j(t)|| < \frac{\varepsilon}{2}$ Fix $j > Q$ and $t > T$ such that $\sup_{t \geq T} ||\phi(t) - \phi_j(t)|| < \frac{\varepsilon}{2}$ for $j > Q$ and $||\phi_j(t) - \lim_t \phi_j(t)|| < \frac{\varepsilon}{2}$ for $t > T$. Then

$$||\phi(t) - \lim_t \phi_j(t)|| \leq ||\phi(t) - \phi_j(t)|| + ||\phi_j(t) - \lim_t \phi_j(t)|| < \varepsilon$$

Finally for any $R > 0$ pick i, j large enough so that $\rho(\phi_i, \phi_j) < R$ which implies that $||\phi_i(t) - \phi_j(t)|| \leq Re^{-\gamma t}$ and taking the limit for i, $||\phi(t) - \phi_j(t)|| \leq Re^{-\gamma t}$ for all t. Then

$$||\phi(t) - \mathbb{1}k|| \leq ||\phi_j(t) - \mathbb{1}k|| + ||\phi_j(t) - \phi(t)|| \leq (R_j + R)e^{-\gamma t}$$

so that $\sup_t e^{\gamma t}||\phi(t) - \mathbb{1}k|| < \infty$. □

Fixed Point Theory is, in fact, a collection of theorems that prove the existence and/or uniqueness of fixed points of mappings in various spaces. Beyond the

[2] the symbol $\rightrightarrows$ stands for uniform convergence.

contraction mapping principle, two important results on linear spaces are presented below as they will also be of use in this work.

Schauder's fixed point theorem In 1930 Schauder published a paper [42] that generalizes the fixed point theorem of Brower to the case of linear spaces.

Theorem 2.5.8. *Let* $\mathbb{M}$ *be a non-empty compact convex subset of a Banach space and let* $\mathcal{Q} : \mathbb{M} \to \mathbb{M}$ *be continuous. Then* $\mathcal{Q}$ *has a fixed point.*

Proof. See [42] or [36]. □

Krasnoselskii's fixed point theorem Krasnoselskii [43] studied a paper of Schauder [44] and obtained the following working hypothesis: The inversion of a perturbed differential operator yields the sum of a contraction and a compact map. Accordingly, he formulated the following fixed point theorem which is a combination of the contraction mapping principle and Schauder's fixed point theorem.

Theorem 2.5.9. *Let* $\mathbb{M}$ *be a closed, convex non-empty subset of a Banach space* $(\mathbb{B}, ||\cdot||)$*. Suppose that* $\mathcal{A}$ *and* $\mathcal{B}$ *map* $\mathbb{M}$ *into* $\mathbb{B}$ *such that*

(i) $\mathcal{A}x + \mathcal{B}y \in \mathbb{M}$, $\forall x, y \in \mathbb{M}$

(ii) $\mathcal{A}$ *is continuous and* $\mathcal{A}\mathbb{M}$ *is contained in a compact set,*

(iii) $\mathcal{B}$ *is a contraction with constant* $\alpha < 1$.

Then there is $y \in \mathbb{M}$ *with* $\mathcal{A}y + \mathcal{B}y = y$.

Proof. See [36]. □

Theorems 2.5.8 and 2.5.9 differ from Theorem 2.5.6 in many ways. A first is that the latter theorem works with complete metric spaces while Theorems 2.5.8 and 2.5.9 work with compact spaces. The second is that the Contraction Mapping Principle proves both existence and uniqueness.

In the study of stability of solutions of differential equations, it is occasionally desirable to derive estimates on the rate of convergence to a limit state. As this will persistently be the case in my thesis, stability by means of fixed point theory proves to be a particularly convenient tool. The existence of a fixed point of a solution operator in a weighted complete metric space implies that the solution (i.e. the fixed point) attains the property of convergence with the prescribed rate (weight). This weight will be a rate function as it was previously defined.

2.6 Stochastic Differential Equations

For given $t_0 \in \mathbb{R}$ and a probability space $(\Omega, \mathcal{U}, \mathbb{P})$, a collection of random variables $\{\mathbf{Y}_t : t \geq t_0\}$, each of which $\mathbf{Y}_t : \Omega \to \mathbb{R}^N$ is $\mathcal{U}$-measurable, consists a stochastic process. The σ-algebra generated by $\mathbf{Y}_t$ is the smallest sub σ-algebra of $\mathcal{U}$ to which $\mathbf{Y}_t$ is $\mathcal{U}$-measurable. Let $\mathbf{B}$ be an N-dimensional Brownian motion defined on $[t_0, \infty)$ and $\mathbf{Y}^0$ is an N-dimensional random variable independent of $\mathbf{B}(t_0)$. The σ-algebra generated by $\mathbf{Y}^0$ and the history of the Brownian motion up to (and including) time $t \geq t_0$ is

$$\mathcal{U}_t := \mathcal{U}\big(\mathbf{B}(s)|_{t_0 \leq s \leq t}, \mathbf{Y}^0\big).$$

The family $\{\mathcal{U}_t\}$ is called a *filtration* and a process $\mathbf{Y}_t$ is *adapted* to $\mathcal{U}_t$ if $\mathbf{Y}_t$ is $\mathcal{U}_t$-measurable for all $t \geq t_0$. The set $(\Omega, \mathcal{U}, \mathcal{U}_t, \mathbb{P})$ consists a complete filtered probability space. Fix $T > t_0$ and let $\mathbf{b} : \mathbb{R}^N \times [t_0, T] \to \mathbb{R}^N, B : \mathbb{R}^N \times [t_0, T] \to M^{N\times N}$ are given vector valued and matrix valued deterministic functions, respectively. Then an $\mathbb{R}^N$-valued stochastic process $\mathbf{Y}_t$ is a solution of the Itô stochastic differential equation

$$\begin{cases} d\mathbf{Y}_t = \mathbf{b}(\mathbf{Y}_t, t)dt + B(\mathbf{Y}_t, t)d\mathbf{B}_t \\ \mathbf{Y}_{t_0} = \mathbf{Y}^0 \end{cases} \qquad t_0 \leq t \leq T$$

provided:

1. $\mathbf{Y}_t$ is a $\mathcal{U}_t$-adapted process.

2. $\mathbb{E}\big[\int_{t_0}^T |b_i(\mathbf{Y}_t, t)|dt\big] < \infty$.

3. $\mathbb{E}\big[\int_{t_0}^T |B_{ij}^2(\mathbf{Y}_t, t)|dt\big] < \infty$.

4. $\forall t \in [t_0, T]$

$$\mathbf{Y}_t = \mathbf{Y}^0 + \int_{t_0}^t \mathbf{b}(\mathbf{Y}_s, s)ds + \int_{t_0}^t B(\mathbf{Y}_s, s)d\mathbf{B}_s, \qquad a.s.$$

The existence and uniqueness (in probability) of a solution to the above initial value problem is guaranteed assuming a local Lipschitz condition on $\mathbf{b}$ and B and a linear sub-growth of $|\mathbf{b}(\mathbf{x}, t)|$ and $|B(\mathbf{x}, t)|$ with respect to $\mathbf{x}$. For more on Itô calculus and explicit types of solutions in certain linear stochastic differential equations as well as in asymptotic behavior of stochastic processes the reader is referred to [31, 32, 33].

2.7 Linear Inequalities

We will draw the following result from [37], Section 22.

Theorem 2.7.1. *Let* $\mathbf{a}_i \in \mathbb{R}^m$ *and* $\alpha_i \in \mathbb{R}$ *for* $i = 1, \dots, m$ *and let* k *be an integer,* $1 \le k \le m$. *Assume that the system*

$$\mathbf{a}_i^T \boldsymbol{\xi} \le \alpha_i$$

for $i = k+1, \dots, m$ *is consistent. Then one and only one of the following alternatives hold:*

1. *There exists a vector* $\boldsymbol{\xi} \in \mathbb{R}^n$ *such that*

$$\mathbf{a}_i^T \boldsymbol{\xi} < \alpha_i, \qquad i = 1, \dots, k$$

$$\mathbf{a}_i^T \boldsymbol{\xi} \le \alpha_i, \qquad i = k+1, \dots, m$$

2. *There exist non-negative real numbers* $\zeta_i|_{i=1}^m$ *such that at least one of* $\zeta_i|_{i=1}^k$ *is not zero and*

$$\sum_{i=1}^m \zeta_i \mathbf{a}_i = 0 \qquad \text{and} \qquad \sum_{i=1}^m \zeta_i \alpha_i \le 0.$$

Chapter 3: Linear Networks

Linear consensus algorithms are perhaps the most fundamental example of a self-organized distributed system. Due to their structural simplicity, they have been particularly famous and they have attracted the attention of researchers in miscellaneous scientific communities.

In its classic version, the setup of a consensus network involves a finite number of agents $N \geq 2$, each agent $i \in [N]$ of which possesses a value of interest. Denoted as $x_i \in \mathbb{R}$, this value evolves under the following averaging schemes, expressed either in discrete or continuous-time:

$$x_i(n+1) = \sum_j a_{ij} x_j(n), \qquad \dot{x}_i(t) = \sum_j a_{ij}\big(x_j(t) - x_i(t)\big). \tag{0.1}$$

The quantities a_{ij}'s are non-negative numbers that model the influence of agent j on i and essentially characterize the interdependence of agents, the connectivity regime and eventually the process of the asymptotic alignment.

For the discrete-time model, $\sum_j a_{ij} \equiv 1$, and for the continuous-time model, $\sum_j a_{ij} \equiv 0$. These two conditions imply that the updated states occurs due to a dynamic convex averaging among the current ones.

The extensive amount of proposed frameworks, much of which is discussed below, are concerned with different versions of the weights a_{ij}. Systems of type (0.1) are also known as 1^{st} *order consensus* schemes. In what follows we conduct a thorough, yet by no means complete, review of the existing models in the literature.

The interest in distributed iterative schemes has a long history in the mathematical community [21, 26, 27, 28] laying the ground for the modern theory of Non-negative Matrices and Markov Chains.

In the control community, distributed computation over networks begins with the work of Tsitsiklis et. al. [17] where problems of asynchronous agreement and parallel computing were considered for (0.1). A theoretical framework for solving consensus problems was introduced by Olfati-Saber et al. in [45] while in their seminal paper Jadbabaie et al. [20] studied a model of asymptotic alignment proposed by Viscek et al. [19]. Both these works consider populations of autonomous agents that exchange information under the assumption of symmetric communication, i.e. $a_{ij} = a_{ji}$. While Olfati-Saber et al. followed algebraic graph theory methods [40], Jadbabaie et al. based their results on the theory of non-negative matrices and non-homogeneous Markov Chains, [28]. The novelty of these works is that it includes switching communication networks, i.e. communications weights $a_{ij}(t)$ that vary over time and may be positive or zero at each t. In [20] this switching connectivity regime asks for a connectivity condition to ensure asymptotic coordination, known as *recurrent connectivity*.

Essentially, any positive value of a_{ij} signifies the existence of connection between j and i (in the sense that j affects i e.g. by signal transmission). Real-world networked systems, however, suffer from various communication failures or creations between nodes. For example, when agents are moving, some existing connections may fail as obstacles may appear between agents or assuming proximity graphs, one agent may enter the effective region of other agents.

An appropriate abstraction is this when agents are connected through a network that changes with time due to link/node failures, packet drops etc. Such variations in topology can happen randomly and this motivates the investigation of consensus problems in a connectivity regime that suffers from stochastic uncertainty. Hatano et al. consider in [46, 25] a linear agreement problem over random information networks where the existence of an information channel between a pair of elements at each time instance is probabilistic and independent of other channels. In [47], Porfiri and Stilwell provide sufficient conditions for reaching consensus almost surely in the case of a discrete linear system, where the communication flow is given by a directed graph derived from a random graph process, independent of other time instances. Under a similar communication topology model, Tahbaz-Salehi and Jadbabaie in [48] provide necessary and sufficient conditions for almost sure convergence to consensus and in [49] they extend the applicability of their necessary and sufficient conditions to strictly stationary ergodic random graphs. In [50] Matei et al. consider the linear consensus system (0.1) under the assumption that the communication flow between agents is modeled by a randomly switching graph. The switching is determined by a homogeneous, finite-state Markov chain and each communication pattern corresponds to a state of the Markov process. Then necessary and sufficient conditions are provided to guarantee convergence to average consensus in the mean square sense and in the almost sure sense.

Motivation and Contribution .Based on the works mentioned above, we make a number of remarks that effectively constitute the contribution of the present chapter.

The non-uniform hypothesis. In their vast majority all the aforementioned works, both in the deterministic and the stochastic framework rely on a fundamental assumption: The exchange of information between any two communicating nodes occurs under established connection with a time varying weight that is, uniformly bounded away from zero. This allows the applicability of an abundance of results from linear algebra, algebraic graph theory, probability theory etc. [28, 40, 23, 25, 26, 27] towards proving asymptotic convergence. The following elementary example shows that if the uniform lower bound assumption is lifted, consensus does not occur.

Example 3.0.2. Consider the 2-D dynamical system

$$\begin{pmatrix} x(n+1) \\ y(n+1) \end{pmatrix} = \begin{pmatrix} 1-f(n) & f(n) \\ g(n) & 1-g(n) \end{pmatrix} \begin{pmatrix} x(n) \\ y(n) \end{pmatrix}$$

where $n \geq 1$, $f(n) = K_f/n^2$, $g(n) = K_g/n^2$ and $K_f, K_g < 1$. Now,

$$\begin{aligned} |x(n+1) - y(n+1)| &= (1 - f(n) - g(n))(x(n) - y(n)) \\ &= |x(0) - y(0)| \prod_{i=0}^{n} (1 - f(i)) \to C \sin(\pi \sqrt{K_f + K_g}) \end{aligned}$$

for some positive constant C according to the Euler-Wallis formula, whenever $|x(0) - y(0)| \neq 0$. So for $\sqrt{K_f + K_g} \notin \mathbb{Z}$ consensus is not achieved.

The importance of this assumption has been noted before [51] and we strenuously remark that whichever work does not explicitly state it, it should be subject to criticism. Distributed consensus systems that bear non-uniform positive weights have appeared in the literature [52, 53] and it is this condition that makes the corresponding stability problems particularly challenging.

Rate of convergence. Most works in the literature are concerned with proving convergence only. Indeed, the linearity of (0.1), together with the aforementioned uniformity assumptions imply uniform asymptotic, hence exponential convergence, [51]. However, to the best of my knowledge, there are no explicit estimates of these rate of convergence for the general, asymmetric case that can include switching communication signals. The results of Chapter 2 critically contribute in the development of a unified theory of linear consensus systems both in discrete and continuous-time with emphasis on the rate of convergence. The theory is general enough to include special connection topologies such as the well-known leader-follower model.

Necessary Conditions. They are less frequent than the sufficient ones. The stochastic frameworks often provide a convenient tool for establishing both necessary and sufficient conditions for convergence as the imposed statistical regularity is especially designed to enforce asymptotic agreement for any initial conditions or disagreement for some special initial conditions [50]. We will make a comment on necessary conditions in the non-stochastic environment. We show that convergence to a common state presupposes a perpetual diffusion of information throughout the network. This propagation is guaranteed only if a critical subset of coupling weights is non-summable.

A new framework for stochastic networks. The stochastic linear consensus models satisfy certain probabilistic laws. We develop a unified framework using measure-preserving dynamical systems and ergodic theory. We demonstrate that all these systems are, in fact, special cases of our framework. We pose the claim that the connection between the deterministic mild connectivity and its probabilistic counterpart is merely semantic.

3.1 Discrete Time Dynamics

Consider N agents with values $x_i \in \mathbb{R}$ and fix $n_0 \in \mathbb{Z}_+$. On each iteration $n \geq n_0$, agent i updates its value $x_i(n) \in \mathbb{R}$ according to

$$i \in [N] : \begin{cases} x_i(n+\eta) - x_i(n) = \eta \sum_{j \in N_i} a_{ij}(n)\big(x_j(n) - x_i(n)\big), \ n \geq n_0 \\ \\ x_i(n_0) = x_i^0 \end{cases} \tag{3.1}$$

for $\eta > 0$ fixed. Hereafter, we set $\eta = 1$. We will discuss the stability of the solution $\mathbf{x}(n) = \big(x_1(n), \ldots, x_N(n)\big)$ and our objective is to derive sufficient conditions for asymptotic convergence to consensus under different connectivity schemes and the following assumption:

Assumption 3.1.1. *The connectivity weights $a_{ij}(n) : [n_0, \infty) \to \mathbb{R}_+$ satisfy*

$$\sum_{j \in N_i} a_{ij}(n) \leq m < 1, \qquad a_{ij}(n) > 0 \Rightarrow a_{ij}(n) \geq f(n), \qquad i \neq j$$

where $f(\cdot)$ is a positive function with the property that there exists $M \in [0, \infty)$ so that $f(n) \in (0, 1-m]$ for $n \geq M$.

Remark 3.1.2. It is an easy exercise to see that if we take $m \leq \frac{N}{N+1}$ for m as defined in Assumption 3.1.1, then $f(n) \leq 1 - m$.

We are interested in the solutions of (3.1) with the property that

$$x_i(n) - x_j(n) \to 0 \;\text{ as }\; n \to \infty$$

$\forall i, j \in [N]$. The next Lemma shows that this is equivalent to

$$\mathbf{x}(n) \to \Delta \;\text{ as }\; n \to \infty.$$

Lemma 3.1.3. *Let Assumption 3.1.1 hold. Then the solution* $\mathbf{x}$ *of* (3.1) *satisfies* $x_i(n) \in W_{n_0,\mathbf{x}}$ *for all* $n \geq n_0$. *Moreover,* $S(\mathbf{x}(n)) \to 0$ *as* $n \to \infty$ *implies* $\lim_n x_i(n) = k,\ \forall i \in [N]$ *for some* $k \in W_{n_0,\mathbf{x}}$.

Proof. Since $a_{ii}(n) = 1 - \sum_j a_{ij}(n) > 0$ it is an easy exercise to see that $x_i(n) \in W_{n_0,\mathbf{x}}$. Indeed, the first time n^* and agent i^* such that $x_i^*(n^*) = \max_i x_i^0$, then $x_i^*(n+1) \leq x_i^*(n)$ and likewise for the lower limit. Consequently, the forward limit set, $\omega(\mathbf{x}^0)$, is non-empty, closed subset of $W_{n_0,\mathbf{x}}$ and it is invariant with respect to (3.1). Then if $x_i(n) - x_j(n) \to 0$, any point in the forward limit set will lie in Δ. Indeed for $\mathbf{x}^0 \in \omega$, we have $\mathbf{x}^0 \in \Delta$ as well and the solution $\mathbf{x}\big(n, n_0, \mathbf{x}^0\big)$ will be in Δ for all $n \geq n_0$. $\square$

In vector form, (3.1) reads:

$$\mathbf{x}(n+1) = P(n)\mathbf{x}(n)$$

with

$$P(n) = \begin{pmatrix} 1 - \sum_{j\neq 1} a_{1j}(n) & a_{12}(n) & \cdots & a_{1N}(n) \\ a_{21}(n) & 1 - \sum_{j\neq 2} a_{2j}(n) & \cdots & a_{2N}(n) \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}(n) & a_{N2}(n) & \cdots & 1 - \sum_{j\neq N} a_{Nj}(n) \end{pmatrix} \tag{3.2}$$

Remark 3.1.4. Under Assumption 3.1.1, $P(n)$ is a stochastic matrix with strictly positive diagonal elements. Consequently, for any $P(n_1), P(n_2)$, the product $P(n_2)P(n_1)$ is structurally similar to (3.2): It is also a stochastic matrix and it has non-zero elements the union of the non-zero ones of $P(n_1)$ and $P(n_2)$.

The analysis is conducted with the following schemes:

* *Type I*: $\exists\ B \geq 1, M \geq n_0$ such that for $n \geq M$, the collection $\{P(s)\}_{s=n}^{n+B-1}$ of stochastic matrices contains at least one scrambling matrix.

* *Type II*: $\exists\ B \geq 1, M \geq n_0$ such that for $n \geq M$, the collection $\{P(s)\}_{s=n}^{n+B-1}$ of stochastic matrices contains at least one regular matrix.

Type I connectivity is significantly stronger than Type II. Reasonably so, the proof of stability is much simpler. This will be made clear in the continuous-time case in §3.2 where this distinction allows for a completely different type of proofs. A deeper insight, in fact, reveals that Type I connectivity supports a non-decentralized communication scheme as there exists a (central) node that affects the rest of the group at the same time. As an illustrative example of Type I connectivity, one may think of a star-graph communication.

Theorem 3.1.5. *Let Assumption 3.1.1 hold. Under Type I connectivity, the solution* $\mathbf{x} = \mathbf{x}(n, n_0, \mathbf{x}^0),\ n \geq n_0$ *of* (3.1) *satisfies*

$$\mathbf{x}(n) \to \Delta \quad \text{as } n \to \infty$$

if $\sum_l f(s_l) = \infty$, *where* $\{s_l\}_{l\geq 1}$ *is the sequence for which* $P(s_l)$ *is scrambling.*

Proof. Consider the sub-sequence $l : s_l \geq M$. We will estimate $\kappa(P(s_l))$ and for this we are asked to find the two rows i, j that minimize

$$\sum_k \min\{a_{ik}(s_l), a_{jk}(s_l)\}.$$

But, by assumption, $P(s_l)$ is scrambling and this means that there exists a column k^* with strictly positive elements. Then for arbitrary i, j

$$\sum_k \min\{a_{ik}, a_{jk}\} \geq \min\{a_{ik^*}, a_{jk^*}\}$$

from which two alternatives occur:

1. $i, j \neq k^*$ and $\min\{a_{ik^*}, a_{jk^*}\} = a_{ik^*}$ so that

$$a_{ik^*} \geq f(s_l),$$

2. $i = k^*$ or $j = k^*$ and $\min\{a_{ik^*}, a_{jk^*}\} = a_{ik^*}$ so that

$$a_{jk^*} \geq a_{k^*k^*} = 1 - d_i(s_l) \geq 1 - m.$$

The structure of $P(n)$, in (3.2) and Assumption 3.1.1 implies that

$$\kappa\big(P(s_l)\big) \leq 1 - \min\big\{1 - m, f(s_l)\big\} = 1 - f(s_l).$$

For any $\mathbf{x}^0 \in \mathbb{R}^N$, the general solution of (3.1) at time n is

$$\mathbf{x}(n) = P(n-1)P(n-2)\cdots P(0)\mathbf{x}^0 = P_{-1,n}\mathbf{x}^0$$

and for $n \geq M + 1$,

$$S\big(\mathbf{x}(n)\big) \leq \kappa\big(P(n-1)\big)S\big(\mathbf{x}(n-1)\big) \leq \prod_{s=M}^{n-1} \kappa\big(P(s)\big)S\big(\mathbf{x}(M)\big) \leq \prod_{s=M}^{n-1} \kappa\big(P(s)\big)S\big(\mathbf{x}^0\big)$$

Fix $\varepsilon > 0$ and pick $\bar{n} \geq 1$ large enough so that $\sum_{l=1}^{\bar{n}-1} f(s_l) > -\ln \frac{\varepsilon}{S(\mathbf{x}^0)}$. Then for $n \geq M + (\bar{n}+1)B - 1$

$$S(\mathbf{x}(n)) \leq \prod_{l=1}^{\bar{n}-1} \big(1 - f(s_l)\big) S(\mathbf{x}^0) \leq e^{-\sum_{l=1}^{\bar{n}-1} f(s_l)} S(\mathbf{x}^0) < \varepsilon.$$

$\square$

Several special cases of interest occur. If $B = 1$ then $P(n)$ is scrambling for any $n \geq M$ and hence the non-summability of $f(\cdot)$ suffices for convergence. On the other hand if $f(n) \geq \underline{f} > 0$ then the convergence occurs exponentially fast, as we shall state below.

Type I connectivity allows for the graph $\mathbb{G}_{P(n)} = ([N], E(n))$ to be time-dependent only to the point where it is sufficiently connected. This is an unnecessarily strong condition.

On the other hand, in Type II connectivity is by no means clear that $\kappa(P(n)) < 1$. Therefore, the contraction effect of $P(n)$ is not captured by κ. It is true, however, that under certain conditions the product $P_{n,h}$, as defined in (2.8), may be scrambling. In the example to follow we illustrate a simple communication condition and the effect of non-uniform connectivity weights.

Example 3.1.6. Consider a network of 4 agents, characterized by the stochastic matrix

$$P(n) = \begin{pmatrix} 1 - a_{12}(n) & a_{12}(n) & 0 & 0 \\ a_{21}(n) & 1 - a_{21}(n) - a_{23}(n) & a_{23}(n) & 0 \\ 0 & a_{32}(n) & 1 - a_{32(n)} - a_{34}(n) & a_{34}(n) \\ 0 & 0 & a_{43}(n) & 1 - a_{43}(n) \end{pmatrix}$$

with $a_{ij}(n) \geq f(n) > 0$ and $f(n)$ a monotonically decreasing function. It is easy to check that $\kappa\big(P(n)\big) = 1$ for all n. In the static case where $B = 1$, any product of such two matrices is scrambling and in our example a straightforward calculation reveals that $P(n)P(n-1)$ is scrambling with the first and fourth row to have nonzero entries that sum up to

$$a_{12}(n)a_{23}(n-1) + a_{32}(n-1)a_{43}(n) \geq 2f^2(n)$$

so that for $n >> 1$

$$\kappa\big(P(n)P(n-1)\big) = 1 - 2f^2(n)$$

and

$$S\big(\mathbf{x}(n)\big) \leq \prod_{i=1}^{n} \kappa\big(P(i)P(i-1)\big) S\big(\mathbf{x}^0\big) \leq S\big(\mathbf{x}^0\big) e^{-2\sum_{i=1}^{[\frac{t}{2}]-1} f^2(2i-1)} \to 0$$

as $n \to \infty$, on condition that $\sum_i f^2(2i-1)$ diverges.

We now make this point more rigorous.

Lemma 3.1.7. *For fixed $p \geq 0, h \geq 1$, consider the matrix product $P_{p,h}$ for each $P(s)|_{s=p+1\ldots p+h}$ defined as in (3.2) with the property that $\kappa\big(P_{p,s}\big) = 1$ for $s \in 1 \ldots h-1$ and $\kappa\big(P_{p,h}\big) < 1$. If $a := \min_{s \in \{p+1 \ldots p+h\}} \{a_{ij}(s)\} \in [0, 1-m]$ where $\sum_j a_{ij}(s) \leq m < 1$, then*

$$\kappa\big(P_{p,h}\big) \leq 1 - a^h.$$

Proof. We will use induction on h. For $h = 2$, $P_{p,2} = P_{p+2}P_{p+1}$. Since $\kappa\big(P_{p+1}\big) = 1$ and $\kappa\big(P_{p,2}\big) < 1$ then there exists a strictly positive column of $P(p+2)P(p+1) =$

$[p_{ij}]$. Let i be this column. Then

$$p_{ii} = (1 - d_i(p+2))(1 - d_i(p+1)) + \sum_j a_{ij}(p+2)a_{ji}(p+1) \geq (1-m)^2$$

$$p_{ji} = \sum_k a_{jk}(p+2)a_{kj}(p+1) \geq \min\{(1-m)a, a^2\} = a^2$$

from the bound on a the result follows. If the statement is true for $h = l$, then for $h = l+1$, similar calculations yield the bounds $p_{ii} \geq (1-m)^{l+1},\ p_{ji} \geq a^{l+1}$ so that $\kappa\big(P_{p,l+1}\big) \leq 1 - a^{l+1}$. □

We elevate the analysis now to consider Type II connectivity. This is the mildest type of communication for convergence to consensus, reported in the literature [20] and it is also known as *recurrent connectivity* condition.

Theorem 3.1.8. *Let Assumption 3.1.1 hold. Under Type II connectivity, the solution* $\mathbf{x} = \mathbf{x}(n, n_0, \mathbf{x}^0),\ n \geq n_0$ *of* (3.1) *satisfies*

$$\mathbf{x}(n) \to \Delta \quad as \quad n \to \infty$$

if $\sum_l f^\sigma(M + l\sigma - 1) = \infty$, *where* $\sigma = l^*([N/2]+1)B$ *and* l^* *with the meaning of Remark 2.3.7.*

Proof. Type II connectivity implies that for some $M > 0$ and $B \geq 1$ it holds that $P_{n,B}$ is regular with $\gamma = \gamma_{n,B}$ or equivalently $\mathbb{G}_{P_{n,B}} \in \mathcal{T}$ and $\mathbb{G}_{P_{n+B,B}} \in \mathcal{T}$ as well. Now,

$$\gamma_{n,2B} \leq \begin{cases} \max\{\gamma_{n,B}, \gamma_{n+B,B}\} - 1, & \mathbb{G}_{P_{n,B}} \subset \mathbb{G}_{P_{n+B,B}} \text{ or } \mathbb{G}_{P_{n+B,B}} \subset \mathbb{G}_{P_{n,B}} \\ \max\{\gamma_{n,B}, \gamma_{n+B,B}\}, & o.w. \end{cases}$$

If $\mathbb{G}_{P_{n,B}}$ is not a sub-graph of $\mathbb{G}_{P_{n+B,B}}$ and $\mathbb{G}_{P_{n,B}}$ is not a sub-graph of $\mathbb{G}_{P_{n+B,B}}$ or vice-versa, it holds that $E^C_{P_{n,B}} \cap E^C_{P_{n+B,B}} \neq \emptyset$. An element of this set is the pair (i,j) such that $[\mathbb{G}_{P_{n,B}}]_{ij} = 0$ and $[\mathbb{G}_{P_{n+B,B}}]_{ij} > 0$ or vice versa. This element, however, will be a member of $E_{P_{n,2B}}$ since $[\mathbb{G}_{P_{n,2B}}]_{ij} \geq [\mathbb{G}_{P_{n+B,B}}]_{ij}[\mathbb{G}_{P_{n,B}}]_{jj} > 0$. From the discussion on the partitioning of $\mathcal{T}$ with respect to the scrambling indexes and for $l^* = \max_i\{l_{i,i-1}\}$,

$$\gamma_{n,l^*B} \leq \max\{\gamma_{n,B}, \gamma_{n+B,B}\} - 1.$$

For $n \geq M$, $P_{n,l^*([N/2]+1)B}$ will be scrambling. Set $\sigma = l^*([N/2]+1)B$ and so by Lemma 3.1.7

$$\kappa(P_{M-1,l\sigma}) < 1 - f^\sigma(M + l\sigma - 1)$$

for any $l \geq 1$. Since

$$\mathbf{x}(n) = P(n)P(n-1)\cdots P(0)\mathbf{x}^0,$$

for fixed $\varepsilon > 0$ and pick $\bar{n}$ large enough so $\sum_{l=1}^{\bar{n}} f^\sigma(M + l\sigma - 1) > -\frac{\ln\varepsilon}{S(\mathbf{x}^0)}$. Finally, for $n \geq M + (\bar{n}+1)\sigma - 1$

$$S\big(\mathbf{x}(n)\big) \leq e^{-\sum_l f^\sigma(M+l\sigma-1)} S(\mathbf{x}^0) < \varepsilon.$$

$\square$

The section is concluded with a straightforward application of Theorems 3.1.5 and 3.1.8 stated as a Corollary without proof and an illustrative example.

Corollary 3.1.9. *Let the system* (3.1) *with* $a_{ij}(n) \geq \underline{f} > 0$ *uniformly in* n. *Set* $\underline{\kappa} = \min\big\{(1-m), \underline{f}\big\} \in (0,1)$. *The following estimates hold for the solution* $\mathbf{x}$ *of* (3.1)*:*

1. *Under Assumption 3.1.1 and conditions of Theorem 3.1.5,* $\mathbf{x}$ *satisfies*

$$S(\mathbf{x}(n)) \leq \frac{S(\mathbf{x}^0)}{1-\underline{\kappa}}(1-\underline{\kappa})^{\frac{n-M}{B}}, \quad n \geq M.$$

2. *Under Assumption 3.1.1 and conditions of Theorem 3.1.5,* $\mathbf{x}$ *satisfies*

$$S(\mathbf{x}(n)) \leq \frac{S(\mathbf{x}^0)}{1-\underline{\kappa}^{\sigma}}(1-\underline{\kappa}^{\sigma})^{\frac{n-M}{\sigma B}}, \quad n \geq M.$$

Example 3.1.10. Consider the system (3.1) and its solution $\mathbf{x}$ with $f(n) \geq \omega(n^{-\alpha})$ i.e. for large n, $f(\cdot)$ dominates a function that vanishes as slow as $n^{-\alpha}$. Then under Assumption 3.1.1, $\mathbf{x}(n)$ converges to Δ, if

(i) Type I connectivity holds and $\alpha \in [0, 1]$, in view of Theorem 3.1.5, or

(ii) Type II connectivity holds and $\alpha \in [0, 1/\sigma]$, in view of Theorem 3.1.8.

Theorems 3.1.5, 3.1.8 and Corollary 3.1.9 form a set of results on discrete-time linear consensus that generalize the existing ones in the literature [20]. In case of uniformly bounded weights all past results are recovered in a more concise manner. The case of static simple connectivity calls for a time invariant scrambling index γ and Theorem 3.1.8 holds for $\sigma = \gamma$. Finally, in the vanishing communication topology, we remark the interdependence between the connectivity regime and the rate at which the connections are allowed to vanish.

Discrete-time dynamics will not be reconsidered but in two circumstances. The first is in §3.3 where the analysis of the effect of stochastic uncertainty in the connectivity regime is held. The second occasion is in §6.1.1 where a 2^{nd} order consensus (flocking) network is analyzed. The latter case is primarily utilized to make the case that discretization of such systems often leads to instabilities.

3.2 Continuous Time Dynamics

The linear continuous-time model reads:

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_j a_{ij}(t)\big(x_j(t) - x_i(t)\big), \ t \geq t_0 \\ \\ x_i(t_0) = x_i^0 \end{cases} \tag{3.3}$$

We recall that $\dot{x}_i = \frac{d}{dt}x_i$ denotes the right Dini time-derivative of x_i. The solution $\mathbf{x}(t) = \mathbf{x}(t, t_0, \mathbf{x}^0)$ is an absolutely continuous function that is defined in $[t_0, \infty)$ and takes values in $\mathbb{R}^N$. The reason for considering $\mathbf{x}$ to be absolutely continuous is that we want to include the case of switching connectivity weights, i.e. the ability of $a_{ij}(t)$ to jump from a non-zero value to a zero in an abrupt manner. As already remarked, such discontinuities in real world applications can model for instance, a technical failure of connection in a communication system and although they prevent x_i from being continuously differentiable, they do not affect either their existence or uniqueness and, most importantly, they play no role on the integral representation of the solutions [54]. The next conditions is the partial continuous-time alternative of Assumption 3.1.1 and it is instrumental for the analysis to follow:

Assumption 3.2.1. *The connectivity weights a_{ij} are upper bounded, right continuous, non-negative functions of time.*

This, although hardly an assumption, together with $N < \infty$ implies that

$$\sup_{t \geq t_0} \max_i \sum_j a_{ij}(t) < \infty.$$

For the purposes of the analysis set

$$m > \sup_{t \geq t_0} \max_i \sum_j a_{ij}(t) \tag{3.4}$$

Recall now the matrix representation of the graph $\mathbb{G}_{P(t)}$ in terms of the degree matrix $D = D(t)$ and the adjacency matrix $A = A(t)$. Then

$$W(t) := mI_{N\times N} - D(t) + A(t) \tag{3.5}$$

is m-stochastic. We begin the discussion with two elementary yet crucial remarks:

Lemma 3.2.2. *Under Assumption 3.2.1 the solution* $\mathbf{x} = \mathbf{x}(t, t_0, \mathbf{x}^0),\ t \geq t_0$ *of* (3.3) *satisfies* $x_i(t) \in W_{t_0,\mathbf{x}},\ \forall\ t \geq t_0, i \in [N]$.

Proof. Let t^* be the first time that $x_i(t^*)$ escapes to the right of $\max_i x_i^0$. Then it should hold that $x_i(t^*) \geq x_j(t^*)$ for all $j \neq i$ and $\dot{x}_i(t^*) > 0$. This is an unfeasible condition from the systems equations (3.3). Similar argumentation can be made for the lower bound $\min_i x_i^0$. □

Lemma 3.2.3. *If* $\mathbf{x} = \mathbf{x}(t, t_0, \mathbf{x}^0),\ t \geq t_0$ *is the solution of* (3.3) *such that* $S\big(\mathbf{x}(t)\big) \to 0$ *as* $t \to \infty$ *then the forward limit set* $\omega(\mathbf{x}^0)$ *is a singleton with a point in* Δ.

Proof. From Lemma 3.2.2 we have that $\omega(\mathbf{x}^0)$ is non-empty, compact and connected and any element of which must lie in Δ. Take $\mathbf{x}^\omega \in \omega(\mathbf{x}^0)$ and consider the solution $\mathbf{x}(t, t_0, \mathbf{x}^\omega)$. Since $\mathbf{x}^\omega \in \Delta$ as well, we have that $\dot{\mathbf{x}} \equiv 0$, i.e. the solution is constant. □

Just like the discrete-time case, the communication signals are classified in Type I and Type II with exactly the same meaning. In Type I, $P(t)$ is scrambling on the "average" and we have a very simple and elegant proof of the consensus problem.

Theorem 3.2.4. *Let Assumption 3.2.1 hold. If*

$$f(t) := \min_{i,j} \sum_s \min\{a_{is}(t), a_{js}(t)\}$$

satisfies

$$\int^{\infty} f(s)\, ds = \infty$$

then the solution $\mathbf{x} = \mathbf{x}(t, t_0, \mathbf{x}^0),\ t \geq t_0$ *of* (3.3) *satisfies*

$$\mathbf{x}(t) \to \Delta \quad as \quad t \to \infty.$$

Proof. We write (3.3) in vector form

$$\dot{\mathbf{x}} = -D(t)\mathbf{x} + A(t)\mathbf{x} = -m\mathbf{x} + \big(mI_{N\times N} - D(t) + A(t)\big)\mathbf{x} = -m\mathbf{x} + W(t)\mathbf{x}$$

equivalently

$$\frac{d}{dt}(e^{mt}\mathbf{x}) = e^{mt}W(t)\mathbf{x}.$$

From Theorem 2.3.1 we obtain the bound

$$S\left(\frac{d}{dt}\big(e^{mt}\mathbf{x}(t)\big)\right) \leq e^{mt}\big(m - f(t)\big)S\big(\mathbf{x}(t)\big)$$

then

$$\begin{aligned}
\frac{d}{dt}S\big(\mathbf{x}(t)\big) &= -me^{-mt}S\big(e^{mt}\mathbf{x}(t)\big) + e^{-mt}\frac{d}{dt}S\big(e^{mt}\mathbf{x}(t)\big) \\
&\leq -mS\big(\mathbf{x}(t)\big) + e^{-mt}S\left(\frac{d}{dt}\big(e^{mt}\mathbf{x}(t)\big)\right) \\
&\leq -mS\big(\mathbf{x}(t)\big) + (m - f(t))S\big(\mathbf{x}(t)\big) \\
&\leq -f(t)S\big(\mathbf{x}(t)\big)
\end{aligned}$$

which implies

$$S\big(\mathbf{x}(t)\big) \le e^{-\int_{t_0}^{t} f(s)\,ds} S\big(\mathbf{x}^0\big)$$

and the result follows in view of the condition on $f(\cdot)$ and Lemma 3.2.3. $\square$

This is a generalization of the results obtained in [55] concerning continuous-time consensus algorithms. On condition that there is always an agent $i = i(t) \in [N]$ that affects every other agent j in the group it then suffices for

$$\int^{\infty} f(s)\,ds = \infty$$

This is a Type I connectivity. The non-integrability condition on f is the continuous-time counterpart of the non-summability of $f(\cdot)$ imposed on Theorem 3.1.5. We escalate the analysis now, with the study of the dynamics of (3.3) under the recurrent connectivity condition. Define for $B \ge 0$, $s \in [t-B, t]$

$$C(t,s) = e^{-mB}\delta(s-(t-B))I_{N\times N} + e^{-m(t-s)}W(s)$$

with $\delta(\cdot)$ the delta function, m as in (3.4) and $W(s)$ as in (3.5).

Proposition 3.2.5. *Let Assumption 3.2.1 hold. For any $B > 0$, $l \ge 1$, the matrix*

$$P_B^{(l)}(t) := \begin{cases} \int_{t-B}^{t} C(t,s_1)\,ds_1, \;\; l = 1 \\ \\ \int_{t-B}^{t} C(t,s_1)P_B^{(l-1)}(s_1)\,ds_1, \;\; l > 1 \end{cases}$$

is stochastic.

Proof. The matrix

$$P_B(t) := \int_{t-B}^{t} \Big(e^{-mB}\delta(s_1-(t-B))I_{N\times N} + e^{-m(t-s_1)}W(s_1)\Big)\,ds_1$$

is stochastic. Indeed, the i^{th} row of $P_B(t)$ consists of the positive diagonal element

$$e^{-mB} + \int_{t-B}^{t} e^{-m(t-s_1)}(m - d_i(s_1))\, ds_1 = 1 - \int_{t-B}^{t} e^{-m(t-s_1)} d_i(s_1)\, ds_1$$

and the non-negative off-diagonal elements

$$\int_{t-B}^{t} e^{-m(t-s_1)} a_{ij}(s_1)\, ds_1.$$

since $d_i(s_1) = \sum_j a_{ij}(s_1)$, $P_B(t)$ is stochastic. We proceed with induction:

For $l = 2$,

$$\begin{aligned}
P_B^{(2)}(t) = \int_{t-B}^{t}\int_{s_1-B}^{s_1} &\left(e^{-mB}\delta(s_1-(t-B))I_{N\times N} + e^{-m(t-s_1)}W(s_1)\right)\cdot \\
&\cdot\left(e^{-mB}\delta(s_2-(s_1-B))I_{N\times N} + e^{-m(s_1-s_2)}W(s_2)\right) ds_2 ds_1 \\
= \int_{t-B}^{t}\int_{s_1-B}^{s_1} & e^{-2mR}\delta(s_1-(t-B))\delta(s_2-(s_1-B))\, ds_2 ds_1 I_{N\times N} + \\
+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} & e^{-mR}\delta(s_1-(t-B))e^{-m(s_1-s_2)}W(s_2)\, ds_2 ds_1 + \\
+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} & e^{-m(t-s_1)}W(s_1)e^{-mR}\delta(s_2-(s_1-B))\, ds_2 ds_1 + \\
+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} & e^{-m(t-s_1)}W(s_1)e^{-m(s_1-s_2)}W(s_2)\, ds_2 ds_1
\end{aligned}$$

and straightforward calculations yield

$$\begin{aligned}
P_B^{(2)}(t) = e^{-2mB}I_{N\times N} &+ \int_{t-2B}^{t-B} e^{-m(t-s_2)}W(s_2)\, ds_2 \\
&+ e^{-mB}\int_{t-B}^{t} e^{-m(t-s_1)}W(s_1)\, ds_1 + \\
&+ \int_{t-B}^{t}\int_{s_1-B}^{s_1} e^{-m(t-s_2)}W(s_1)W(s_2)\, ds_2 ds_1
\end{aligned}$$

Now, every element of $P_B^{(2)}(t)$ is non-negative as a sum of non-negative matrices. It is only left to verify that $\sum_j [P_B^{(2)}(t)]_{ij} = 1$ for any i. Indeed, the first matrix contributes e^{-2mB}, the second and the third $e^{-mB} - e^{-2mB}$ and the fourth $(1-e^{-mB})^2$, so eventually

$$e^{-2mB} + 2(e^{-mB} - e^{-2mB}) + (1 - 2e^{-mB} + e^{-2mB}) = 1$$

Let $P_B^{(l)}(t)$ be stochastic. Then the elements of $P_B^{(l+1)}(t)$ are non-negative by the same reasoning as above and finally, since

$$\begin{aligned}
&P_B^{(l+1)}(t) = \int_{t-B}^{t} C(t, s_1) \cdots \int_{s_l - B}^{s_l} C(s_l, s_{l-1})\, ds \\
&= e^{-mB} P_B^{(l)}(t) + (1 - e^{-mB}) P_B^{(l)}(t) - \\
&- \int_{t-B}^{t} \int_{s_1-B}^{s_1} \cdots \int_{s_l-B}^{s_l} C(t, s_1) C(s_1, s_2) \cdots \big(D(s_{l+1}) - A(s_{l+1})\big)\, ds_{l+1} \ldots d_{s_1} \\
&= P_B^{(l)}(t) - \\
&- \int_{t-B}^{t} \int_{s_1-B}^{s_1} \cdots \int_{s_l-B}^{s_l} C(t, s_1) C(s_1, s_2) \cdots \big(D(s_{l+1}) - A(s_{l+1})\big)\, ds_{l+1} \ldots d_{s_1}
\end{aligned}$$

the sum of the i^{th} row of $P_B^{(l+1)}(t)$ equals 1 because the corresponding sum in the final integrand is zero (as it is a left multiplication of a matrix with a Laplacian matrix). □

Remark 3.2.6. It can be easily seen that whereas $P_B^{(l)}(t)$ is stochastic, the collection of matrices

$$\int_{s_1 \in I_t} \int_{s_2 \in I_{s_1}} \cdots \int_{s_l \in I_{s_{l-1}}} C(t, s_1) C(s_1, s_2) \ldots C(s_l, s_{l-1})\, ds_l \ldots ds_1$$

for $I_{s_l} \subset I_{s_{l-1}}$ and $I_t \in [t, t-B]$ is a member of $\mathcal{M}$ as it was defined in § 2.3.

Assumption 3.2.7. *$\exists B > 0, M \geq t_0$ such that $\forall\ t \geq M$, $\mathbb{G}_{P_B(t)} \in \mathcal{T}$.*

In addition to the upper boundedness of $a_{ij}(t)$, the weights are also assumed to satisfy the dwelling time condition [20]. This ensures that in any subset of $\mathbb{R}_+$ with positive and bounded measure, the number of discontinuities must be finite.

Assumption 3.2.8. *For any $t \geq t_0$ there exists $\epsilon > 0$ independent of t such that $a_{ij}(t) \neq 0 \Rightarrow a_{ij}(s) \geq f(s)$ for $s \in I_\epsilon(t^*) = [t^* - \epsilon, t^* + \epsilon]$ for some $t^* \in \mathbb{R}$ and $t \in I_\epsilon(t^*)$.*

Theorem 3.2.9. *Let Assumptions 3.2.1, 3.2.7 and 3.2.8 hold. The solution $\mathbf{x} = \mathbf{x}(t, t_0, \mathbf{x}^0)$, $t \geq t_0$ of (3.3) satisfies*

$$\mathbf{x}(t) \to \Delta \ \ as \ \ t \to \infty$$

if $\exists \ \{t_n\}_{n\geq 1}, t_n \geq M$ with $t_{n+1} - t_n \geq \sigma B$, such that $\sum_n f^\sigma(t_n) = \infty$, with $f(\cdot)$ as defined in Assumption 3.2.8, $\sigma = l^*([N/2] + 1)$ and l^* with the meaning of Remark 2.3.7.

If, in addition, the connectivity is static (with time-varying weights), then $\sigma = \gamma$, where γ is the scrambling index of $\mathbb{G}_{P_B(t)}$ and $B > 0$ can be chosen arbitrarily small.

Proof. The solution $\mathbf{x}$ of (3.3) satisfies

$$\begin{aligned}
&\dot{\mathbf{x}} = -m\mathbf{x} + \big(mI - D(t) + A(t)\big)\mathbf{x} \Rightarrow \frac{d}{dt}\big(e^{mt}\mathbf{x}\big) = e^{mt}W(t)\mathbf{x} \\
&e^{mt}\mathbf{x}(t) - e^{m(t-B)}\mathbf{x}(t - B) = \int_{t-B}^{t} e^{ms}W(s)\mathbf{x}(s)\, ds
\end{aligned}$$

By Assumption 3.2.7, $P_B(t) = \int_{t-B}^{t} C(t, s_0)\, ds_0$ is routed-out branching, for any $t \geq t_0 + B$. Then

$$\begin{aligned}
\mathbf{x}(t) &= \int_{t-B}^{t} \Big(e^{-mB}\delta(s - (t - B))I + e^{-m(t-s)}W(s)\Big)\mathbf{x}(s)\, ds \\
&= \int_{t-B}^{t} C(t, s_1)\mathbf{x}(s_1)\, ds_1 \\
&= \int_{t-B}^{t}\int_{s_1-B}^{s_1} \cdots \int_{s_{\sigma-1}-B}^{s_{\sigma-1}} C(t, s_1)C(s_1, s_2)\cdots C(s_{\sigma-1}, s_\sigma)\mathbf{x}(s_\sigma)\, ds_\sigma \cdots ds_1
\end{aligned}$$

for $t \geq t_0 + \sigma B$. Choosing σ large enough, $P_B^{(\sigma)}(t)$ should be scrambling. At this point the proof is identical as this of Theorem 3.1.8, thus It can be shown that for $\sigma = l^*([N/2]+1)$ the matrix $P_B^{(\sigma)}(t)$ is scrambling. In view of Remark 3.2.6, Theorem 2.3.4 applies and by Lemma 3.2.2 and Proposition 3.2.5 we have

$$S\big(\mathbf{x}(t)\big) \leq \kappa\big(P_B^{(\sigma)}(t)\big) S\big(\mathbf{x}\big(t - \sigma B\big)\big)$$

for $\sigma = l^*([N/2]+1)$ and $\kappa\big(P_B^{(\sigma)}(t)\big) < 1$ on the assumption of static connectivity. The next step is to estimate $\kappa\big(P_B^{(\sigma)}(t)\big)$. Since $P_B^{(\sigma)}(t)$ is scrambling, there exists $j^* \in V$ such that $[P_B^{(\sigma)}(t)]_{j^*i} > 0$. By direct calculations we have:

$$\begin{aligned}
&[P_B^{(\sigma)}(t)]_{j^*j^*} \geq \\
&\int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_{\sigma-1}-B}^{s_{\sigma-1}}\prod_{k=1}^{\sigma}\Big(e^{-mB}\delta\big(s_k-(s_{k-1}-B)\big)+ \\
&\qquad\qquad\qquad + e^{-m(s_{k-1}-s_k)}\big(m-d_i(s_k)\big)\Big)\,ds_\sigma\cdots ds_1 \\
&> e^{-\sigma m B}
\end{aligned}$$

and for $i \neq j^*$

$$\begin{aligned}
[P_B^{(\sigma)}(t)]_{j^*i} &\geq \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots \\
&\quad \cdots\int_{s_{\sigma-1}-B}^{s_{\sigma-1}}\sum_{l_0,\ldots,l_{\sigma-1}} e^{-m(t-s_\sigma)}a_{il_0}(s_1)a_{l_0l_1}(s_2)\ldots a_{l_{\sigma-1}j^*}(s_\sigma)\,ds_\sigma\cdots ds_1 \\
&> \int_{t-B}^{t}\int_{s_1-B}^{s_1}\cdots\int_{s_{\sigma-1}-B}^{s_{\sigma-1}} e^{-m(t-s_\sigma)}\,ds_\sigma\cdots ds_1 f^\sigma(t) = \frac{(1-e^{-m\epsilon})^\sigma}{m^\sigma}f^\sigma(t)
\end{aligned}$$

where f and $\epsilon > 0$ have the meaning of Assumption 3.2.8. For $t' \geq M$ large enough so that $f(t) \leq \frac{me^{-mB}}{1-e^{-m\epsilon}}$ whenever $t \geq t'$, from the definition of κ we obtain the estimate:

$$\kappa\big(P_B^{(\sigma)}(t)\big) \leq 1 - \frac{(1-e^{-m\epsilon})^\sigma}{m^\sigma}f^\sigma(t) \tag{3.6}$$

Then, for the aforementioned sequence $\{t_n\}$, for any $t \geq t'$, there exists $\bar{n}$ such that $t \in [t_{\bar{n}}, t_{\bar{n}+1}]$ so that

$$S\big(\mathbf{x}(t)\big) \leq S\big(\mathbf{x}(t_{\bar{n}})\big) \leq \left(1 - \frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}} f^{\sigma}(t_{\bar{n}})\right) S\big(\mathbf{x}(t_{\bar{n}} - \sigma B)\big)$$
$$\leq \left(1 - \frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}} f^{\sigma}(t_{\bar{n}})\right) S\big(\mathbf{x}(t_{\bar{n}-1})\big)$$

For any $\varepsilon > 0$, pick n_1 and n_2 large enough so that $t_{n_1} \geq t'$ and $\sum_{j=n_1}^{n_2} f(t_j) \geq \big[\frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}}\big]^{-1} \log(\frac{\varepsilon}{S(\mathbf{x}^0)})$. Then for $t \geq t_{n_1}$

$$S\big(\mathbf{x}(t)\big) \leq \prod_{k=i_1}^{i_2} \left(1 - \frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}} f^{\sigma}(t_k)\right) S\big(\mathbf{x}^0\big)$$
$$\leq e^{-\frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}} \sum_{k=i_1}^{i_2} f^{\sigma}(t_k)} S\big(\mathbf{x}^0\big) \leq \varepsilon.$$

□

The rate of convergence is dictated by the non-summability of $\sum_n f^{\sigma}(t_n)$ and in the general case implies convergence to consensus in a non-uniform sense. Just as in the discrete case an elegant result is obtained if we assume $f(t)$ to be uniformly lower bounded.

Corollary 3.2.10. *Let the conditions of Theorem 3.2.9 hold with $M = t_0$ and B, ϵ, σ the same parameters defined in its statement. If $f(t)$ as defined in Assumption 3.2.8 satisfies $f(t) \geq \underline{f} > 0$, then*

$$S\big(\mathbf{x}(t)\big) \leq \frac{S(\mathbf{x}^0)}{1-\underline{\kappa}} e^{-\theta(t-t_0)}$$

for $\theta = -\frac{\ln(1-\underline{\kappa})}{\sigma B}$, $\underline{\kappa} = \min\big\{e^{-\sigma m B}, \big(\frac{(1-e^{-m\epsilon})\underline{f}}{m}\big)^{\sigma}\big\} \in (0,1)$, m as in (3.4).

Corollary 3.2.10 is a direct application of Theorem 3.2.9 and its proof is omitted. We remark that Theorem 3.2.4, 3.2.9 and Corollary 3.2.10 unify and extend previous results [20, 51, 55]. Examples and simulations that illustrate the results of this section are postponed until Chapter 4.

3.2.1 A leader-follower scheme

In a consensus network, a *leader* is considered to be an agent that only affects the rest of the group, yet it cannot be affected by it. Non-negative matrix theory assures that the corresponding graph $\mathbb{G}$ of such a network can be routed-out branching if there is at most one leader [28] (which is actually the root of the graph). Take, without loss of generality, the leader to be agent number 1. Then the network dynamics can be written as

$$i \in [N] : \begin{cases} \dot{z}_1(t) = g\big(t, z_1(t)\big), & t \geq t_0 \\ \dot{z}_i(t) = \sum_j a_{ij}(t)\big(z_j(t) - z_i(t)\big),\ i \neq 1, & t \geq t_0 \\ z_i(t_0) = z_i^0, & t = t_0 \end{cases} \tag{3.7}$$

The dynamics of the leader's state $z_1(t)$ are assumed to evolve free of any interaction with the rest of the group under the assumption that they eventually converge to a constant value. Thus, we assume, without loss of generality, that the following condition is true:

$$|z_1(t) - k| \leq \frac{1}{h(t)} \tag{3.8}$$

for some appropriate rate function h. Then the long-run behavior of (3.7) is associated with this of (3.3) via a stability in variation argument. The idea is for all

$i \in [N]$ such that $a_{i1} \neq 0$ to write

$$\begin{aligned}\dot{z}_i(t) &= \sum_{j \neq 1} a_{ij}(t)\big(z_j(t) - z_i(t)\big) + a_{i1}(t)\big(z_1(t) - z_i(t)\big) \\ &= \sum_{j \neq 1} a_{ij}(t)\big(z_j(t) - z_i(t)\big) + a_{i1}(t)\big(k - z_i(t)\big) + a_{i1}(t)\big(z_1(t) - k\big)\end{aligned}$$

so that we can introduce a consensus system with a virtual leader of constant state z_∞, the perturbation of which will produce $\mathbf{z}$. Convergence is then guaranteed if there exists sufficient connectivity among the network of the virtual leader.

Theorem 3.2.11. *Let the solution $\mathbf{z} = \mathbf{z}(t, t_0, \mathbf{z}^0)$ $t \geq t_0$ of* (3.7) *and the dynamics of the leader together with condition* (3.8) *to hold. Assume uniform lower bounds of the connectivity weights and the connectivity conditions of Corollary 3.2.10. Let $\theta > 0$ to characterize the rate of convergence of* (3.3) *with leader. Assume that there exists a rate function c with the properties*

1. $\sup_{t \geq t_0} e^{-\theta(t-t_0)} c(t) < \infty$ *and*

2. $\sup_{t \geq t_0} c(t) \int_{t_0}^{t} \frac{e^{-\theta(t-s)}}{1-\kappa} \max_i \frac{a_{i1}(s)}{h(s)}\, ds < \infty.$

Then there exists $K > 0$ such that

$$||\mathbf{z}(t) - \mathbb{1}k||_\infty \leq \frac{K}{c(t)}.$$

Proof. The proof relies on an elementary variation argument. Equation (3.3) can be written in vector form

$$\dot{\mathbf{x}} = -L(t)\mathbf{x}$$

The dynamics of $\mathbf{x}$ are governed by the fundamental matrix $\Phi(t, s)$ such that $\mathbf{x}(t, t_0) = \Phi(t, t_0)\mathbf{x}^0$. Theorem 3.3 and the existence of leader implies that $\Phi(t, s)$ is also endowed with the stability property:

$$\big|\big|\big(\Phi(t,s) - \mathbb{1}(1,0,\ldots,0)\big)\mathbf{x}^0\big|\big|_\infty \leq \frac{S(\mathbf{x}^0)}{1-\kappa}e^{-\theta(t-s)},$$

a bound obtained by the equivalence of norms in $\mathbb{R}^N$. Based on these remarks we will study the dynamics of

$$\dot{\mathbf{z}} = -L(t)\mathbf{z} + \boldsymbol{\eta}(t) \tag{3.9}$$

where $\boldsymbol{\eta}(t) = \big(\eta_1(t), \ldots, \eta_N(t)\big)$ with

$$\eta_i(t) = \begin{cases} a_{i1}(t)\big(z_1(t) - k\big), & 1 \text{ affects } i \\ 0, & o.w. \end{cases}$$

Now, $\boldsymbol{\eta}$ is a state independent perturbation which vanishes as fast as $\frac{1}{h(t)}$. Using the linear variation of constants formula, the solution $\mathbf{z}$ satisfies

$$\mathbf{z}(t, t_0, \mathbf{z}^0) = \mathbf{x}(t, t_0, \mathbf{z}^0) + \int_{t_0}^{t} \Phi(t,s)\boldsymbol{\eta}(s)\, ds.$$

Now since $\Phi(t,s)$ projects any vector to Δ, exponentially fast with the common element, the first component of the vector but $\boldsymbol{\eta}(s)$ is by construction orthogonal to $(1,0,\ldots,0)$, we see that

$$\int_{t_0}^{t} U(t,s)\boldsymbol{\eta}(s)\, ds = \int_{t_0}^{t} \Big(\Phi(t,s) - \mathbb{1}(1,0,\ldots,0)\Big)\boldsymbol{\eta}(s)\, ds$$

also $\big|\big|\mathbf{x}(t,t_0,\mathbf{z}^0) - \mathbb{1}k\big|\big|_\infty \leq \frac{S(\mathbf{z}^0)}{1-\kappa}e^{-\theta(t-t_0)}$ so $\mathbf{z}$ satisfies

$$\mathbf{z}(t) - \mathbb{1}k = \big(\mathbf{x}(t,t_0,\mathbf{z}^0) - \mathbb{1}k\big) + \int_{t_0}^{t} \Big(\Phi(t,s) - \mathbb{1}(1,0,\ldots,0)\Big)\boldsymbol{\eta}(s)\, ds$$

and this implies

$$||\mathbf{z}(t) - \mathbb{1}k||_\infty \leq \frac{S(\mathbf{z}^0)}{1-\kappa}e^{-\theta(t-t_0)} + \int_{t_0}^{t} \frac{e^{-\theta(t-s)}}{1-\kappa} \max_i \frac{a_{i1}(s)}{h(s)}\, ds$$

In view of the imposed conditions and (3.8)

$$\sup_{t\geq t_0} c(t)||\mathbf{z}(t)-\mathbb{1}k||_\infty < \infty$$

and the proof is concluded. $\square$

3.2.2 Symmetric coupling rates

Let us take a digression and consider (3.3) with the additional condition that

$$a_{ij}(t) = a_{ij}(t), \quad \forall j \in N_i,\ t \geq t_0.$$

The analysis of this system is then significantly simplified if we consider the metric

$$\Lambda(\mathbf{x}) = \frac{1}{2}\sum_{j<i:j\in N_i}(x_i - x_j)^2 = \mathbf{x}^T\mathbf{x}$$

for $\mathbf{x} \notin \Delta$. $\Lambda(\cdot)$ is a smooth function and the standard derivative can be used to show that along the solution $\mathbf{x}(t)$ of (3.3)

$$\frac{d}{dt}\Lambda\big(\mathbf{x}(t)\big) = \sum_j (x_i(t) - x_j(t))\big(\dot{x}_i(t) - \dot{x}_j(t)\big) = -\mathbf{x}^T L\mathbf{x}$$

From §2.2.1.1 we have that

$$\begin{aligned}\mathbf{x}^T L\mathbf{x} = \sum_{a_{ij}(t)\neq 0} a_{ij}(t)(x_i - x_j)^2 &\geq \min_{a_{ij}\neq 0} a_{ij}(t) \sum_{a_{ij}(t)\neq 0}(x_i - x_j)^2 \\ &\geq \min_{a_{ij}\neq 0} a_{ij}(t)\lambda \cdot \mathbf{x}^T\mathbf{x}\end{aligned}$$

where λ is the Fiedler number of the simply connected topological graph $\mathbf{G}$. Finally,

$$\frac{d}{dt}\Lambda\big(\mathbf{x}(t)\big) \leq -f(t)\lambda\Lambda\big(\mathbf{x}(t)\big)$$

so that

$$\Lambda\big(\mathbf{x}(t)\big) \leq e^{-\lambda\int_{t_0}^t f(s)\,ds}\Lambda(\mathbf{x}^0).$$

This result highlights the discrepancy between the contraction coefficient used so far and the Fiedler number. The symmetry of the weights favors the use of the smooth metric Λ and stronger results are provided. In the simple connectivity case, while Theorem 3.2.9 effectively asks for

$$\int^{\infty} f^{\gamma}(s)\, ds = \infty$$

the Fiedler estimate asks for

$$\int^{\infty} f(s)\, ds = \infty.$$

The latter condition is what the contraction coefficient delivers only under the Type I connectivity assumption. Moreover note that $\frac{\mathbb{1}^T}{N}\mathbf{x}$ is an integral of motion and therefore the consensus point is $\frac{\mathbb{1}^T}{N}\mathbf{x}^0$. That's why the symmetry hypothesis is also known as *average consensus*. This, in turn, is a special case of the family of networks with connectivity weights that define Laplacian matrices $L(t)$ under a common left eigenvector, $\mathbf{c}$, of the zero eigenvalue, i.e. $\mathbf{c}^T L(t) \equiv 0$. Then the integral of motion of this system is $\mathbf{c}^T\mathbf{x}(t)$ and the limit point is $\mathbf{c}^T\mathbf{x}^0$.

3.2.3 A note on necessary conditions

The discussion has so far exclusively revolved around sufficient conditions for asymptotic convergence to Δ. Analysis on necessary conditions usually requires tedious arguments and are, therefore, more rare. Here we comment on this aspect of the consensus problem. For the sake of the argument we adopt symmetrical weights $a_{ij}(t) = a_{ji}(t)$ so that $a_{ij}(t) \neq 0 \Rightarrow a_{ij}(t) \geq \underline{f} > 0$ with a simple static connectivity regime. This way we avoid the, previously noted, unnecessary discrepancy of the results while at the same time can ensure a convenient exponential behavior for the "sufficiently" connected part of the network.

Theorem 3.2.12. *Consider the system* (3.3) *and its solution* $\mathbf{x} = \mathbf{x}(t, t_0, \mathbf{x}^0),\ t \geq t_0$ *with Assumption 3.2.1 to hold and the communication graph to be simply connected. Assume that over a population of N autonomous agents there is a cut of $[N]$ such that $[N] = \mathcal{V}_1 \sqcup \mathcal{V}_2$ so that for any $(i,j) \in \mathcal{V}_1 \times \mathcal{V}_2$ or $(j,i) \in \mathcal{V}_1 \times \mathcal{V}_2$, $a_{ij}(t) > 0$ implies $\int^{\infty} a_{ij}(s)\, ds < \infty$. If for any $l_1, l_2 \in [N]$, $x_{l_1}(t) - x_{l_2}(t) \to 0$ implies $|x_{l_1}(t) - x_{l_2}(t)| \leq e^{-\theta t} S(\mathbf{x}^0)$ then there exist initial conditions such that $S\big(\mathbf{x}(t)\big) > 0$ for any $t \geq t_0$*

Proof. Let the initial conditions be set such $x_l(0) < x_n(0)$ for $l \in \mathcal{V}_1$ and $n \in \mathcal{V}_2$. Consider then the subset $\mathcal{V}_{11}$ of $\mathcal{V}_1$ and accordingly the subset $\mathcal{V}_{22}$ of $\mathcal{V}_2$ which by assumption they must have connections between them. Let $i \in \mathcal{V}_{11}$ such that $x_i(t) \leq x_{i^*}(t)$ for any $i^* \in \mathcal{V}_{11}$ and $j \in \mathcal{V}_{22}$ such that $x_j(t) \geq x_{j^*}(t)$ for any $j^* \in \mathcal{V}_{22}$.

$$\dot{x}_i(t) \leq d_{ij}\big(x_j(t) - x_i(t)\big) + z_i(t)$$

$$\dot{x}_j(t) \geq d_{ji}\big(x_i(t) - x_j(t)\big) + z_j(t)$$

where $z_i(t) = \sum_{l \in \mathcal{V}_1} a_{il}\big(x_l(t) - x_i(t)\big), z_j(t) = \sum_{l \in \mathcal{V}_2} a_{jl}\big(x_l(t) - x_j(t)\big)$ are functions that signify the interconnections among agents on the separated subsets. Now,

$$\frac{d}{dt}\big(x_j(t) - x_i(t)\big) \geq -\big(d_{ij}(t) + d_{ji}(t)\big)\big(x_j(t) - x_i(t)\big) + z_j(t) - z_i(t)$$

if either $z_i(t)$ or $z_j(t)$ do not vanish then $S(\mathbf{x}(t))$ will not converge to zero and there is nothing to prove. On the other hand we have by assumption that $|z_i(t) - z_j(t)| \leq$

$2(N-1)Ce^{-\gamma t}S(\mathbf{x}^0)$ for $(N-1)C$ to play the role of the uniform upper bound of $a_{ij}(t)$ according to Assumption 3.2.1. Next we set for simplicity $Q(s) = \big(d_{ij}(s) + d_{ji}(s)\big)$

Now, $\int_{t_0}^{\infty} Q(s)\,ds < \infty$ and this means that there is a sequence $\{t_n\}_{n\geq 1}$ and a constant $J_1 > 0$ such that

$$\int_0^{t_n} Q(s)\,ds \geq J_1.$$

Since

$$x_j(t) - x_i(t) \geq e^{-\int_0^t Q(s)\,ds}(x_j^0 - x_i^0) + \int_0^t e^{-\int_w^t Q(s)\,ds}\big(f_j(w) - f_i(w)\big)dw,$$

we have that

$$|x_j(t_n) - x_i(t_n)| \geq \left| e^{-J_1}|x_{ji}^0| - \int_0^{t_n} e^{-\int_w^{t_n} Q(s)\,ds} 2(N-1)Ce^{-\gamma w} dw S(\mathbf{x}^0) \right|.$$

Choosing $\left| e^{-J_1}|x_{ji}^0| - \frac{2(N-1)CS(\mathbf{x}^0)}{\theta} \right| > \epsilon$ we obtain $|x_{ij}(t)| > \epsilon$ for infinitely many t and the proof is concluded. □

3.3 Stochastic Consensus

As already mentioned in the introduction, the stochastic part in consensus systems was initially implemented for the purpose of modeling the uncertainty in the interconnection status among agents. Heuristically speaking, the statistical regularity smooths out the condition of recurrent connectivity, that was characterized by several researchers as too stringent [46]. Indeed, imposing strong and time-invariant statistical rules, the dynamics of inter-connections, the assumption of connectivity over uniformly bounded intervals of time is satisfied almost surely and it can be essentially omitted.

The purpose of this section is to re-formulate the discrete-time linear consensus problem with modeling the communication topology in a measure theoretic framework. Thus, we recall the discussion, the notations and the terminology developed in §2.4.

We claim that this setting is general enough to unify many results proposed in the literature. The idea is to consider an appropriate finite measure space and a dynamical shift T over infinite products of stochastic matrices. Asymptotic consensus depends on the probability of a particular T invariant event and the properties of the imposed measure that preserves T. If, in addition, the shift T is ergodic then any T-invariant event is of either full or zero measure. In the examples section we will review related results from the literature and show how they constitute special cases of our setup.

As in the deterministic case, it is important to separate the existence of a connection among two agents from the strength (weight) of this connection, exactly because we are under the non-uniform lower bound condition. It is this condition that plays a critical role in the analysis of the system. In fact, unless the probabilistic regime concerning the connection failures is trivial, asymptotic consensus is never guaranteed in full probability whenever the weights are free to vanish.

3.3.1 Topology driven by measure-preserving dynamical systems

For $N < \infty$ we define the discrete set $\mathbb{Y}$ of all possible (directed) connections among N nodes. The cardinality of $\mathbb{Y}$ is finite. The product measure space is

$(\mathbb{X}, \mathcal{B}, \mu) = \prod_{n\geq 0}(\mathbb{Y}, 2^{\mathbb{Y}}, m)$ for some measure function m and μ is the induced product measure as it was discussed in §2.4. The shift operator $T : \mathbb{X} \to \mathbb{X}$ is a measure preserving transformation since for any $A \in \mathcal{J}$, $\mu(T^{-1}A) = \mu(A)$ (see also [30]). We understand χ_i as an $N \times N$, $0-1$ matrix with all diagonals zero and the off-diagonal values to attain value 1 if there is a connection from the agent of the column to the agent of the row, otherwise attain the zero value, as well. Recall now (3.1) and its solution

$$\mathbf{x}(n) = P(n-1)P(n-2)\dots P(0)\mathbf{x}(0) \tag{3.10}$$

for $\mathbb{G}_{P(n)} \in \mathcal{S}$. We would want the steering force that generates the matrices $P(n)$ at every instant n, to be effectively characterized by the projection map of the shift $T : \mathbb{X} \to \mathbb{X}$. Let the family of functions $\{a_{ij}(n)\}_{i\neq j}$, so that for $n \geq 0$ and any $i \neq j \in [N]$, $a_{ij}(n) \in [f(n), a)$ for a uniform finite upper bound a and some fixed non-increasing positive function $f(\cdot)$ that vanishes as $n \to \infty$. Let the stochastic matrix

$$P(n) = \phi(\{T^n\chi\}) \tag{3.11}$$

to be defined through the following measurable function $\phi : \mathbb{X} \to \mathcal{S}$:

$$[\phi(\{T^n\chi\})]_{ij} = \begin{cases} \frac{a_{ij}(n)}{\varepsilon \sum_{l:[\{T^t x\}_{il}]=1} a} & \text{if } \left[\{T^n\chi\}\right]_{ij} = 1 \text{ and } i \neq j \\ 0 & \text{if } \left[\{T^n\chi\}\right]_{ij} = 0 \text{ and } i \neq j \end{cases}$$

for some fixed $\varepsilon > 1$ so that $[\phi(T^n\chi)_{ii}] := 1 - \sum_j \left[\phi(\{T^n\chi\})\right]_{ij} > 0$. We are interested in the the following set

$$Q_B = \left\{\chi \in \mathbb{X} : \mathbb{G}_{P_{n,B}} \text{ is routed-out branching } \forall n \geq 0\right\} \in \mathcal{B},$$

where $P_{n,B}$ is defined in (2.8). Because of the uncertainty on the connection status we chose a slightly different version on the weights. We re-scaled the connectivity weights, in order to preserve the stochastic structure for $P(n)$, regardless of the probabilistic generator that controls the existence of connections. It should be noted that $P(n)$ may not symmetric.

The setting clearly proposes that the solution $\mathbf{x}$ is a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ closely related to $(\mathbb{X}, \mathcal{B}, \mu)$. The consensus problem becomes equivalent to the convergence of $\mathbf{x}$ to Δ. We are basically interested in convergence in the almost sure sense.[1]

Consensus may be achieved if μ assigns a positive value to Q_B for some B and in particular, it is this value of $\mu(Q_B)$ the probability with which consensus occurs, in exactly the same way as in Theorem 3.1.8 and therefore we can readily state the following result without proof:

Theorem 3.3.1. *Let $(\mathbb{X}, \mathcal{B}, \mu)$ be the direct product measure space on products of stochastic matrices and $T : \mathbb{X} \to \mathbb{X}$ a shift. Consider the solution $\mathbf{x}(n)$ of (3.10) with $P(n)$ as in (3.11) and also consider the set Q_B. Then $\mathbf{x}(t)$ satisfies*

$$\mathbf{x}(n) \to \Delta, \quad n \to \infty, \quad w.p. \quad \mu(Q_B)$$

if

$$\sum_l f^\sigma(l\sigma) = \infty$$

with $\sigma = l^*([N/2] + 1)B$ and l^* to have the meaning of Remark 2.3.7.

[1]Since $\mathbb{P}\big(|\mathbf{x}(n)| \leq N \max_i |x_i(0)|\big) = 1$ we have $E[|\mathbf{x}(n)|^r <]\infty$ and from these two facts we have that almost sure convergence implies convergence in the r^{th} mean for any $r \geq 0$.

This theorem is simply the measure theoretic analogue of Theorem 3.1.8 and little does it contribute to our discussion. It illustrates, however, the interdependence between the non-uniform lower bounds of a_{ij}, the induced statistical regularity and it is only of theoretical interest. Almost sure convergence is ensured if the event $\bigcup_{B\geq 1} P_B$ is of full measure. The most common processes in the literature (e.g. i.i.d, markov or stationary) obey probability laws that are invariant in time and they yield almost sure consensus only under the uniform bound condition (i.e. $a_{ij}(n) \neq 0 \Rightarrow a_{ij}(n) \geq \underline{f} > 0$). It is exactly this case where there is no difference between the existence of connection and its weight, when one studies the asymptotic convergence to Δ.

For this reason, in the rest of this section we will strengthen to $a_{ij}(n) \in \{0\} \cup (0, 1)$ uniformly in n so that we can focus on the processes, produced by the shift T, which guarantee the asymptotic convergence of $P_{0,n}$ to a rank-1 matrix.

Corollary 3.3.2. *Let $T : \mathbb{X} \to \mathbb{X}$ be an ergodic shift on the product space $(\mathbb{X}, \mathcal{B}, \mu)$, $P(n)$ with the form of* (3.11) *and $a_{ij}(n) \in \{0\} \cup (0, 1)$ for $i \neq j$. Then the solution $\mathbf{x}$ of* (3.10) *converges to consensus with probability one if $\mu(Q_B) > 0$ for some $B \geq 0$.*

Proof. At first we show that the set $W = \bigcup_B Q_B$ is T-invariant. Fix $B > 0$. Then for $\chi \in Q_B$ we have $T^{-1}\chi \in Q_{B+1}$.

$$T^{-1}W = T^{-1} \bigcup_B Q_B = \bigcup_B T^{-1} Q_B \subset \bigcup_B Q_{B+1} \subset W$$

Also, $\chi \in Q_B \Rightarrow T\chi \in Q_B$ so that $Q_B \subset T^{-1}Q_B$ and this is true over the union for all $B \geq 0$. Consequently

$$W \subset T^{-1}W$$

and we conclude that W is T-invariant. The ergodicity condition makes T an indecomposable transformation on T invariant sets, i.e. $\mu(W) = 0$ or $\mu(W) = 1$ but the first case is excluded because $\mu(W) \geq \mu(Q_B) > 0$. Then the only realization of shifting over $\mathbb{X}$ is this concerning processes with routed-out branching graphs over B intervals for some $B < \infty$. Any other event occurs with zero probability and the result follows in view of the uniformly bounded weights. $\square$

It should be noted here that $P(n)$ is not a stationary process. By construction the measure μ does not concern the weights $a_{ij}(n)$. The stationarity property can be observed in $\mathbb{G}_{P(n)}$ which, as we mentioned above, is the only key feature for the stability analysis.

Example 3.3.3 (Stationary Ergodic processes [49])**.** The problem of consensus over stationary ergodic processes assumes that the matrix $P(n)$ is essentially such a process. It is very well known that a measure preserving shift can be used to generate stationary processes and, conversely, that any stationary process is equal (in distribution) to a process generated by a measure preserving shift [33]. Given a stationary ergodic process that produces stochastic matrices $P(n)$ process one can easily verify whether this particular shift is ergodic after applying Birkhoffs ergodic theorem: If for some $B > 0$

$$\lim_{\bar{n}} \frac{1}{\bar{n}} \sum_{n=0}^{\bar{n}-1} \mathbf{1}_{\mathcal{T}}(P_{n,B}) > 0$$

where $\mathbf{1}_A(s)$ is a dual function that takes value 1 if $s \in A$ and 0 otherwise, $\mathcal{T}$ is the set of routed-out branching graphs, then the corresponding shift that is ergodic and consensus is proved in the almost sure sense. Corollary 3.3.2 reproduces the results of [49] but in a broader setting as not only does it allow for connectivity over B intervals of time, but it is also not concerned with the stationarity of the weighted graph. It exclusively describes the existence of a connection and not the strength of it.

Our setting may also include stationary processes that occur from deterministic systems that exhibit a non-trivial stochastic behavior, such as chaotic maps or nonlinear differential equations, so long as their solutions produce a (natural) invariant measure on the state space (see [29]). Then one can read these dynamics as stochastic and consider the consensus problem with communication topology driven by chaotic signals.

Example 3.3.4 (IID processes [46, 48])**.** One of the first works on the topic of probabilistic consensus in [46], formulated (3.1), as a stochastic linear equation with symmetric connectivity weights ($a_{ij} = a_{ji}$) to randomly take values at each time $n \in \mathbb{N}$. The partition of interest would be $a_{ij}(n) \neq 0$ with probability p and $a_{ij}(n) = 0$ with probability $1 - p$, independently of the rest of the connections and times.

Let us digress for a moment and see $P(n) = P_n(y)$ as a random process defined on a probability space $(\mathbb{Y}, 2^{\mathbb{Y}}, \mathbb{P})$. Then $P(n)$ takes values in the space of stochastic matrices with positive diagonals and uniformly bounded weights. Then the backward product $P_{n,B}(y)$ is a homogeneous sequence of independent random trails and it forms a stationary process. By the independence assumption it is easy to directly calculate the probability of the event the corresponding graph $\mathbb{G}_{P_{n,B}}$ to be routed-out branching: If p is the probability that $a_{ij}(t) \neq 0$ then the probability of j affecting i through a B time interval is by the binomial theorem $1-(1-p)^B$. For $\mathbb{G}$ a graph on N nodes, let $q \in [1, N(N-1))$ denote the minimal number of edges so that each additional edge will keep $\mathbb{G}$ routed-out branching. Then for Q_B as defined before

$$\begin{aligned}
\mathbb{P}\big(Q_B\big) &> \sum_{l=q}^{N(N-1)} \binom{N(N-1)}{l} \big(1-(1-p)^B\big)^l (1-p)^{B((N-1)^2-l)} \\
&= 1 - \sum_{l=0}^{q-1} \binom{N(N-1)}{l} \big(1-(1-p)^B\big)^l (1-p)^{B(N(N-1)-l)} \\
&= 1 - \mathcal{O}\big((1-p)^B\big) \to 1, \text{ as } B >> 1
\end{aligned}$$

To see why the event $E = \Big\{ \sup_B \mathbb{G}_{P_{n,B}} \text{is not connected}, \forall n \geq 0 \Big\}$ is a zero probability event, note that P_B are nested for B decreasing and for this reason $\mathbb{P}(E^c) = \lim_{B\to\infty} \mathbb{P}(Q_B)$.

To adapt this example to our framework we work as follows. Let the set $\{0,1\}$ and $(p, 1-p)$ the probability vector for some fixed $p \in (0,1)$, so that $\{0\}$ is assigned to $1-p$ and $\{1\}$ is assigned to p. This is an elementary measure space. On this space, we define the triplet $(\mathbb{Y}, 2^{\mathbb{Y}}, m)$ over $N(N-1)$ pairs of nodes (i.e. without self-connections) each fixed pair of which will be considered connected and take values in an open subset of $[0,1]$ with probability p or it will be zero with probability $1-p$, independently of the rest of the pairs. Eventually, $(\mathbb{X}, \mathcal{B}, \mu) = \prod_{j=0}^{\infty} (\mathbb{Y}, 2^{\mathbb{Y}}, m)$ is the product space of interest on which the shift $T : \mathbb{X} \to \mathbb{X}$ is defined, as $T(\chi_0\chi_1\chi_2\dots) = \chi_1\chi_2\dots$. If $\mathcal{J}$ is the semi-algebra of all measurable rectangles then $\mu(T^{-1}A) = \mu(A)$ for any $A \in \mathcal{J}$ and by Theorem 1.1 of [30], T is measure preserving. It is a standard exercise to show that T is ergodic [30]. It is only left to show that for some $B > 0$, $\mu(Q_B) > 0$, a calculation very similar to the one carried before and Corollary 3.3.2 applies.

Example 3.3.5 (Markov processes [50])**.** The authors considered (3.1) with a switching communication topology driven by a Markovian jump process and in particular a process on a homogeneous Markov chain over l states defined by a stochastic matrix Z, each state of which, corresponds to a connectivity regime among N nodes. The result is summarized as follows: Unconditional asymptotic consensus is achieved if and only if Z is irreducible and the union of states of the chain correspond to a routed-out branching graph. We note that the irreducibility of Z implies the existence of an invariant measure $\boldsymbol{\pi} \in \mathbb{R}^l > 0$ with $\sum_i \pi_i = 1$ with the property that $\boldsymbol{\pi}^T Z = \boldsymbol{\pi}$. In the shift oriented framework, we have a transformation T on $(\boldsymbol{\pi}, Z)$ known as Markov shift which is ergodic if and only if Z is irreducible [30]. Then the event of connectivity over a B-interval of times is dictated by the invariant measure to be of positive measure and Corollary 3.3.2 applies.

Continuous time   The results of the previous section can be modified to deal with the problem in continuous-time and there are numerous different settings to choose upon. Let us recall the deterministic case and the system (3.3). The stability of its solution with respect to Δ is decided upon the product of the matrices $P_B^{\gamma}(t)$. Just as in the discrete-time case, the stochastic nature is implemented exclusively to model the communication failure. Taking into account the necessary dwelling time condition (Assumption 3.2.8) we will use the same measure preserving shift on the same product space which will operate every $\epsilon > 0$ time: More specifically, consider continuous deterministic functions a_{ij} that satisfy Assumptions 3.2.1 and m as define (3.4). For fixed $\epsilon > 0$, $q \in \mathbb{N}$ and $t \in [t_0 + (q-1)\epsilon, t_0 + q\epsilon)$, we define the $m-$stochastic matrix $W(t) = mI_{N\times W} - D(t) + A(t)$ with

$$[W_{ij}(t)] = \begin{cases} a_{ij}(t)[\{T^q\chi\}_{ij}], & i \neq j \\ \\ m - \sum_j w_{ij}(t), & i = j \end{cases}$$

where $T : \mathbb{X} \to \mathbb{X}$ is the measurable transformation defined in the preceding section. Consequently the matrices $P_B^{(l)}(t) : \mathbb{X} \to \mathcal{S}$ from Proposition 3.2.5 are well-defined processes as measurable mappings.

Example 3.3.6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for $\omega \in \Omega$, $\mathbf{x}(t,\omega) \in \mathbb{R}^N \times \mathcal{F}$ a random process such that $\mathbf{x}(t_0) = \mathbf{x}^0$, it is sample continuous, product measurable, it has a sample right derivative and it is a solution of the stochastic differential equation

$$\dot{x}_i(t,\omega) = \sum_j a_{ij}(t,\omega)\big(x_j(t,\omega) - x_i(t,\omega)\big), \qquad x_i(t_0) = x_i^0$$

if it satisfies this equation with probability one for all $t \geq t_0$. The stochastic part of this equation lies in $a_{ij}(t,\omega)$ which are assumed to be stochastic processes generated by the shift $T : \mathbb{X} \to \mathbb{X}$ and in particular we assume to be just as the one described in the Example 3.3.4. Then we are interested in the integral

$$\int_{t_0}^{\infty} f(s,\omega)\,ds$$

where $f(t,\omega) = \min_{i,j} \sum_l \min\{a_{il}(t,\omega), a_{jl}(t,\omega)\}$, since

$$\mathbb{P}\big(\omega \in \Omega : \lim_{t\to\infty} S\big(\mathbf{x}(t,\omega)\big) \neq 0\big) = 1 - \mathbb{P}\left(\omega \in \Omega : \int_{t_0}^{\infty} f(s,\omega)\,ds = \infty\right)$$

and it can be easily calculated that at every $I_q = [t_0 + (q-1)\epsilon, t_0 + q\epsilon)$,

$$\mu\big(x \in \mathbb{X} : \mathbb{G}_{P(t)} \text{ scrambling}, t \in I_q\big) > p^{N-1} > 0$$

and by the non-summability and independence of the above events, the sum over q diverges and the Borel-Cantelli Lemma assures that $P(t)$ will be scrambling for infinitely many ϵ intervals of time. Hence $\mathbb{P}\big(\omega \in \Omega : \int_{t_0}^{\infty} f(s,\omega)\, ds = \infty\big) = 1$ and almost sure asymptotic consensus occurs.

3.4 Supplementary Remarks

An issue that needs clarification with this implementation is the seemingly mystic role of $B > 0$. Inverting the system of differential equations to a system of integral equations, we asked for a positive number $B > 0$, that will allow the solution to "run" over $[t-B, t]$ first. This is a central part of the analysis. Interestingly enough, B has its roots in the classification of states in finite state continuous-time Markov Chains. For the sake of simplicity, recall a static connectivity version of (3.3). Then as $B > 0$ is the necessary length interval of time for the contraction to act, so is the necessary time needed for the state classification into communication classes: Indeed, a pair of states i, j of the chain belongs to the same communication class if the probability starting from i to arrive in j is strictly positive for positive times, i.e. $\mathbb{P}(X_{t+B} = j | X_t = i) > 0$ for $B > 0$ (see p. 260 of [32]). In other words, $B > 0$ is the necessary time the system needs to identify the communication classes in the network.

The use of the contraction coefficient allows for asymmetric couplings but the price to pay is weak estimates; an undesirable and unavoidable situation. The symmetric connectivity achieves stronger estimates as it relies on the Fiedler number; a spectral tool that of topological (un-directed) graphs. Additionally, the dropping of the uniform hypothesis allows for nonuniform type of convergence to Δ, very important turn of events for the analysis of the 2^{nd} order consensus (flocking) networks in Chapter 6. We will see that the analysis of the linear model developed here, can serve as an appropriate base towards a unified approach in distributed co-operation and agreement dynamics.

Simulation examples of Theorems 3.2.4 and 3.2.9 are postponed for §4.4 where the effect of delays will be taken into account.

For the stochastic case we developed a general framework and we outlined the fact that the statistical regularity imposed, is in fact of little help to ease the, characterized as too demanding, recurrent connectivity regime of the deterministic case. Measure preserving ergodic dynamical systems generate precisely the statistical alternative of the recurrent connectivity. Finally, the analysis of stochastic networks was restricted to the discrete-time case for technical reasons only as we described the appropriate modifications needed for a continuous-time setting analysis.

Chapter 4: The Effect Of Delays

In many applications a central hypothesis is that the system under consideration is governed by a principle of causality; that is, future states of the system are independent of the past states and it is determined solely by the present. If, in addition, the system is governed by an equation involving both the state and the rate of change of the state then, generally, one is considering either ordinary or partial differential equations. However, under closer scrutiny, it becomes apparent that this setup is often only a first approximation to the true situation and that a more realistic model would include some of the past states of the system. The simplest type of past dependence in a differential equation is that in which the past dependence is through the state variable, the so-called retarded functional differential equations, expressed in the general form

$$\dot{x}(t) = F\big(t, x(t), x(t-\tau)\big)$$

The investigator almost always feels that delays cause difficulties. But there are striking incidents of simplifications wrought by delays. A delay can cause boundedness, stability, continuation, integrability or oscillation.

Example 4.0.1. The initial value problem

$$\begin{cases} \dot{x}(t) = x^2(t), \quad t \geq 0 \\ \\ x(0) = 1 \end{cases}$$

has a solution that escapes to infinity in finite time. But if $\tau(t)$ is positive for all $t \geq 0$ then

$$\dot{x} = x^2\big(t - \tau(t)\big)$$

has all solutions exist for all $t \geq 0$. To see this, let an initial function in $[-\tau(0), 0]$ be given and let $x(t)$ be a solution defined on $[0, T)$ with $\limsup_{t \to T^-} |x(t)| = +\infty$. Then there is a $\bar{\tau} > 0$ with $\tau(t) \geq \bar{\tau}$ on $[0, T]$ so that $t - \tau(t) \leq T - \bar{\tau}$ for $t \leq T$. But $x(t)$ is continuous on $[0, T - \bar{\tau}]$ and hence it is bounded by some number M. Thus $|\dot{x}| \leq M^2$ for $0 \leq t \leq T$ and so $|x(t) - x(0)| \leq M^2 T$, a contradiction.

For a thorough discussion on the theory of functional differential equations there is an abundance of solid textbooks [56, 34, 57, 58, 59].

4.1 Delays In Consensus Networks

Time delays are inevitable in the study of real-world complex systems. Modern research advances persistently point to the direction of understanding delays and explaining their deeper effect on the behavior of dynamical systems. The literature is steadily enriched with advances regarding the modeling of time-delay systems, the stability of their solutions as well as efficient and/or optimal control techniques in order to respond to nowadays increasing need for knowledge about of these systems from the Applied Science community [60, 61, 62, 63, 64]

A most important member of the complex systems are the networked cooperative systems with the consensus dynamics standing among the most distinguished ones. Delays in consensus networks are the result of either finite speed of information propagation between agents, known as *communication* or *propagation* delays, or finite speed of information processing, known as *input* or *processing* delay. In both cases, delays tend to weaken the performance of the system and in some cases destabilize it [45]. In consensus networks the mathematical model that incorporates delays reads

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_{j \in N_i} a_{ij}\big(x_j(t-\tau) - x_i(t-\nu)\big), & t \geq 0 \\ x_i(s) = \phi_i(s), & s \in [-\max\{\nu, \tau\}, 0] \end{cases}$$

In real-world applications it is often essential to study the stability problem not only on the level of simple convergence but also to be able to estimate the rate of convergence. From an application point of view it is desirable to obtain mathematically tractable expressions on the effect of delays in the performance of these systems. Expressions that are explicit functions of the system's parameters.

Nevertheless, there are no strong results in the literature concerning the effect of delays in distributed consensus systems. To the best of my knowledge we mention a number of relative results. A simple delayed consensus algorithm was proposed and discussed in [45] where the model

$$\dot{x}_i(t) = \sum_j a_{ij}(x_j(t-\tau) - x_i(t-\tau))$$

With $\tau > 0$ constant and uniform for all agents, a frequency method analysis was carried through. The authors used frequency methods to show that a necessary and sufficient condition for convergence is

$$\tau < \frac{2\pi}{\lambda_N}$$

where λ_N is the largest eigenvalue of the Laplacian matrix (see § 2.2.1). The problem with this method is that it does not apply if the delays are multiple and incommensurate or if the system is time-varying or even nonlinear. In [65, 17] the authors consider a discrete time version of (4.24) without processing delays and with time-varying information propagation delays $\tau(t)$. On condition that the delay is uniformly bounded from above, the strategy towards this problem is to extend the state space by adding artificial agents which played no actual role in the dynamics other than transmitting a pre-described delayed version of an agent's state. It is unclear how one would extend the state space in a continuous-time framework, unless the system is discretized. Next, in [66] the authors discuss the convergence properties of a nonlinear model which has the form

$$\dot{x}_i(t) = \sum_j a_{ij} f_{ij}\big(x_j(t-\tau) - x_i(t)\big)$$

Using passivity assumptions on f_{ij} they apply Invariance Principles to derive delay-independent convergence results. The main setback of that approach, however, is that nothing can be said for either the rate of convergence to the consensus point or the consensus point itself. Another similar argument is made in [51] for the linear consensus model with information propagation delays. The author relies on the linearity argument to conclude exponential stability from asymptotic stability.

The last family of models concerns rendezvous type of algorithms. In [67] the authors propose a second order consensus based algorithms, where agents asymptotically meet in a common place as their speed vanishes to zero. This algorithm is of the form

$$\dot{v}_i(t) = -cv_i(t) + \sum_j a_{ij}(r_j(t - \tau_j^i) - r_i(t))$$

The authors establish a Lyapunov-Krasovskii argument based on Invariance Principles and give delay-dependent results. Again, nothing can be said about the rate of convergence of this system since Invariance principles guarantee convergence.

In a very different vein, multi-dimensional systems with delays appear in the study of Monotone Dynamical Systems. In his monograph [18], Smith discusses systems of the type

$$\begin{aligned} \dot{x}_1(t) &= -ax_1(t) + ax_2(t - \tau) \\ \dot{x}_2(t) &= -bx_2(t) + bx_1(t - \tau) \end{aligned} \tag{4.1}$$

The dynamical systems involved, are categorized either as competitive or as cooperative, depending on their monotonicity (here again the sign of the parameters, a and b in the example above). Whenever a, b are positive the system is cooperative and the asymptotic behavior is a constant value for any bounded τ. Systems of the type of (4.1) are known from the control community as linear distributed agreement (consensus) dynamics and in the un-delayed case ($\tau = 0$) are treated with Algebraic Graph Theory methods [24] and form the core of Networked Control Theory [25]. Despite the abundance of results in the control community, the case of distributed delayed dynamics is treated only on the part of simple convergence results. It is still an open problem to estimate the rate of convergence to a constant value for the general case as a function of the delay.

Fixed points. The lack of strong results on the effect of delay on the rate of convergence of the consensus systems motivated me to develop a general framework based on stability by fixed point theory involving linear and nonlinear versions of consensus systems [3, 4, 5, 7, 6]. Stability by Fixed Point Theory has been recently developed with the seminal work of T.A. Burton [34]. This is a Lyapunov-free method that consists of three crucial steps. Firstly, the solution form needs to be expressed in a non-trivial way, usually with the use of a variation of parameters formula [68]. Secondly, the researcher is called upon to decide on an appropriate metric space. Having defined the solution operator based on the solution form from the variation of parameters, the third step is to derive sufficient conditions for showing using a fixed point theorem to prove that the operator has a fixed point in this metric space. This point is in our case the solution of the differential equation and it automatically inherits all the properties of the members of the metric space it is proved to exist in. If the metric space is endowed with stability properties, proving the existence of a fixed point of a solution operator in a metric space is a de-facto solution of the stability problem.

The sufficient conditions derived in the aforementioned works were proved to be too demanding. For example in the linear time-invariant system

$$i \in [N] : \begin{cases} \dot{x}_i = \sum_{j \in N_i} a_{ij}\big(x_j(t-\tau) - x_i(t)\big), & t \geq 0 \\ x_i(t) = \phi_i(t), & t \in [-\tau, 0] \end{cases} \tag{4.2}$$

This system has no processing delays but it has a uniform constant propagation delay. The derivation of the solution operator was based on a naive approximation of (4.2) with the undelayed linear time-invariant model of (2.3). This led the authors impose severe conditions on the the smallness of the allowed delay in order for the solution operator to satisfy the necessary contraction property. Such a condition is unreasonable since it is very well known that processing delays do not alter the asymptotic stability of such systems. To give a perspective, the approach asked for

$$\tau < \frac{1}{(N-1)\max_{i,j} a_{ij}}$$

The explanation that the severity of condition is only due to the variational argument is not totally convincing. Such severe conditions can also be found in fairly simple scalar differential equations of the type

$$\dot{x} = -ax + bx(t-\tau). \tag{4.3}$$

In the above scalar delayed equation the exponential stability with respect to $x = 0$ is guaranteed for any $\tau < \infty$, if $|b| < a$ and $a > 0$. At $b = a$, this condition seizes to exist.

Motivation and contribution. The contribution of this chapter concerns scalar and multidimensional linear functional differential equations. The analysis has two parts. The first part concerns the study of a class of important scalar functional equations. The second part deals with consensus networks where we apply the techniques developed for the scalar case. Our persistent goal is to prove asymptotic stability emphasizing on providing explicit estimates on the rate of convergence.

More specifically, we take a long digression studying the stability of the solutions of scalar functional equations. We will attempt to find answers on the above, seemingly trivial, peculiarity concerning the behavior when $b = a$. Our quest is leading us to the study of old models on biological growth and the large discussion in the applied mathematics community that followed and lasted for over a decade. We consider the scalar problem in two versions and adopt two different solution approaches.

1. The first is the time-invariant one: It involves constant weights and multiple delays. We develop a fixed point theory argument with the use of a combination of solution forms. We provide general and unified results and we show by example that the rate estimates are significantly improved. Its main drawback is that it is difficult to be implemented in linear time-varying or nonlinear systems. It generally fails to yield delay-independent results, despite its strong estimates on the rate of convergence.

2. The second version is the general linear one: It involves multiple time-varying weights and delays. Here, a Lyapunov-Razumikhin type of argument is developed and asymptotic stability with respect to Δ is proved and explicit estimates on the rate of convergence are stated. We provide conditions on the sign of the time-varying weights so that the convergence is delay-independent. The derived estimates recover older results from the literature in a much simpler way. The disadvantage is that these estimates are in general very weak. We will compare our results by example and simulations the two approaches.

For the second part, we apply the scalar results for the study of the multidimensional alternative. The analysis is split to the time invariant and the time-varying case, accordingly. The linear time-invariant (LTI) consensus networks are considered with multiple constant propagation and processing delays. Similar to the scalar version, a fixed point theory argument is developed on the base of a combination of two integral representations of the solution. We establish sufficient conditions for stability with emphasis on the rate of convergence. The second part concerns the development of general theory of time-varying consensus dynamics with time-varying delays along the lines of the scalar time-varying case and the use of the contraction coefficient from Chapter 3. Several examples and simulations are provided in every step in order to illustrate and compare the various results.

4.2 A Remarkable Scalar Equation

In the theory of delayed differential equations, probably the simplest example one can discuss is (4.3). There is, indeed, hardly a textbook in the field not mentioning this equation at the chapter of stability of solutions [57, 58, 59]. It is in that particular chapter where the text focuses on equations of the type (4.3) with the parameters a, b satisfying $a > 0$ and $|b| \leq a$. The zero solution is then shown to be stable for any $\tau > 0$. In the vast majority of these texts, the next sentence goes pretty much as follows:

"Moreover, if $|b| < a$, then the zero solution is asymptotically stable.[...]".

In any advanced textbook or technical paper one may find the proof that the asymptotic stability of such systems is exponential. This can be shown using any of the conventional stability analysis methods. It is the purpose of the next subsection, to briefly review these tools in the stability of (4.3) for $|b| < a$.

4.2.1 The case $|b| < a$.

Equation (4.3) is stated as the following initial value problem

$$\begin{cases} \dot{x}(t) = -ax(t) + bx(t-\tau), & t \geq 0 \\ x(t) = \phi(t), & t \in [-\tau, 0] \end{cases} \tag{4.4}$$

where $a > 0$ and $b \in \mathbb{R}$ such that $|b| < a$. It is the latter condition which guarantees that both $x(t)$ and $e^{\theta t}x(t)$ are qualitatively equal for $\theta > 0$ small enough. If $x(t)$ is a solution of (4.4) then $y(t) = e^{\theta t}x(t)$ satisfies:

$$\begin{cases} \dot{y}(t) = -(a-\theta)y(t) + be^{\theta\tau}y(t-\tau), & t \geq 0 \\ y(t) = e^{\theta t}\phi(t), & t \in [-\tau, 0] \end{cases} \tag{4.5}$$

We choose $\theta > 0$ small enough so that

$$\frac{|b|e^{\theta\tau}}{a-\theta} < 1 \tag{4.6}$$

and such θ always exist only because $|b| < a$.

We will review the main stability analysis methods for the asymptotic behavior of y with respect to the zero solution as asymptotic stability of y implies exponential stability of x with rate θ.

Frequency Methods. We choose the direct method (§2.2.3 of [58]). The quasi-polynomial of (4.5) is

$$a(s, e^{-\tau s}) = s + (a-\theta) - be^{\theta\tau}e^{-s\tau} = a_0(s) + a_1(s)e^{-s\tau}$$

with $a_0(s) = s + (a - \theta)$ and $a_1(s) = -be^{\theta\tau}$. At $\tau = 0$, $a(s, 1) = s + (a - \theta) - b$ and it is stable $a(s, 1) = 0$ if and only if $\Re(s) < 0$. Then

$$\left|\frac{a_1(j\omega)}{a_0(j\omega)}\right| = \frac{|b|e^{\theta\tau}}{\sqrt{\omega^2 + (a-\theta)^2}} < 1, \qquad \omega \in \mathbb{R}.$$

in view of (4.6). Then there is no solution of $a(s, e^{s\tau}) = 0$ with $\Re(s) > 0$ and the stability result follows.

Liapunov-Krasovskii[59]. For the problem (4.5), the appropriate functional to be selected is $V(\phi) = \frac{1}{2}\phi^2(0) + \mu \int_{-\tau}^{0} \phi^2(\theta) d\theta$ so that

$$\dot{V}(\phi) = -(a - \theta - \mu)\phi^2(0) + |b|e^{\theta\tau}\phi(0)\phi(-\tau) - \mu\phi^2(-\tau) \leq 0$$

if $\mu \in \left(\frac{|b|e^{\theta\tau}}{2}, a - \theta - \frac{|b|e^{\theta\tau}}{2}\right)$ which exists in view of (4.6). Then the result follows in view of (4.6) and Theorem 2.1 of [59].

Liapunov-Razumkhin[59]. The Lyapunov function in this case is $V(x) = \frac{x^2}{2}$. Then

$$\dot{V}(x(t)) \leq -(a - \theta)x^2(t) + |b|e^{\theta\tau}|x(t)| \cdot |x(t-\tau)| \leq -(a - \theta - |b|e^{\theta\tau})x^2(t) \leq 0$$

whenever $|x(t)| \geq |x(t - \tau)|$. Then the result follows in view of (4.6) and Theorem 4.1 of [59].

Fixed Point Theory[34]. This technique does not directly rely on the transformation $y(t) = e^{\theta t}x(t)$. The condition $|b| < a$ suffices to prove that the solution operator defined by inverting (4.4) as

$$(\mathcal{Q}x)(t) = \begin{cases} e^{-at}\phi(0) + b\int_0^t e^{-a(t-s)}x(s-\tau)\, ds, & t \geq 0 \\ \phi(t), & t \in [-\tau, 0] \end{cases} \tag{4.7}$$

is a contraction in the complete metric space $(\mathbb{M}, \rho)$ where

$$\mathbb{M} = \{x \in C^0[-\tau, \infty) : x = \phi|_{[-\tau,0]}, \ \sup_{t \geq -\tau} e^{\theta t}|x(t)| < \infty\}$$

and $\rho(x_1, x_2) = \sup_{t\geq 0} e^{\theta t}|x_1(t) - x_2(t)|$, whenever condition (4.6) holds. Then by Theorem 2.5.6, $\mathcal{Q}$ attains a unique fixed point in $\mathbb{M}$. This is a de facto proof of the asymptotic convergence of the solution to zero (see p. 41 of [34]).[1]

There is no doubt that the critical condition which characterizes (4.3) as "simple" is $|b| < a$ and consequently condition (4.6). Neither the constancy of the weights a and b nor the nature of delay (large, time/state dependent) play any substantial role in the asymptotic behavior of the solution x. So long as the magnitude of the undelayed term dominates the magnitude of the retarded term the effect of delay acts as a harmless perturbation of the the solutions in the qualitative sense. This property can be readily generalized to more complex systems and it is therefore exploited in different contexts for the establishment of delay-independent stability results [58]. The crucial condition (4.6) seizes to hold when $a = b$.

4.2.2 The case $b = a$.

At this "bifurcation" value, a number of new phenomena occur. Since condition (4.6) does not hold, one cannot use the transformation $y(t) = e^{\theta t}x(t)$. Also (4.3) reads

$$\dot{x}(t) = -ax(t) + ax(t - \tau) \tag{4.8}$$

[1]This approach does not include the step of the stability of solutions with respect to the classical definition. This part must be handled separately (usually with an $\epsilon - \delta$ argument).

and every real constant is a solution. In particular if from (4.4), $\phi \in \Delta_I$ then the solutions stay in Δ_I for all times. However, no solution in Δ_I is asymptotically stable in the classical sense: if $y_1, y_2 \in \Delta$ with $|y_1 - y_2| < \epsilon$ for any $\epsilon > 0$, with $\phi = y_1$ then $y(t) \equiv y_1$ and it will never converge to y_2. As a result, none of the above methods is applicable any more. For instance, the direct method of Section 4.2.1 gives for (4.4)

$$\frac{a}{\sqrt{\omega^2 + a^2}} \leq 1, \quad \omega \in \mathbb{R}$$

which by no means imply asymptotic stability of the solutions, let alone the convergence rate. However, we may still conclude stability. The same result occurs for the Liapunov-Krasovski and Liapunov-Razumikhin methods, whereas the mapping $\mathcal{Q}$ in the Fixed Point Theory approach seizes to be a contraction for any value of τ.

It is this class of delayed differential equations for which Invariance Principles may take over in the Liapunov-based approaches and prove asymptotic stability of the solutions with respect to the invariant set Δ [69]. These techniques however suffer from the standard drawback that they provide no information on the rate at which the solutions convergence to such an invariant subset. A feature that is of utmost importance for real-world problems.

The Cook-Yorke hypothesis. To the best of my knowledge, the first occurrence of this equation dates back in 1973 with the seminal work of Cooke and Yorke [70]. The authors developed a theory of biological growth and epidemics by introducing and analyzing the system

$$\dot{x} = g(x(t)) - g(x(t-\tau)) \tag{4.9}$$

where g is an arbitrary Lipschitzian function. The authors proved that whenever the solution of (4.9) exists in the large and does not diverge, it approaches asymptotically a constant value. Their work has ever since attracted enormous attention from the mathematical community and caused an abundance of convergence results of these types of functional differential equations, known in the literature as *equations with asymptotic constancy of solutions* [71, 72, 73, 34, 74, 75, 76, 77, 78].

A similar field of study where such equations appear is this of the motion of a classical radiating electron [79]. In the following, we will review a number of past works that emphasize both on the asymptotic stability and on the rate of convergence.

In [72, 71] the authors develop conditions which ensure that all solutions of certain functional differential equations are asymptotically constant as $t \to \infty$. Equation (4.8) is a special case of their work for which it must hold that

$$|a\tau| < 1 \tag{4.10}$$

so that solutions x tend to a constant so that $x \in L^1\big([t_0, \infty), \mathbb{R}\big)$. In [71], it is proved that the rate is exponential with exponent $\theta \in (0, -\frac{\ln(a\tau)}{\tau})$.

In his monograph [34], T.A. Burton explains that, since the solution operator of (4.8) can be expressed as

$$x(t) = -a \int_{t-\tau}^{t} x(s)\, ds + x(0) + a \int_{-\tau}^{0} \phi(s)\, ds,$$

condition (4.10) ensures that this solution form can be a contraction in the complete metric space $(\mathbb{M}, \rho)$ defined in § 4.2.1 with the origin translated to the fixed constant

$$k = \frac{\phi(0) + a \int_{-\tau}^{0} \phi(s)\, ds}{1 + a\tau} \tag{4.11}$$

and θ small enough to satisfy $|a|\frac{e^{\theta\tau}-1}{\theta} < 1$. The approach of Burton is more general in the sense that it sheds light upon the asymptotic value k as well as it can be readily applied to nonlinear versions of (4.8) such as (4.9).

Finally, Krisztin [78], developed a Liapunov-Razumikhin argument, based on the monotonicity of (4.8) (i.e. the fact that we can take $a > 0$). He estimated the rate of convergence without the condition (4.10). In particular, assuming

$$a\tau < \infty$$

he proved that the rate of convergence of solutions to a constant is exponential with rate proportional to

$$\frac{\ln(1 - e^{-a\tau})}{2\tau} \tag{4.12}$$

which is a delay-independent result. The main conclusion of the discussion should be that as far as (4.8) is concerned, the convergence to a constant value is independent of the magnitude of the delay τ only when $a > 0$ but is restricted to conditions like (4.10) whenever $a < 0$.

The fixed-point argument. Within the context of fixed point theory this peculiarity occurs as a result of the way the solutions are expressed and the corresponding solutions operators are defined.

Motivated by this shortcoming, we discuss the scalar problem expressed in (4.8). In fact, within the framework of fixed point theory we develop a new form of the solution operator on the base of the following observation: (4.10) reveals that this condition simply neglects the sign of a. Indeed, the solution x for (4.8) also converges to a constant exponentially fast if (4.10) holds. Moreover one may suspect that instability occurs at $|a\tau| = 1$ exactly because it may mean $a\tau = -1$ for which value k as defined in (4.11), is not finite. The first remark is that for the solution of (4.8) can be expressed either as

$$x(t) = -a \int_{t-\tau}^{t} x(s)\, ds + \left(\phi(0) + a \int_{-\tau}^{0} \phi(s)\, ds \right) \tag{4.13}$$

or as

$$x(t) = e^{-a(t-t_0)} x(t_0) + \int_{t_0}^{t} e^{-a(t-s)} a x(s-\tau)\, ds \tag{4.14}$$

for any t, t_0 with $t \geq t_0$. These are two forms with different information on the dynamics of x. (4.13) shows that the value of $x(t)$ exclusively depends on the information of x in $[t-\tau, t]$ and (4.14) shows that $x(t)$ is based on the information of x in $[t_0 - \tau, t_0]$ whereas this form also exploits the dissipative nature of the dynamics, due to $a > 0$. The problem with (4.14) is that, unlike (4.7), it is not a contraction in any useful metric space with regards to asymptotic stability. This is exactly because $a = b$ from (4.3). Since in the FPT framework it is the representation of the solution as of much importance as the form of the Lyapunov function in the Lyapunov theory, we will combine (4.13) and (4.14), to obtain a new representation of the solution that in many cases will yield asymptotic stability results independent of the magnitude of the delays.

Notation. Fix $N < \infty$ and $t_0 \in \mathbb{R}$. $\tau_i \in [t_0, \infty) \to \mathbb{R}_+$ is a continuous function such that $\lambda_i(t) := t - \tau_i(t)$ is non-decreasing with $\lim_{t\to\infty} \lambda_i(t) = \infty$, for any $i = 1, \ldots, N$. Whenever the subscript i is omitted we understand the maximum over i, i.e. $\tau(t) := \max_i \tau_i(t)$ whereas $\lambda(t) = t - \tau(t) = \min_i \lambda_i(t)$. Also $\lambda^{(j)}(t)$

will denote the j^{th} composition of $\lambda(t)$ and for $t \geq t_0$ $I_{\lambda^{(j-1)}t} := [\lambda^{(j)}(t), \lambda^{(j-1)}(t)]$ with the convention $I_{\lambda^{(0)}(t)} = I_t$. For any $a \in \mathbb{R}$, $a^+ := \max\{0, a\}$ and $a^- := \min\{0, a\}$. A solution of a scalar functional differential equation is denoted as $x(t, t_0, \phi)$ where ϕ is the initial function so that $x(t_0, t_0, \phi) = \phi$ as and equivalent to $x(t_0, t_0, \phi)(s) = \phi(s)$ for $s \in I_{t_0}$. In general $x_t = x(t+s)$ for $s \in I_t$. Also $W_{I_t,\phi} = \left[\min_{s\in I_t} \phi(s), \max_{s\in I_t} \phi(s)\right]$ and $S_{I_t}(\phi) = \max_{s\in I_t} \phi(s) - \min_{s\in I_t} \phi(s)$ in agreement with the notation in § 2.1 from where we recall for the rest of the symbols.

4.2.3 Time-invariant dynamics

The first set of results concerns the following initial value problem

$$\begin{cases} \dot{x}(t) = -\sum_{i=1}^{N} a_i x(t) + a_i x(\lambda_i(t)), & t \geq t_0 \\ x(t) = \phi(t), & t \in I_0 \end{cases} \tag{4.15}$$

with $a_i \in \mathbb{R}$ constant $\lambda_i(t) = t - \tau_i$ for τ_i constant and ϕ given initial datum. We re-write (4.15) as

$$\dot{x} = -ax(t) + \sum_i a_i^+ x(\lambda_i(t)) + \sum_i |a_i^-| \frac{d}{dt} \int_{\lambda_i(t)}^t x(s)\, ds$$

so as to separate the dissipation dynamics from the non-dissipation ones. Then the solution x satisfies both

$$\begin{aligned} x(t) =& e^{-a\tau} x(\lambda(t)) + \int_{\lambda(t)}^t e^{-a(t-s)} \sum_i a_i^+ x(\lambda_i(s))\, ds + \\ &+ \int_{\lambda(t)}^t e^{-a(t-s)} \sum_i |a_i^-| \frac{d}{ds} \int_{\lambda_i(s)}^s x(w)\, dw ds, \qquad t \geq \tau \end{aligned} \tag{4.16}$$

and

$$x(t) = c_0 - \sum_{i=1}^{N} a_i \int_{\lambda_i(t)}^{t} x(s)\, ds \tag{4.17}$$

with $c_0 := \phi(0) + \sum_i a_i \int_{-\tau_i}^{0} \phi(s)\, ds$. For $t \geq \tau$ we substitute $x(\lambda(t))$ of the first form with the right hand-side equivalent of the second form and we derive the following form of the solution of (4.15):

$$\begin{aligned}
x(t) &= e^{-a\tau} c_0 + \sum_i \int_{\lambda_i(\lambda(t))}^{\lambda(t)} \left(e^{-a(t-w-\tau_i)} - e^{-a\tau}\right) a_i^+ x(w)\, dw \\
&\quad + \sum_i \int_{\lambda(t)}^{\lambda_i(t)} e^{-a(t-w-\tau_i)} a_i^+ x(w)\, dw + \sum_i |a_i^-| \int_{\lambda_i(t)}^{t} x(w)\, dw \\
&\quad - \sum_i \int_{\lambda(t)}^{t} e^{-a(t-s)} a|a_i^-| \int_{\lambda_i(s)}^{s} x(w)\, dwds \\
&= e^{-a\tau} c_0 + \sum_i \int_{\lambda_i(\lambda(t))}^{\lambda(t)} \Bigg[\left(e^{-a(t-w-\tau_i)} - e^{-a\tau}\right) a_i^+ - \\
&\qquad\qquad - \int_{\lambda(t)}^{g_i(w)} e^{-a(t-s)} a|a_i^-|\, ds\Bigg] x(w)\, dw \\
&\quad + \sum_i \int_{\lambda(t)}^{\lambda_i(t)} \left[e^{-a(t-w-\tau_i)} a_i^+ - \int_{w}^{g_i(w)} e^{-a(t-s)} a|a_i^-|\, ds\right] x(w)\, dw \\
&\quad + \sum_i \int_{\lambda_i(t)}^{t} |a_i^-| \left[1 - \int_{w}^{t} a e^{-a(t-s)}\, ds\right] x(w)\, dw
\end{aligned} \tag{4.18}$$

where the last step is due to the change of the order of integration on the term $\sum_i \int_{t-\tau}^{t} e^{-a(t-s)} a|a_i^-| \int_{s-\tau_i}^{s} x(w)\, dwds$. This will be our solution operator, for $t \geq \tau$. We exploit the monotonicity of the integrand functions to arrive in the following condition

Assumption 4.2.1. *There exists $\alpha \in [0, 1)$ such that*

$$1 - e^{-a\tau} - \sum_{i=1}^{N} \left[\tau_i |a_i| e^{-a\tau} - \frac{|a_i^-|}{a}\left(1 - e^{-a\tau_i} - e^{-a\tau} + e^{-a(\tau-\tau_i)}\right)\right] \leq \alpha$$

Remark 4.2.2. We outline the following two special cases:

1. $a_i^- \equiv 0$: The condition reduces to $1 - e^{-a\tau} - \sum_i a_i \tau_i e^{-a\tau} =: \alpha < 1$ and it is satisfied for any magnitude of $\tau_i, a_i < \infty$.

2. $a_i^+ \equiv 0$: Then $a = 0$ and the condition reduces to $\sum_i |a_i \tau_i| \leq \alpha < 1$.

For $\theta \in (0, a)$ we define the quantities

$$\begin{aligned}
\Theta_1(i,\theta) &= |a_i^-| e^{a\tau_i} e^{-(a-\theta)\tau} \frac{1 - e^{-(a-\theta)\tau_i}}{a-\theta} - |a_i^-| e^{\theta\tau} \frac{e^{\theta\tau_i} - 1}{\theta} \\
\Theta_2(i,\theta) &= |a_i^-| (e^{a\tau_i} - 1) \frac{e^{-(a-\theta)\tau_i} - e^{-(a-\theta)\tau}}{a-\theta} \\
\Theta_3(i,\theta) &= |a_i^-| \frac{e^{(a-\theta)\tau_i} - 1}{a-\theta}
\end{aligned}$$

and we pick θ small enough so that

$$\sum_i a_i^+ \left(e^{\theta\tau_i} \frac{1 - e^{-a\tau}}{a-\theta} - e^{-a\tau} \frac{e^{\theta\tau_i} - 1}{\theta} \right) + \sum_{j=1}^{3} \Theta_j(i,\theta) \leq 1. \tag{4.19}$$

One can always find such θ to satisfy this condition since in the limit $\theta \downarrow 0$ (4.19) coincides with Assumption 4.2.1.

Theorem 4.2.3. *Under Assumption 4.2.1, the solution $x = x(t, 0, \phi),\ t \geq 0$ of* (4.15) *is exponentially asymptotically stable with respect to Δ. More specifically, x converges to*

$$k = \frac{\phi(0) + \sum_{i=1}^{N} a_i \int_{-\tau_i}^{0} \phi(s)\, ds}{1 + \sum_{i=1}^{N} a_i \tau_i} \tag{4.20}$$

exponentially fast with exponent θ that satisfies (4.19).

Proof. At first, we see that k is well-defined as if, $1 + \sum_{i=1}^{N} a_i \tau_i \downarrow 0$ then the left hand side of the expression in the Assumption 4.2.1 becomes greater than 1.

Next, we prove stability of the solution with respect to Δ. Fix $\epsilon > 0$ and k. We pick ϕ and $\delta_\tau = \delta_\tau(\epsilon, k) < \epsilon$ so that $|x(t,\phi) - k| < \delta_\tau$ for $t \in [-\tau, \tau]$. Such a δ_τ can always be found by the continuous dependence on initial conditions. Next we pick $\delta \leq \delta_\tau$ satisfying $\delta(1 + \sum_i a_i\tau_i)e^{-a\tau} + \alpha\epsilon < \epsilon$, consider the first time $t^* \geq \tau$ such that $|x(t^*, \phi) - k| = \epsilon$ and express x as in (4.18) to arrive in a contradiction. Finally, we prove exponential convergence to k by a fixed point argument. Consider the metric space $(\mathbb{M}, \rho)$ with

$$\mathbb{M} = \left\{ y \in C^0([-\tau, \infty), \mathbb{R}) : y = \tilde{x}|_{[-\tau,\tau]},\ \sup_{t\geq\tau} e^{\theta t}|y(t) - k| < \infty \right\}$$

and $\rho(x_1, x_2) = \sup_{t\geq\tau} e^{\theta t}|x_1(t) - x_2(t)|$. From Proposition 2.5.7, $(\mathbb{M}, \rho)$ is a complete metric space. Next, we define the operator

$$(\mathcal{P}x)(t) = \begin{cases} \tilde{x}(t), & t \in [-\tau, \tau] \\ \\ x_{(4.18)}(t), & t \geq \tau \end{cases}$$

To show that $\mathcal{P}$ is a member of $\mathbb{M}$ we observe that it is continuous and it agrees in $[-\max_i \tau_i, \tau]$ with any member of $\mathbb{M}$, by definition. Next $\sup_{t\geq\tau} e^{\theta t}|(\mathcal{P}x)(t) - k|$ is finite, for $\theta < a$. Finally, we show that $\mathcal{P}$ is a contraction in $(\mathbb{M}, \rho)$: For $x_1, x_2 \in \mathbb{M}$ we calculate an upper bound of $e^{\theta t}\big|(\mathcal{P}x_1)(t) - (\mathcal{P}x_2)(t)\big|$ and we arrive at

$$e^{\theta t}\big|(\mathcal{P}x_1)(t) - (\mathcal{P}x_2)(t)\big| \leq \Theta(t, \theta)\rho(x_1, x_2)$$

where $\sup_{t\geq\tau} \Theta(t, \theta)$ is equal to (4.19). Then Theorem 2.5.6 can be applied concluding the proof. $\square$

4.2.4 Time-varying dynamics

The natural extension of (4.15) is

$$\begin{cases} \dot{x}(t) = -\sum_{i=1}^{N} a_i(t)x(t) + a_i(t)x\big(\lambda_i(t)\big), & t \geq t_0 \\ x(t) = \phi(t), & t \in I_{t_0}. \end{cases} \tag{4.21}$$

where $a_i(t)$ are functions to be determined and $\lambda_i(t) = t - \tau_i(t)$ as defined in § 4.2.2. There are a number of reasons to focus on the dynamics of the first derivative of the state $\dot{x}$ for (4.21), rather than the study of x itself. It should be intuitively clear that there is no hope to try to express the asymptotic value k in a closed form because the time-varying components of the system supply the orbit with new information. Hence it is not possible for an integral of motion to be derived. Since x always converges to $\mathbb{R}$ whenever $\int^{\infty} \dot{x}(s)\, ds$ exists, it is desirable to seek the solution $\dot{x}$ in (4.21) in $L^1_{[t_0,\infty)}$. To outline this method, consider the simplified system

$$\dot{x}(t) = -a(t)x(t) + a(t)x(\lambda(t)).$$

We note that $\dot{x}$ satisfies both

$$\dot{x}(t) = -a(t) \int_{\lambda(t)}^{t} \dot{x}(s)\, ds$$

and

$$\ddot{x}(t) = -a(t)\dot{x}(t) + a(t)\dot{x}(\lambda(t))\dot{\lambda}(t) - \dot{a}(t) \int_{\lambda(t)}^{t} \dot{x}(s)\, ds$$

Anyone trying to apply the same approach as before, will find themselves with an integrodifferential equation. The analysis is held in [14]. Unfortunately, the results are far from satisfactory. The method yields strong delay-dependent conditions, even for $a(t) \geq 0$; it becomes too dysfunctional and it is thus abandoned. For (4.21) we follow a different approach that restricts for $a_i(t) \geq 0$.

Lemma 4.2.4. *The solution $x = x(t, t_0, \phi)$, $t \geq t_0$ of (4.21) satisfies $x(t) \in W_{I_{t_0},\phi}$ for all $t \geq t_0$.*

Proof. Let t^* be the first time that x escapes to the right. Then $x(t^*) = max_{s \in I_{t^*}} x(s)$ and $\dot{x}(t^*) > 0$. This is a contradiction in view of (4.21) and a similar argument can be made for the lower bound. □

A first crucial remark is that

$$S_{I_{\bar{t}}}(x) \leq S_{I_{\underline{t}}}(x) \tag{4.22}$$

for any $\bar{t} \geq \underline{t} \geq t_0$.

Theorem 4.2.5. *Consider the system (4.21) and assume that*

$$m = \sup_{t \geq t_0} \int_{\lambda(t)}^{t} a(s)\, ds < \infty$$

and that $\lambda(t) \to \infty$ as $t \to \infty$. Then the solution $x = x(t, t_0, \phi)$, $t \geq t_0$ of (4.21) satisfies

$$S_{I_t}(x) \leq e^{\ln(1-e^{-m}) l_t} S_{I_{t_0}}(\phi)$$

for l the maximum natural number so that $\lambda^{(l)}(t) \geq t_0$.

Proof. For any $t \geq t_0$ consider the interval $I_t = [t - \tau(t), t]$ and pick the points $t_1, t_2 \in I_t$ so that $x(t_2) = \max_{s \in I_t} x(s)$, $x(t_1) = \min_{s \in I_t} x(s)$. Moreover, assume without loss of generality that $t_2 > t_1$. Then

$$x(t_2) = e^{-\int_{t_1}^{t_2} a(s)\,ds} x(t_1) + \int_{t_1}^{t} e^{-\int_s^t a(s)\,ds} \sum_i a_i(s) x(s - \tau_i(s))\,ds$$
$$= \int_{t_1}^{t_2} e^{-\int_s^t a(s)\,ds} \sum_i a_i(s)\big(x(s - \tau_i(s)) - x(t_1)\big)\,ds + x(t_1)$$

by Lemma 4.22. Then

$$S_{I_t}(x) = x(t_2) - x(t_1) \leq \big(1 - e^{-\int_{t_1}^{t_2} a(s)\,ds}\big) S_{I_{\lambda^{(2)}(t)}}(x)$$

and the result follows from a simple recursive argument. $\square$

Note that for $a(t) \leq a$ we must have $\tau(t) \leq \tau$ so that $m = a\tau$ from Theorem 4.2.5 and

$$S_{I_t}(x) \leq \frac{S_{I_{t_0}}(\phi)}{1 - e^{-m}} e^{\frac{\ln(1-e^{-m})}{2\tau}(t - t_0)}, \tag{4.23}$$

implying exponential convergence. The above result is a simple and elegant proof of stability for (4.21) and it reaffirms the rates obtained in [78], following a much simpler way. Note that here we ask for $a_i \geq 0$ so that the crucial Lemma 4.22 holds. In [14] the same bounds are obtained via a different, much harder, way. One way to treat (4.21) for general functions $a_i(t)$ is if we separate the positive from the negative parts of a_i and make a stability in variation argument. Indeed, one can set $a_i^+(t) = \max\{0, a_i(t)\}$ and $a_i^-(t) = \min\{0, a_i(t)\}$ so that

$$\dot{x}(t) = -\sum_i a_i^+(t) x_i(t) + \sum_i a_i^+(t) x\big(\lambda_i(t)\big) - \sum_i a_i^-(t) x_i(t) + \sum_i a_i^-(t) x\big(\lambda_i(t)\big)$$

then x can be expressed as a variation of solution y that satisfies

$$\dot{y}(t) = -\sum_i a_i^+(t) y(t) + \sum_i a_i^+(t) y\big(\lambda_i(t)\big)$$

using appropriate techniques presented in [68]. The details of this method are discussed in [14] where several examples are worked to illustrate and compare the strength of the different rates obtained. Some of them are presented right below.

4.2.5 Examples and Simulations

Example 4.2.6 (Stability bounds in LTI systems)**.** Consider the initial value problem

$$\begin{cases} \dot{x}(t) = -1.2x(t) + 1.2x(t-\tau) + 2.3x(t) - 2.3x(t-\nu), & t \geq 0 \\ x(t) = \phi(t), & t \in [-\max\{\tau,\nu\},0] \end{cases}$$

We discuss the asymptotic behavior of solutions with respect to τ and ν and we compare the results of § 4.2.3 and § 4.2.4.

$\underline{\nu = 0}$: Both methods are applicable, so we will compare the estimates that Theorems 4.2.3 and 4.2.5 provide for $\tau = 1, \dots, 10$.

Theorem 4.2.3 asks for

$$\Theta(\tau,0) = 1 - e^{-1.2\tau} - 1.2\tau e^{-1.2\tau} < 1$$

which is always satisfies and thus we have delay-independent exponential stability with respect to Δ with rate θ which satisfies

$$G(\tau,\theta) = 1.2e^{\theta\tau}\frac{1-e^{-1.2\tau}}{1.2-\theta} - 1.2e^{-1.2\tau}\frac{e^{\theta\tau}-1}{\theta} \leq 1.$$

Theorem 4.2.5 calculates from (4.23)

$$|x(t) - k| \leq \frac{S_{I_0}(\phi)}{1 - e^{-1.2\tau}} e^{\frac{1}{2\tau}\ln(1-e^{-1.2\tau})t}$$

so that the estimated rate is $\theta = -\frac{1}{2\tau}\ln(1 - e^{-1.2\tau})$.

In Fig. 4.2.6 we compare the two curves

$$\theta : G(\tau, \theta) = 1$$

$$\theta = -\frac{1}{2\tau}\ln(1 - e^{-1.2\tau})$$

for $\tau = 1, \dots, 10$. We conclude that the estimates of Theorem 4.2.3 clearly outperform these of Theorem 4.2.5. Simulations for $\tau = 5$ indicate that, in general both estimates are still away from the numerically verified ones by an order of 10 (see Fig. (4.2.6)).

$\underline{\nu > 0}$: Here only Theorem 4.2.3 applies and convergence is guaranteed for small values of ν. In particular Assumption 4.4.1 asks for $F(\tau, \nu) < 1$ where

$$\Theta(\tau, \nu) := 1 - e^{-1.2\tau} - \Big[1.2\tau e^{-1.2\tau} + 2.3\nu e^{-1.2\tau} - \\ - 1.9\big(1 - e^{-1.2\nu} - e^{-1.2\tau} + e^{-1.2(\tau-\nu)}\big)\Big]$$

In Fig. (4.2.6) the function $\nu = \nu(\tau)$ is plotted for the values such that $F(\tau, \nu(\tau)) = 1$ which is the stability bounds of the system.

As a numerical application take $\tau = 1$. Fig. (4.2.6) yields stability for $\nu \leq 0.325$ so we fix $\nu = 0.32$. Then condition (4.19) of Theorem 4.2.3 yields a rate equal to $\theta = 0.175$. Numerical inspection of this condition gives for $\nu \approx 0.001$ an estimate near 0.071. Fig. (4.2.6) has the solution for $\tau = 1, \nu = 0.32$ and the estimated rate is near 0.4. Both simulations were carried through with `MATLAB` and the `ddesd`

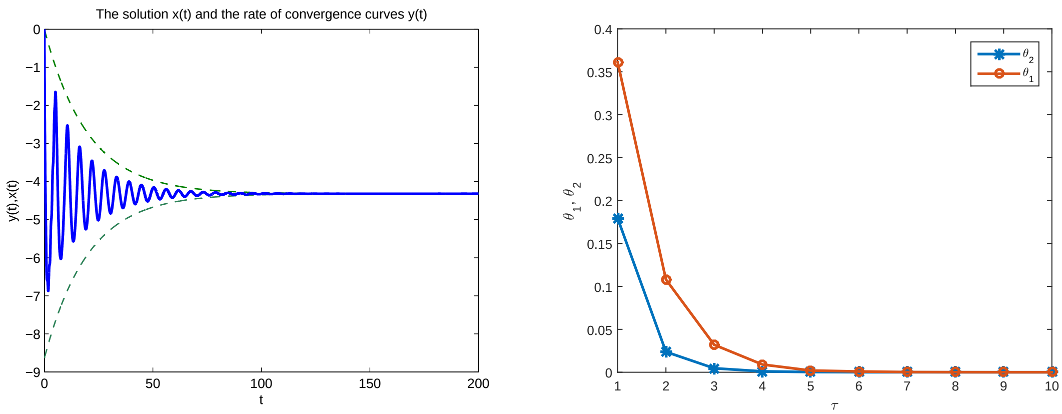

Figure 4.1: Example 1. (a): $\nu = 0, \tau = 5$ and $\phi(s) = \sin(8s) + 2s$, $s \in [-5, 0]$ and the $(t, y(t))$ curve defined by $y(t) = \pm(x(0) - k)e^{-0.05t} + k$. (b) The rate estimates between Theorem 4.2.3 (θ_1) and Theorem 4.2.5 (θ_2).

function.

Example 4.2.7 (Direct linearization of a nonlinear system)**.** In this example we will show how our results can be applied in the study of nonlinear systems

$$\begin{cases} \dot{x} = a(t)f\big(x(t-\tau(t)) - x(t)\big), t \geq t_0 \\ \\ x(s) = \phi(s), s \in I_{t_0} \end{cases}$$

where $0 \leq a(t) \leq a$, $\tau(t) \leq \tau$ and f satisfies $\frac{f(q)}{q} > 0$ for $q \neq 0$ and $|f(q)| \leq |q|$.

Claim 4.2.8. *The solutions of the system satisfy $x(t) \in W_{I_{t_0},\phi}$.*

Proof. Let the first time $t^* \geq t_0$ such that $x(t^*) = \max_{s \in I_{t_0}} \phi(s)$ with $\dot{x}(t^*) > 0$. But $\dot{x}(t^*) = a(t^*)f\big(x(t^* - \tau(t^*)) - x(t^*)\big) < 0$ by virtue of the property of f, a contradiction. Similarly for the lower bound. $\square$

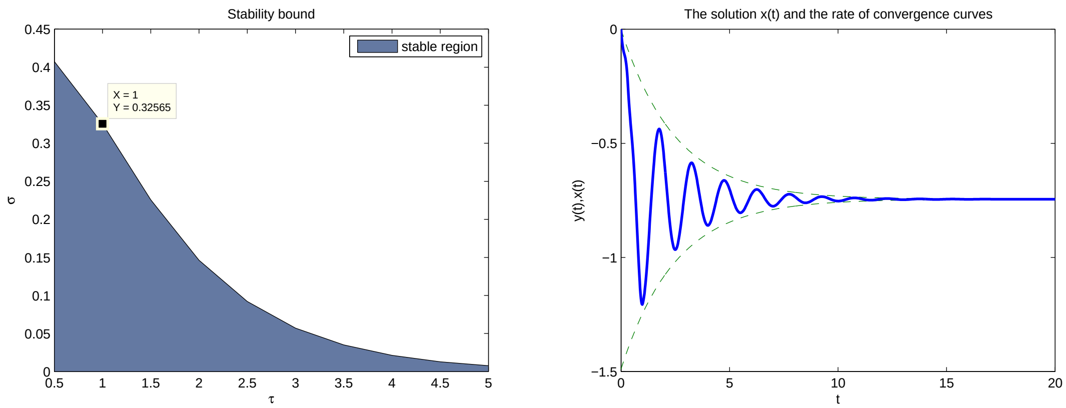

Figure 4.2: Example 1. (a) $\nu \neq 0$. and the stability curve $\nu = \nu(\tau)$. The shaded region is the region of stability i.e. $(\nu < \nu(\tau))$. (b) The solution of Example 1 with $\tau = 1, \nu = 0.32$. The rate curves $(t, y(t))$ are defined by $y(t) = \pm(x(0) - k)e^{-0.4t} + k$.

Given the unique solution x of the system $\dot{x} = f(t, x_t)$ we define the linear system

$$\dot{z} = a(t)c(t)\big(z(t - \tau(t)) - z(t)\big)$$

with $c(t) = c(x(t)) := \frac{f(x(t-\tau(t))-x(t))}{x(t-\tau(t))-x(t)} > 0$. We observe that the solution z is indistinguishable from x. From Claim 4.2.8 we conclude that $c_M := \sup_{x \in M_{t_0,\phi}} \frac{f(x)}{x} > 0$ is well defined and thus $c(t) \geq c_M$. Then Theorem 4.2.5 can be applied with (4.23) and we get the estimate:

$$|z(t) - k| \leq M_{t_0,\phi}(1 - e^{-c_M a\tau})^{-1} e^{\frac{1}{2\tau}\ln(1-e^{-c_M a\tau})(t-t_0)}.$$

We can extend this example to include negative $a(t)$ and work similarly to the case (1) if we ask f to grow sub-linearly, e.g. $f(x) = \sin(x)$.

4.3 LTI Delayed Consensus Networks

In this section, we will apply the argument of § 4.2.3 for the multidimensional case and in particular a consensus network under linear time-invariant static connectivity and multiple constant processing and propagation delays. For this, we recall the discussion held in § 2.2 with the notation used. Fix $\tau^i_j, \nu^j_i \in [0, \infty)$ with $\tau := \max_{i,j}\{\tau^i_j\}$, $\nu := \max_{i,j}\{\nu^j_i\}$ and let $I_0 = [-\max\{\tau, \nu\}, 0]$. Consider the following initial value problem:

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_j a_{ij}\big(x_j(t - \tau^i_j) - x_i(t - \nu^j_i)\big), & t \geq 0 \\ \\ x_i(t) = \phi_i(t), & t \in I_0 \end{cases} \tag{4.24}$$

where $\phi_i \in C^0(I_0, \mathbb{R})$ are given initial data. For the analysis we will make use of the following three hypotheses:

Assumption 4.3.1. *The communication graph $\mathbb{G}$ is routed-out branching.*

Assumption 4.3.2. *The system parameters $a_{ij}, \tau^i_j, \nu^j_i$ and $\mathbf{c}$ satisfy*

$$1 + \sum_{i,j=1}^{N} c_i a_{ij}\big(\tau^i_j - \nu^j_i\big) > 0$$

From Proposition 2.2.1 and Assumption 4.3.1 we can conclude that $e^{-Lt}L = (e^{-Lt} - \mathbb{1}\mathbf{c}^T)L$ is a matrix function with elements that converge to zero exponentially fast. In particular, $e^{-Lt}L = [\kappa_{ij}(t)]_{ij}$ satisfies $\sum_{j \neq i} \kappa_{ij} = -\kappa_{ii}$ and $\kappa_{ij}(t) \to 0$ in the order of $\mathcal{O}(e^{-\Re\{\lambda\}t})$. These functions are calculated from the connectivity weights and are assumed known. Define

$$g_{il}(\theta) := \sup_{t \geq \tau} \int_0^{t-\tau} |\kappa_{il}(t - \tau - s)| e^{\theta(t-s)}\, ds \tag{4.25}$$

a quantity that is well-defined for $\theta < \Re\{\lambda\}$ and

$$h_{l,j,i}(\theta) := |a_{lj} - a_{ij}|\frac{e^{\theta\tau_j^l} - 1}{\theta} + a_{ij}\frac{e^{\theta\max\{\tau_j^i,\tau_j^l\}} - e^{\theta\min\{\tau_j^i,\tau_j^l\}}}{\theta} \tag{4.26}$$

Define

$$\begin{aligned} f_{ij}(\theta) :=& a_{ij}e^{\theta\tau_j^i}\frac{1 - e^{-(d_i-\theta)\tau}}{d_i - \theta} - a_{ij}e^{-(d_i-\theta)\tau}\frac{e^{\theta\tau_j^i} - 1}{\theta} + e^{-d_i\tau}\sum_{l\neq i} h_{l,j,i}(\theta)g_{il}(\theta) + \\ &+ e^{-d_i\tau}\sum_{l} g_{ij}(\theta)a_{jl}\frac{e^{\theta\nu_j^l} - 1}{\theta} \end{aligned}$$

if $i \neq j$ and

$$\begin{aligned} -f_{ii}(\theta) :=& 1 - \sum_{l\neq i} a_{il}\frac{e^{\theta\nu_i^l} - 1}{\theta}\left(1 + d_i\frac{1 - e^{-(d_i-\theta)\tau}}{d_i - \theta}\right) - e^{-d_i\tau}\sum_{l\neq i} h_{l,i,i}(\theta)g_{il}(\theta) \\ &- e^{-d_i\tau}\sum_{l\neq i} g_{ii}(\theta)a_{il}\frac{e^{\theta\nu_i^l} - 1}{\theta}. \end{aligned}$$

Finally form the matrix $F(\theta) = [f_{ij}(\theta)]$.

Assumption 4.3.3. *There exists* $\mathbf{q} \in \mathbb{R}_+$ *so that*

$$F(0)\frac{1}{\mathbf{q}} < 0$$

where $\frac{1}{\mathbf{q}} := (1/q_1, \ldots, 1/q_N)$.

We are now ready to state the central theorem of this section:

Theorem 4.3.4. *Under Assumptions 4.3.1, 4.3.2 and 4.3.3, the solution* $\mathbf{x} = \mathbf{x}(t, 0, \boldsymbol{\phi}),\ t \geq 0$ (4.24) *converges to* $\mathbb{1}k$ *where*

$$k = \frac{\sum_{i=1}^N c_i\big[\phi_i(0) + \sum_{j\in N_i} a_{ij}\big(\int_{-\tau_j^i}^0 \phi_j(s)\,ds - \int_{-\nu_i^j}^0 \phi_i(s)\,ds\big)\big]}{1 + \sum_{i=1}^N c_i\sum_{j\in N_i} a_{ij}\big(\tau_j^i - \nu_i^j\big)} \tag{4.27}$$

exponentially fast with rate $\theta \in (0, \min\{d_i, \Re\{\lambda\}\})$ *that satisfies*

$$F(\theta)\frac{1}{\mathbf{q}} \leq 0.$$

Assumption 4.3.1 is a minimal connectivity condition which requires that there can be at most one $i^* \in [N]$ with $d_{i^*} = 0$, usually called the leader of the group, in which case for $\mathbf{c} = (c_1, \ldots, c_N)^T$ the the left eigenvector of $L = L_G$ associated with $\lambda_1 = 0$, we have $\mathbf{c} = e_{i^*}$. This means that i^* remains unaffected and all other agents' states asymptotically converge to the state of i^*. For the rest of the paper we will quietly assume that $d_i > 0, \forall i \in [N]$ without essentially effecting any of the derived results. The hypothesis on the boundedness of the delays is necessary in order to maintain exponential convergence of the system to the consensus state. This state is, in turn, analytically calculated from the initial conditions and the topological structure of the graph. Employing FPT methods in the study of the stability of the solutions of (4.24), we extract sufficient conditions for their asymptotic behavior of its solutions.

Although our analysis is in principle based on the smallness of the maximum allowed delays ν and τ (Assumption 4.3.3) in the Examples section we will consider special cases where the conditions become delay-independent.

<u>Proof of Theorem 4.3.4.</u> The proof is an application of Theorem 2.5.6 and it consists of several steps. These involve the preparation of the solution operator, the definition of an appropriate metric space, the proof that the solution operator maps this space into itself and the proof that this operator is a contraction.

Preparing the solution operator. Following § 4.2.3, we will express the solution $\mathbf{x}(t,0,\boldsymbol{\phi})$ of the system (4.24) in two different ways and we will combine both of them to create a new one which will serve as the solution operator. For $i \in [N]$ we write from (4.24):

$$\dot{x}_i = \sum_{j\in N_i} a_{ij}\big(x_j - x_i\big) - \sum_{j\in N_i} a_{ij}\frac{d}{dt}\left[\int_{t-\tau_j^i}^{t} x_j(s)\,ds - \int_{t-\nu_i^j}^{t} x_i(s)\,ds\right]$$

In vector form this equation reads

$$\dot{\mathbf{x}} = -L\mathbf{x} - \sum_{i,j} A_{ij}\frac{d}{dt}\int_{t-\tau_j^i}^{t}\mathbf{x}(s)\,ds + \sum_{i,j} B_{ij}\frac{d}{dt}\int_{t-\nu_i^j}^{t}\mathbf{x}(s)\,ds$$

where $A_{ij} = [a_{lk}\delta_{li}\delta_{kj}]$ and $B_{ij} = [a_{ij}\delta_{li}\delta_{ki}]$. Using the variation of constants and the integration by parts formulae we see that the solution of (4.24) satisfies

$$\begin{aligned}
\mathbf{x}(t) &= e^{-Lt}\boldsymbol{\phi}(0) - \int_0^t e^{-L(t-s)}\sum_{i,j}\Bigg[A_{ij}\frac{d}{ds}\int_{s-\tau_j^i}^{s}\mathbf{x}(w)\,dw - \\
&\qquad\qquad - B_{ij}\frac{d}{ds}\int_{s-\nu_i^j}^{s}\mathbf{x}(w)\,dw\Bigg]\,ds \\
&= e^{-Lt}\mathbf{r_0} - \sum_{i,j}\left[A_{ij}\int_{t-\tau_j^i}^{t}\mathbf{x}(s)\,ds - B_{ij}\int_{t-\nu_i^j}^{t}\mathbf{x}(s)\,ds\right] + \\
&\quad + \int_0^t e^{-L(t-s)}L\sum_{i,j}\left[A_{ij}\int_{s-\tau_j^i}^{s}\mathbf{x}(w)\,dw - B_{ij}\int_{s-\nu_i^j}^{s}\mathbf{x}(w)\,dw\right]ds
\end{aligned} \tag{4.28}$$

where

$$\mathbf{r_0} := \boldsymbol{\phi}(0) + \sum_{i,j}\left[A_{ij}\int_{-\tau_j^i}^{0}\boldsymbol{\phi}(s)\,ds - B_{ij}\int_{-\nu_i^j}^{0}\boldsymbol{\phi}(s)\,ds\right] \tag{4.29}$$

An alternative way to express the solution of (4.24) in vector form is

$$\dot{\mathbf{x}}(t) = -D\mathbf{x}(t) + \sum_{i,j} A_{ij}\mathbf{x}(t-\tau_j^i) + \sum_{i,j} B_{ij}\frac{d}{dt}\int_{t-\nu_i^j}^{t}\mathbf{x}(s)\,ds$$

and inversion from $t-\tau$ to t yields

$$\begin{aligned}
\mathbf{x}(t) = e^{-D\tau}\mathbf{x}(t-\tau) + \int_{t-\tau}^{t} e^{-D(t-s)}\sum_{i,j}\Bigg[A_{ij}\mathbf{x}(s-\tau_j^i) + \\
+ B_{ij}\frac{d}{ds}\int_{s-\nu_i^j}^{s}\mathbf{x}(w)\,dw\Bigg]\,ds
\end{aligned} \tag{4.30}$$

Following the discussion in § 4.2.3 we combine these two forms of solution to a new one. From (4.30)

$$\begin{aligned}
&\mathbf{x}(t) = \\
&= e^{-D\tau}\mathbf{x}(t-\tau) + \int_{t-\tau}^{t} e^{-D(t-s)} \sum_{i,j} \left[A_{ij}\mathbf{x}(s-\tau_j^i) + B_{ij}\frac{d}{ds}\int_{s-\nu_i^j}^{s} \mathbf{x}(w)\,dw \right] ds \\
&= e^{-D\tau}\mathbf{x}_{(4.28)}(t-\tau) + \int_{t-\tau}^{t} e^{-D(t-s)} \sum_{i,j} \left[A_{ij}\mathbf{x}(s-\tau_j^i) + B_{ij}\frac{d}{ds}\int_{s-\nu_i^j}^{s} \mathbf{x}(w)\,dw \right] ds
\end{aligned}$$

So that for $t \geq \tau$

$$\begin{aligned}
&\mathbf{x}(t) := e^{-D\tau}e^{-L(t-\tau)}\mathbf{r_0} + \\
&+ \sum_{i,j} \Bigg[\int_{t-\tau-\tau_j^i}^{t-\tau} \left(e^{-D(t-\tau_j^i-s)} - e^{-D\tau} \right) A_{ij}\mathbf{x}(s)\,ds + \\
&\qquad\qquad + \int_{t-\tau}^{t-\tau_j^i} e^{-D(t-\tau_j^i-s)} A_{ij}\mathbf{x}(s)\,ds \Bigg] + \\
&+ \sum_{i,j} \left[B_{ij} \int_{t-\nu_i^j}^{t} \mathbf{x}(w)\,dwds - \int_{t-\tau}^{t} e^{-D(t-s)} D B_{ij} \int_{s-\nu_i^j}^{s} \mathbf{x}(w)\,dwds \right] + \\
&+ e^{-D\tau} \int_0^{t-\tau} e^{-L(t-\tau-s)} L \sum_{i,j} \Bigg[A_{ij} \int_{s-\tau_j^i}^{s} \mathbf{x}(w)\,dw - \\
&\qquad\qquad - B_{ij} \int_{s-\nu_i^j}^{s} \mathbf{x}(w)\,dw \Bigg] ds
\end{aligned} \tag{4.31}$$

For asymptotic stability with prescribed convergence estimate we will need an especially designed metric space.

The space of solutions. Fix $\boldsymbol{\psi} \in C^0([-p,\tau],\mathbb{R}^N)$, $\theta > 0$, $k \in \mathbb{R}$. By $\mathbb{B} = C^0([-p,\infty),\mathbb{R}^N)$ we define the set of continuous bounded (in the sense of the supremum norm) functions defined in $[-p,\infty)$ and which take values in $\mathbb{R}^N$. Define

$$\mathbb{M} = \left\{ \mathbf{y} \in \mathbb{B} : \mathbf{y} = \boldsymbol{\psi}|_{[-p,\tau]},\ \sup_{t\geq\tau} e^{\theta t}||\mathbf{y}(t) - \mathbb{1}k||_{\mathbf{q}} < \infty \right\} \tag{4.32}$$

together with the function

$$\rho(\mathbf{y}_1, \mathbf{y}_2) := \sup_{t \geq \tau} e^{\theta t} ||\mathbf{y}_1(t) - \mathbf{y}_2(t)||_{\mathbf{q}} \tag{4.33}$$

where we define

$$||\mathbf{x}||_{\mathbf{q}} = \max_i q_i |x_i|$$

for some $\mathbf{q}$, a vector with strictly positive elements.

This is the space of $\mathbb{R}^N$-valued continuous functions each member of which agrees on $[-p, \tau]$ with a prescribed function ϕ and it converges to $\mathbb{1}k$ exponentially fast with rate θ. We can readily attach this space to our problem by picking $\tau = \max_{i,j} \tau_j^i$, $p = \max\{\nu, \tau\}$, $\boldsymbol{\psi}(t) = \boldsymbol{\phi}(t)$ for $t \in [-p, 0]$ and $\boldsymbol{\psi}(t) = \mathbf{x}_{(4.24)}(t)$ $t \in [0, \tau]$, i.e. the unique solution of (4.24) in $[0, \tau]$. This formulation considers the existence and uniqueness for the solution of (4.24) in $[0, \tau]$ which is hardly an assumption due to the linearity of the model. It is easy to see that the pair of equations (4.32) and (4.33) is equivalent to the pair in (2.9), (2.10) in Proposition 2.5.7, from § 2.5 so that $(\mathbb{M}, \rho)$ is a complete metric space.

The solution operator. We define $\mathcal{P} : \mathbb{M} \to \mathbb{B}$ as follows:

$$(\mathcal{P}\mathbf{x})(t) = \begin{cases} \boldsymbol{\psi}(t), & t \in [-\max\{\tau, \nu\}, \tau] \\ \mathbf{x}_{(4.31)}(t), & t \geq \tau. \end{cases} \tag{4.34}$$

where $\mathbf{x}_{(4.31)}(t)$ is the right hand-side of (4.31). The next step is to show that $\mathcal{P}$ maps $\mathbb{M}$ into itself. The following lemma ensures that this is true for specific k and θ.

Proposition 4.3.5. *$\mathcal{P} : \mathbb{M} \to \mathbb{M}$ if k is defined as in* (4.27) *and $\theta < \Re(\lambda)$.*

Proof. The proof consists of two parts. The first is the calculation k and the second is the estimate of θ. For the first part we will need the following result:

Lemma 4.3.6. *Let L be the weighted Laplacian matrix of a routed-out branching graph G with $\mathbf{c}$ its normalized left eigenvector as defined in Proposition 2.2.1. Let $\mathbf{z} \in C^1([0,\infty),\mathbb{R}^N)$ such that $\lim_{t\to\infty} \mathbf{z}(t) \in \mathbb{R}^N$. Then*

$$\lim_{t\to\infty} \int_0^t e^{-L(t-s)} L\mathbf{z}(s)\, ds = (I_{N\times N} - \mathbb{1}\mathbf{c}^T)\mathbf{z}(\infty)$$

We calculate the t limit of $(\mathcal{P}\mathbf{x})(t)$ as the sum of the four quantities defined in (4.35).

For $\mathbf{x} \in \mathbb{M}$, $\lim_{t\to\infty} \mathcal{P}_i(t)$ yields

$$\lim_t \mathcal{P}_1(t) = \lim_t e^{-D\tau}\big(e^{-L(t-\tau)} - \mathbb{1}\mathbf{c}^T\big)\mathbf{r_0} + e^{-D\tau}\mathbb{1}\mathbf{c}^T\mathbf{r_0} = e^{-D\tau}\mathbb{1}\mathbf{c}^T\mathbf{r_0}$$

$$\begin{aligned}
\lim_t \mathcal{P}_2(t) &= e^{-D\tau}\Big(\sum_{j\in N_1} \frac{a_{1j}}{d_1}\big(e^{d_1\tau_j^1} - 1\big), \ldots, \sum_{j\in N_N} \frac{a_{Nj}}{d_N}\big(e^{d_N\tau_j^N} - 1\big)\Big)^T k \\
&- e^{-D\tau}\Big(\sum_{j\in N_1} a_{1j}\tau_j^1, \ldots, \sum_{j\in N_N} a_{Nj}\tau_j^N\Big)^T k \\
&+ \left(\mathbb{1} - e^{-D\tau}\Big(\sum_{j\in N_1} \frac{a_{1j}}{d_1} e^{d_1\tau_j^1}, \ldots, \sum_{j\in N_N} \frac{a_{Nj}}{d_N} e^{d_N\tau_j^N}\Big)^T\right) k \\
\lim_t \mathcal{P}_3(t) &= e^{-D\tau}\Big(\sum_{j\in N_1} a_{1j}\nu_1^j, \ldots, \sum_{j\in N_N} a_{Nj}\nu_N^j\Big)^T k
\end{aligned}$$

Finally, from Lemma 4.3.6 with $\mathbf{z}(t) = \sum_{i,j}\left[A_{ij}\int_{t-\tau_j^i}^t \mathbf{x}(s)\,ds - B_{ij}\int_{t-\nu_i^j}^t \mathbf{x}(s)\,ds\right]$ we obtain

$$\lim_t \mathcal{P}_4(t) = e^{-D\tau}(I - \mathbb{1}\mathbf{c}^T)\Big(\sum_{j\in N_1} a_{ij}(\tau_j^1 - \nu_1^j), \ldots, \sum_{j\in N_N} a_{Nj}(\tau_j^N - \nu_N^j)\Big)^T k$$

Take $\mathbf{w} = \big(\sum_{j\in N_1} a_{ij}(\tau_j^1 - \nu_1^j), \ldots, \sum_{j\in N_N} a_{Nj}(\tau_j^N - \nu_N^j)\big)^T$ and cancel the common terms to obtain

$$\lim_t (\mathcal{P}\mathbf{x})(t) = \mathbb{1}k + e^{-D\tau}\mathbb{1}\Big(\mathbf{c}^T\mathbf{r}_0 - k - \mathbf{c}^T\mathbf{w}k\Big) = \mathbb{1}k$$

if k is defined as in (4.27). For the second part, Proposition 2.2.1 implies

$$\big|\big|\big(e^{-Lt} - \mathbb{1}\mathbf{c}^T\big)\mathbf{r_0}\big|\big| = \mathcal{O}(t^{r(\lambda)-1}e^{-\Re(\lambda)t})$$

whereas for $\mathbf{x} \in \mathbb{M}_{\theta,k}$ the rest of the terms of $(\mathcal{P}\mathbf{x})(t)$ are of order $\mathcal{O}(e^{\theta t})$. Then $\sup_{t\geq p} e^{\theta t}||(\mathcal{P}\mathbf{x})(t) - \mathbb{1}k||$ is finite for any $\theta < \Re(\lambda)$. □

Proof of Lemma 4.3.6. Set $\mathbf{z}^c(t) := \mathbf{z}(t) - \mathbb{1}\mathbf{c}^T\mathbf{z}(t)$. Then

$$\begin{aligned} Q(t) :&= \int_0^t e^{-L(t-s)}L\mathbf{z}(s)\,ds = \int_0^t \big(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T\big)L\mathbf{z}^c(s)\,ds \\ &= \int_0^t \frac{d}{ds}\big(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T\big)\mathbf{z}^c(s)\,ds = \int_0^t d(e^{-L(t-s)})\mathbf{z}^c(s)\,ds \\ &= \mathbf{z}^c(t) - (e^{-Lt} - \mathbb{1}\mathbf{c}^T)\mathbf{z}^c(0) - \int_0^t (e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T)\dot{\mathbf{z}}^c(s)\,ds. \end{aligned}$$

The result follows from Assumption 4.3.1 and hence the observation that the integral asymptotically vanishes because it is a convolution of an L^1 function with a function that tends to zero. □

The quantity k as defined in (4.27) is the consensus point and it is, as expected, a function only of the parameters of the system, the initial data and it is well-defined in view of Assumption 4.3.2.

Finally, it is shown that $\mathcal{P}$ satisfies the contraction property for some θ.

Lemma 4.3.7. *Under Assumption 4.3.3, $\mathcal{P} : \mathbb{M} \to \mathbb{M}$ is a contraction for some $\theta \in (0, \min\{d_i, \Re(\lambda)\})$.*

Proof of Lemma 4.3.7. From the definition of $\mathcal{P}$ in (4.34) we observe that for $t \geq \tau$ it can be written as the sum

$$\mathcal{P} = \mathcal{P}_1 + \mathcal{P}_2 + \mathcal{P}_3 + \mathcal{P}_4 \tag{4.35}$$

where

$$\begin{aligned}
\mathcal{P}_1(t) &:= e^{-D\tau} e^{-L(t-\tau)} \mathbf{r_0} \\
\mathcal{P}_2(t) &:= \sum_{i,j} \int_{t-\tau-\tau_j^i}^{t-\tau} \left(e^{-D(t-\tau_j^i - s)} - e^{-D\tau} \right) A_{ij} \mathbf{x}(s)\, ds + \\
&\qquad\qquad + \sum_{i,j} \int_{t-\tau}^{t-\tau_j^i} e^{-D(t-\tau_j^i-s)} A_{ij} \mathbf{x}(s)\, ds \\
\mathcal{P}_3(t) &:= \sum_{i,j} B_{ij} \int_{t-\nu_i^j}^{t} \mathbf{x}(s)\, ds - \int_{t-\tau}^{t} e^{-D(t-s)} D B_{ij} \int_{s-\nu_i^j}^{s} \mathbf{x}(w)\, dw ds \\
\mathcal{P}_4(t) &:= e^{-D\tau} \int_0^{t-\tau} e^{-L(t-\tau-s)} L \sum_{i,j} \left[A_{ij} \int_{s-\tau_j^i}^{s} \mathbf{x}(w)\, dw - B_{ij} \int_{s-\nu_i^j}^{s} \mathbf{x}(w)\, dw \right] ds
\end{aligned}$$

Take $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{M}$. Then

$$\begin{aligned}
\rho((\mathcal{P}\mathbf{x}_1), (\mathcal{P}\mathbf{x}_2)) &= \sup_{t \geq \tau} e^{\theta t} ||(\mathcal{P}\mathbf{x}_1)(t) - (\mathcal{P}\mathbf{x}_2)(t)||_{\mathbf{q}} \\
&\leq \sum_{l=1}^{4} \sup_{t \geq \tau} e^{\theta t} ||(\mathcal{P}_l \mathbf{x}_1)(t) - (\mathcal{P}_l \mathbf{x}_2)(t)||_{\mathbf{q}}.
\end{aligned}$$

The contribution of each $\mathcal{P}_l \mathbf{x}_1 - \mathcal{P}_l \mathbf{x}_2$ is studied separately. At first, $\mathcal{P}_1 \mathbf{x}_1 - \mathcal{P}_1 \mathbf{x}_2$ contributes nothing. For the rest we work as follows.

$\mathcal{P}_2$: We estimate the upper bound of $e^{\theta t} q_i |(\mathcal{P}_2 \mathbf{x}_1)^{(i)}(t) - (\mathcal{P}_2 \mathbf{x}_2)^{(i)}(t)|$. Observe that $e^{-d_i(t-s-\tau_j^i)} - e^{-d_i \tau}$ is non-negative for $s \in [t-\tau-\tau_j^i, t-\tau]$ and for convenience set

$\mathbf{x}_{12}(s) := \mathbf{x}_1(s) - \mathbf{x}_2(s)$ and $\rho := \sup_{t\geq\tau} e^{\theta t}||\mathbf{x}_{12}(t)||_{\mathbf{q}}$.

$$\bigg|\int_{t-\tau-\tau_j^i}^{t-\tau}\big(e^{-d_i(t-s-\tau_j^i)} - e^{-d_i\tau}\big)a_{ij}\mathbf{x}_{12}^{(j)}(s)\,ds + \int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)}a_{ij}\mathbf{x}_{12}^{(j)}(s)\,ds\bigg| \leq$$
$$\leq \bigg[\int_{t-\tau-\tau_j^i}^{t-\tau}\big(e^{-d_i(t-s-\tau_j^i)} - e^{-d_i\tau}\big)a_{ij}e^{-\theta s}\,ds + \int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)}a_{ij}e^{-\theta s}\,ds\bigg]\frac{1}{q_j}\rho$$

$\Rightarrow$ where the two integrals give

$$e^{\theta t}\int_{t-\tau-\tau_j^i}^{t-\tau}\big(e^{-d_i(t-s-\tau_j^i)} - e^{-d_i\tau}\big)a_{ij}e^{-\theta s}\,ds =$$
$$= \frac{a_{ij}}{d_i-\theta}\bigg(e^{d_i(\tau_j^i-\tau)+\theta\tau} - e^{-(d_i-\theta)\tau+\theta\tau_j^i}\bigg) - \frac{a_{ij}}{\theta}e^{-(d_i-\theta)\tau}\big(e^{\theta\tau_j^i}-1\big),$$
$$e^{\theta t}\int_{t-\tau}^{t-\tau_j^i} e^{-d_i(t-s-\tau_j^i)}a_{ij}e^{-\theta s}\,ds = \frac{a_{ij}}{d_i-\theta}\bigg(e^{\theta\tau_j^i} - e^{d_i(\tau_j^i-\tau)+\theta\tau}\bigg)$$

Check that the first and last term cancel and that summing over $j \in N_i$ we obtain the following estimate

$$q_i\sum_{j\in N_i}\bigg[a_{ij}e^{\theta\tau_j^i}\frac{1-e^{-(d_i-\theta)\tau}}{d_i-\theta} - a_{ij}e^{-(d_i-\theta)\tau}\frac{e^{\theta\tau_j^i}-1}{\theta}\bigg]\frac{1}{q_j}\rho \tag{4.36}$$

Remark 4.3.8. As $\theta \downarrow 0$, the last expression becomes

$$q_i\sum_{j\in N_i}a_{ij}\bigg[\frac{1-e^{-d_i\tau}}{d_i} - \tau_j^i e^{-d_i\tau}\bigg]\frac{1}{q_j}\rho$$

$\mathcal{P}_3$: Similar manipulation yields

$$q_ie^{\theta t}\bigg|\sum_{j\in N_i}a_{ij}\int_{t-\nu_i^j}^{t}\mathbf{x}_{12}^{(i)}(s)\,ds\bigg| + e^{\theta t}\bigg|\sum_{j\in N_i}a_{ij}\int_{t-\tau}^{t}d_ie^{-d_i(t-s)}\int_{s-\nu_i^j}^{s}\mathbf{x}_{12}^{(i)}(w)\,dwds\bigg|$$
$$\leq \sum_{j\in N_i}a_{ij}\frac{e^{\theta\nu_i^j}-1}{\theta}\rho + \sum_{j\in N_i}a_{ij}\frac{e^{\theta\nu_i^j}-1}{\theta}d_i\frac{1-e^{-(d_i-\theta)\tau}}{d_i-\theta}\rho$$

Note that in this case the weights q_i's are canceled and so we get the estimate

$$\sum_{j\in N_i}a_{ij}\frac{e^{\theta\nu_i^j}-1}{\theta}\bigg(1+d_i\frac{1-e^{-(d_i-\theta)\tau}}{d_i-\theta}\bigg)\rho. \tag{4.37}$$

Remark 4.3.9. As $\theta \downarrow 0$, the last expression becomes

$$\sum_{j\in N_i} a_{ij}\nu_i^j\big(2 - e^{-d_i\tau}\big)\rho$$

$\mathcal{P}_4$: Finally,

$$\mathcal{P}_4(t) := e^{-D\tau}\int_0^{t-\tau} e^{-L(t-\tau-s)}L\sum_{l,m}\bigg[A_{lm}\int_{s-\tau_m^l}^{s}\mathbf{x}_{12}(w)\,dw - \\ - B_{lm}\int_{s-\nu_l^m}^{s}\mathbf{x}_{12}(w)\,dw\bigg]\,ds$$

Let κ_{ij} be the $(i,j)^{th}$ element of $e^{-Lt}L$ so that $\kappa_{ii} = -\sum_{j\neq i}\kappa_{ij}$. A careful calculation for on the i^{th} row of $e^{-Lt}L\sum_{l,m}A_{lm}\int_{s-\tau_m^l}^{s}\mathbf{x}_{12}(w)\,dw$ yields

$$\sum_{l=1}^{N}\kappa_{il}\sum_{j=1}^{N}a_{lj}\int_{s-\tau_j^l}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw =$$
$$\kappa_{ii}\sum_{j=1}^{N}a_{ij}\int_{s-\tau_j^i}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw + \sum_{l\neq i}^{N}\kappa_{il}\sum_{j=1}^{N}a_{lj}\int_{s-\tau_j^l}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw =$$
$$\sum_{l\neq i}\kappa_{il}\bigg[\sum_{j=1}^{N}\bigg(a_{lj}\int_{s-\tau_j^l}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw - a_{ij}\int_{s-\tau_j^i}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw\bigg)\bigg] =$$
$$\sum_{l\neq i}\kappa_{il}\bigg[\sum_{j=1}^{N}\bigg(\big(a_{lj}-a_{ij}\big)\int_{s-\tau_j^l}^{s}\mathbf{x}_{12}^{(j)}(w)\,dw + a_{ij}\int_{s-\max\{\tau_j^l,\tau_j^i\}}^{s-\min\{\tau_j^l,\tau_j^i\}}\mathbf{x}_{12}^{(j)}(w)\,dw\bigg)\bigg]$$

Recall the notations g_{il} and $h_{l,j,i}$ as in (4.25) and (4.26) respectively. Then the first bound is

$$q_ie^{-d_i\tau}\sum_{l\neq i}\sum_{j=1}^{N}h_{l,j,i}(\theta)g_{il}(\theta)\frac{1}{q_j}\rho.$$

The second bound is

$$q_ie^{-d_i\tau}\sum_{l=1}^{N}\sum_{j=1}^{N}g_{il}(\theta)a_{lj}\frac{e^{\theta\nu_l^j}-1}{\theta}\frac{1}{q_l}\rho.$$

We add them both to obtain

$$q_i e^{-d_i\tau}\left[\sum_{l\neq i}\sum_{j=1}^{N} h_{l,j,i}(\theta)g_{il}(\theta)\frac{1}{q_j}+\sum_{l=1}^{N}\sum_{j=1}^{N} g_{il}(\theta)a_{lj}\frac{e^{\theta\nu_l^j}-1}{\theta}\frac{1}{q_l}\right]\rho. \tag{4.38}$$

Remark 4.3.10. As $\theta\downarrow 0$, the last expression becomes

$$q_i e^{-d_i\tau}\left[\sum_{l\neq i}\sum_{j=1}^{N} h_{l,j,i}(0)g_{il}(0)\frac{1}{q_j}+\sum_{l=1}^{N}\sum_{j=1}^{N} g_{il}(0)a_{lj}\nu_l^j\frac{1}{q_l}\right]\rho.$$

Combine Remarks 4.3.8, 4.3.9, 4.3.10 and reorder the weights q_i to obtain the condition of Assumption 4.3.3.

□

4.3.1 Examples and simulations

We will review now a number of illustrative applications of Theorem 4.3.4, compare it with older results and/or with other approaches. In particular we will be propose a way to prove the existence of the weights q_i and hence define an appropriate metric of proving stability under supplementary symmetry assumptions.

Example 4.3.11. For $N=2$, (4.24) becomes

$$\begin{aligned}\dot{x}(t)&=-ax(t-\nu_1)+ay(t-\tau_1)\\ \dot{y}(t)&=-by(t-\nu_2)+bx(t-\tau_2)\end{aligned} \tag{Ex.1}$$

where with a,b so that Assumption 4.3.1 holds and $\nu_i,\tau_i\geq 0$. Without loss of generality we take $\tau_1\geq\tau_2$. For convenience set $\Lambda=\frac{ab}{a+b}$. Since $\mathbf{c}=\{\frac{b}{a+b},\frac{a}{a+b}\}$ Assumption 4.3.2 requires

$$1+\Lambda\big(\tau_1+\tau_2-\nu_1-\nu_2\big)>0 \tag{4.39}$$

Next,

$$e^{-Lt}L = \begin{bmatrix} a & -a \\ -b & b \end{bmatrix} e^{-(a+b)t}$$

The matrix $G(\theta) = \{g_{ij}(\theta)\}$ reads

$$G(\theta) = \begin{bmatrix} a & a \\ b & b \end{bmatrix} \frac{e^{\theta\tau_1}}{a+b-\theta}$$

whenever $\theta < a+b$. Then we consider the elements of $F(\theta)$

$$\begin{aligned} f_{12} &:= ae^{-\theta\tau_1}\frac{1-e^{-(a-\theta)\tau_1}}{a-\theta} - ae^{-(a-\theta)\tau_1}\frac{e^{\theta\tau_1}-1}{\theta} + \\ &\qquad + ae^{-a\tau_1}\frac{e^{\theta\tau_1}-1}{\theta}\frac{a}{a+b-\theta}e^{\theta\tau_1} + e^{-a\tau_1}\frac{abe^{\theta\tau_1}}{a+b-\theta}\frac{e^{\theta\nu_2}-1}{\theta} \\ f_{11} &:= -\Big(1 - a\frac{e^{\theta\nu_1}-1}{\theta}\big(1+\frac{a}{a-\theta}(1-e^{-(a-\theta)\tau_1})\big) - \\ &\qquad - e^{-a\tau_1}\Lambda\big(\frac{e^{\theta\tau_2}-1}{\theta} + \frac{a}{b}\frac{e^{\theta\nu_1}-1}{\theta}\big)\Big) \\ f_{21} &:= be^{-\theta\tau_2}\frac{1-e^{-(b-\theta)\tau_1}}{b-\theta} - be^{-(b-\theta)\tau_1}\frac{e^{\theta\tau_2}-1}{\theta} \\ &\qquad + be^{-b\tau_1}\frac{e^{\theta\tau_2}-1}{\theta}\frac{b}{a+b-\theta}e^{\theta\tau_1} + e^{-b\tau_1}\frac{abe^{\theta\tau_1}}{a+b-\theta}\frac{e^{\theta\nu_1}-1}{\theta} \\ f_{22} &:= -\Big(1 - b\frac{e^{\theta\nu_2}-1}{\theta}\big(1+\frac{b}{b-\theta}(1-e^{-(b-\theta)\tau_1})\big) - \\ &\qquad - e^{-b\tau_1}\Lambda\big(\frac{e^{\theta\tau_1}-1}{\theta} + \frac{b}{a}\frac{e^{\theta\nu_2}-1}{\theta}\big)\Big) \end{aligned}$$

and for $\theta \downarrow 0$, the elements of $F(0)$

$$\begin{aligned} f_{12} &:= \big(1 - e^{-a\tau_1} - \Lambda\tau_1 e^{-a\tau_1} + e^{-a\tau_1}\Lambda\nu_2\big) \\ f_{11} &:= -\Big(1 - a\nu_1\big(2 - e^{-a\tau_1}\big) - e^{-a\tau_1}\Lambda\big(\tau_2 + \frac{a}{b}\nu_1\big)\Big) \\ f_{21} &:= \big(1 - e^{-b\tau_1} - \Lambda\tau_2 e^{-b\tau_1} + e^{-b\tau_1}\Lambda\nu_1\big) \\ f_{22} &:= -\Big(1 - b\nu_2(2 - e^{-b\tau_1}) - e^{-b\tau_1}\Lambda\big(\tau_1 + \frac{b}{a}\nu_2\big)\Big). \end{aligned}$$

Then Assumption 4.3.3 requires finding $q_1, q_2 > 0$ so that

$$f_{12}\frac{1}{q_2} < (-f_{11})\frac{1}{q_1} \quad \text{and} \quad f_{21}\frac{1}{q_1} < (-f_{22})\frac{1}{q_2}$$

a set of linear inequalities which is consistent if and only if

$$f_{11}f_{22} > f_{12}f_{21} \tag{4.40}$$

whenever $f_{ii} < 0, f_{ij} > 0$. Eqs. (4.39) and (4.40) describe the allowed bounds for the processing and propagation delays.

As a numerical application we take $a = 0.5$ and $b = 1.3$, $\tau_1 = 1$ $\tau_2 = 0.2$ and $\nu_1 = \nu_2 = 0.215$. Then (4.40) remains consistent for $\theta \leq 0.045$ which is our estimate for the rate of convergence. Next we focus on two extreme cases:

No propagation delays. In this case the bounds are rightfully much stricter. As the necessary condition by Assumption 4.3.2 suggests it is impossible to obtain stability for any bounded ν. The conditions imposed are

$$\nu_1 + \nu_2 < \frac{1}{\Lambda}$$

by Assumption 4.3.2 and

$$1 + 2ab\nu_1\nu_2 > \frac{2(a^2\nu_1 + b^2\nu_2) + ab(\nu_1 + \nu_2)}{a + b}$$

from condition (4.40). More specifically, for $\nu_1 = \nu_2 = \nu$ we ask

$$\nu < \frac{1}{2}\frac{a + b}{ab}, \qquad \text{and} \qquad \nu^2 - \frac{a^2 + b^2 + (a + b)^2}{2ab(a + b)}\nu + \frac{1}{2ab} > 0.$$

No processing delays. If we ignore the processing delays, the requirement of Assumption 4.3.2 is automatically satisfied and condition (4.40) is simplified to

$$\left(1-\Lambda\tau_1 e^{-b\tau_1}\right)\left(1-\Lambda\tau_2 e^{-a\tau_1}\right) > \left(1-e^{-a\tau_1}-\Lambda\tau_1 e^{-a\tau_1}\right)\left(1-e^{-b\tau_1}-\Lambda\tau_2 e^{-b\tau_1}\right)$$

Consequently, so long as the above inequality is satisfied, there are always $q_1, q_2 > 0$ and hence a weighted metric ρ so that the operator as defined in (4.34) is a contraction in $(\mathbb{M}, \rho)$.

Note that the above inequality does not hold for any a, b, τ_1, τ_2. It is true however either when $a = b > 0$ and arbitrary $\tau_1, \tau_2 \geq 0$ or when $\tau_1 = \tau_2$ and arbitrary $a, b \geq 0$. This allows us to establish bounds around any nominal value w^* in the vicinity of which for any given τ_1, τ_2 there exists a radius r so that a, b can lie in $B(w^*, r)$ and similarly for τ_1, τ_2. In our numerical example we take $\nu_1 = \nu_2 = 0$ and the rate estimate we get is $\theta = 0.4545$.

Remark 4.3.12. In [7] we considered (4.24) with $\nu = 0$ and we used an operator based exclusively on the solution expression of (4.28). Then under the L^1 norm, $||\cdot||_1$ we derived the following contraction condition

$$\tilde{A}\tau\left(1+\frac{\sqrt{N}||L||_1}{\lambda_2}\right) < 1 \tag{4.41}$$

where $\tilde{A} = \sum_{i=1}^{N}\sum_{j\in N_i} a_{ij}$ and $a_{ij} = a_{ji}$. It should be intuitively clear that 4.41 is much stricter than the condition (4.3.3) and we will illustrate this difference within this example. Applying the bound (4.41) stability is ensured if

$$\tau < \frac{1}{2(1+\sqrt{2})a}$$

Remark 4.3.13. Taking $a = b$ and $\tau_i = \nu_i$ we can compare our results with [45] where the necessary and sufficient condition is

$$\tau < \frac{2\pi}{\lambda_N} = \frac{\pi}{a}$$

while our bounds ask

$$\tau < \frac{1}{2a}$$

Remark 4.3.14. We conclude this example by mentioning a candidate Lyapunov functional in the case $\tau_1 = \tau_2 = \tau$.

$$V(x, y) = bx^2 + ay^2 + ab \int_{t-\tau}^{t} x^2(s) + y^2(s)\, ds$$

This functional on $\mathbb{R}^2$ is obviously continuous and

$$\dot{V} = -ab[(x - y_t)^2 + (y - x_t)^2]$$

Then the set, S, such that $\dot{V}(t) \equiv 0$ is the one where $x(t) = y(t-\tau)$ and $y(t) = x(t-\tau)$ for all t. The largest subset of S that is invariant with respect to these dynamics is Δ. Then standard Invariance Theory arguments yield asymptotic convergence to the consensus subspace (Section 5.3 of [59]) independent of the magnitude of the delay (but without any estimate on the rate of convergence).

Example 4.3.15 (A Delayed Complete Graph.)**.** Let (4.24) with $a_{ij} \equiv \theta$, $\nu_i^j \equiv 0$ $\tau_j^i = \tau_i$. Under this communication scheme every agent is connected with everyone else with identical connectivity weight and they receive the signals with propagation delay that depends on agent i only. The complete graph on N agents has the edge set $E = \{(i,j) : i \neq j\}$. The Laplacian of this graph has the spectrum $\lambda_1 = 0, \lambda_i = N\theta|_{i=2}^{N}$. Then

$$[e^{-Lt}L]_{ij} = \begin{cases} \theta \frac{N-1}{N} e^{-N\theta t}, & i = j \\ \\ \theta \frac{-1}{N} e^{-N\theta t}, & i \neq j \end{cases}$$

and $d_i \equiv (N-1)\theta$ and the left eigenvector of L is $\mathbf{c} = \mathbb{1}\frac{1}{N}$. Assumption 4.3.2 is automatically satisfied whereas for Assumption 4.3.3 we have

$$f_{ij}(0) = \frac{1}{N-1}(1 - e^{-(N-1)\theta\tau}) - \theta\tau_i e^{-(N-1)\theta\tau} + \\ + \frac{\theta}{N^2} e^{-(N-1)\theta\tau} \sum_{l \neq i} (\max\{\tau_i, \tau_l\} - \min\{\tau_i, \tau_l\})$$

$$f_{ii}(0) = -1 + \frac{\theta}{N^2} e^{-(N-1)\theta\tau} \sum_{l \neq i} \tau_l$$

We will apply Theorem 2.7.1 and for this we will need the following condition. Let for $i \in [N]$, N_i denote the number of $l \neq i$ such that $\tau_l \geq \tau_i$. We ask that

$$1 + \frac{N^2 + 2N_i - N}{N} \theta\tau_i - \frac{N-1}{N^2} \theta \left[\sum_{l:\tau_l \geq \tau_i} \tau_l - \sum_{l:\tau_l < \tau_i} \tau_l \right] > 0 \tag{4.42}$$

Note that this condition is satisfied if for example $\tau_i \equiv \tau$. From Theorem 2.7.1 we choose $m = 2N$, $\mathbf{a}_i := \big(f_{i1}, \ldots, f_{iN}\big)$ for $i = 1, \ldots, N$, $\mathbf{a}_i < 0$ elementwise for $i = N+1, \ldots, 2N$, $\alpha_i = 0$ for $i = 1, \ldots, N$ and $\alpha_i < 0$ for $i = N+1, \ldots, 2N$. Then for the sake of contradiction if the second case holds there exist $\xi_i|_{i=1}^{2N} \geq 0$ so that at least one of $\xi|_{i \leq N}$ is positive and $\sum_{i=1}^{2N} \alpha_i \xi_i = \sum_{i=N+1}^{2N} \alpha_i \xi_i \leq 0$ and $\xi_i|_{i \geq N+1} \geq 0$. From the second condition

$$\sum_{i=1}^{2N} \mathbf{a}_i \xi_i = 0 \Rightarrow \sum_{j=1}^{N} f_{ij}\xi_j + \sum_{j=1}^{N} a_{ij}\xi_j = 0,\ \forall i$$

Since the second part of the last equation is non-positive from the imposed condition (4.42) it can be verified that it implies $\sum_j f_{ij} < 0$ for all i, hence a contradiction because not all $\xi|_{i\leq N}$ can be zero. So there exists a set of positive numbers (i.e. a weighted metric) to satisfy the Assumption 4.3.3.

Example 4.3.16 (Uniform Delays in a topological Star Graph)**.** We consider the star graph among N agents and we enumerate the central node of the graph to be the first agent. We take the communication weights identically equal to the unity. It can be shown that

$$[e^{-Lt}L]_{ij} = \begin{cases} (N-1)e^{-Nt}, & i = j = 1 \\ -e^{-Nt}, & i = 1, j \neq 1 \text{or} j = 1, i \neq 1 \\ \frac{1}{N-1}e^{-Nt} + \frac{N-2}{N-1}e^{-t}, & i = j, i \neq 1 \\ \frac{1}{N-1}e^{-Nt} - \frac{1}{N-1}e^{-t}, & o.w. \end{cases}$$

so that $g_{ii}(0) = \frac{N-1}{N}$ and $g_{ij}(0) = \frac{1}{N}$ otherwise. Assumption 4.3.2 asks for

$$1 + \frac{1}{N}\sum_{j\neq 1}(\tau_j^1 - \nu_1^j) + \frac{1}{N}\sum_{j\neq 1}(\tau_1^j - \nu_j^1) > 0, \qquad i \in [N]$$

In particular if $\tau_j^i \equiv \tau$ and $\nu_i^j \equiv 0$ the condition is automatically satisfied. Finally, for Assumption 4.3.3 we calculate the elements of $F(0)$

$$f_{1j} = \begin{cases} -1 + \frac{N-1}{N}\tau e^{-(N-1)\tau} & j = 1 \\ \frac{1}{N-1}(1 - e^{-(N-1)\tau}) - \tau e^{-(N-1)\tau} + \frac{N-1}{N}\tau e^{-(N-1)\tau} & o.w. \end{cases}$$

and for $i > 1$

$$f_{ij} = \begin{cases} -1 & j = i \\ 1 - e^{-\tau} - \tau e^{-\tau} + \frac{\tau}{N} e^{-\tau} & j = 1 \\ \frac{\tau}{N} e^{-\tau} & o.w. \end{cases}$$

A simple calculation shows that $\sum_j f_{ij} < 0$ and the argument proceeds as in the Example 4.3.15.

Let us now turn to a numerical example where FPT methods suffer from the asymmetry of the delays and the weighted topology and hence the derived conditions ask for a very small bound on the maximum allowed delay τ.

Example 4.3.17. Consider a weighted graph with 4 nodes and the symmetric Laplacian matrix L

$$L = \begin{bmatrix} 6.3 & 0 & -2 & -4.3 \\ 0 & 4.8 & -3 & -1.8 \\ -2 & -3 & 6.1 & -1.1 \\ -4.3 & -1.8 & -1.1 & 7.2 \end{bmatrix}$$

We also assume the distribution of delays:

$$T_1 = \begin{bmatrix} 0 & 0.6241 & 0.9880 & 0.7962 \\ 0.5211 & 0 & 0.0377 & 0.0987 \\ 0.2316 & 0.3955 & 0 & 0.2619 \\ 0.4889 & 0.3674 & 0.9133 & 0 \end{bmatrix} \tau$$

with the control parameter τ . We calculate the allowed bound for (4.41) $\Re(\lambda) = 4.534$, $||L||_1 = 14.400$, $\tilde{A} = 24.400$ and we obtain $\tau_{(4.41)} < 0.0057$. Now we will Apply Theorem 4.3.4. We execute the following calculations.

$$e^{-Lt}L = \begin{bmatrix} -20.66 & -11.80 & 8.07 & 24.39 \\ -11.80 & -2.51 & 4.84 & 9.47 \\ 8.07 & 4.84 & -3.15 & -9.76 \\ 24.39 & 9.47 & -9.76 & -24.10 \end{bmatrix} e^{-4.53t} +$$

$$+ \begin{bmatrix} 0.06 & -0.37 & 0.61 & -0.3 \\ -0.37 & 1.82 & -3.01 & 1.56 \\ 0.61 & -3.01 & 5.00 & -2.60 \\ -0.30 & 1.56 & -2.60 & 1.34 \end{bmatrix} e^{-8.23t} +$$

$$+ \begin{bmatrix} 26.88 & 12.17 & -10.67 & -28.38 \\ 12.17 & 5.50 & -4.84 & -12.83 \\ -10.67 & -4.84 & 4.25 & 11.26 \\ -28.38 & -12.83 & 11.26 & 29.95 \end{bmatrix} e^{-11.63t}$$

where we used Putzer's Algorithm [80] and MAPLE. Next for $\tau_{\max} = 0.988\tau$ the matrix $G(0)$ is approximated as

$$G(0) = \begin{bmatrix} 2.454 & 1.603 & 0.993 & 2.999 \\ 1.603 & 0.413 & 0.494 & 1.215 \\ 0.993 & 0.494 & 0.535 & 1.502 \\ 2.999 & 1.215 & 1.502 & 2.826 \end{bmatrix}$$

also and with the use of MAPLE we calculate $F(0)$ as a function of τ:

$$F := [\mathbf{F}_1 : \mathbf{F}_2 : \mathbf{F}_3 : \mathbf{F}_4]$$

where

$$\mathbf{F}_1 = \begin{bmatrix} 6.76\tau e^{-6.22\tau} - 1 \\ 2.78\tau e^{-4.74\tau} \\ 0.32(1 - e^{-6.02\tau}) + 3.25\tau e^{-6.02\tau} \\ 0.59(1 - e^{-7.11\tau}) + 9.55\tau e^{-7.11\tau} \end{bmatrix}, \mathbf{F}_2 = \begin{bmatrix} 3.16\tau e^{-6.22\tau} \\ 1.38\tau e^{-4.74\tau} - 1 \\ 0.49(1 - e^{-6.02\tau}) + 2.72\tau e^{-6.02\tau} \\ 0.25(1 - e^{-7.11\tau}) + 5.68\tau e^{-7.11\tau} \end{bmatrix}$$

$$\mathbf{F}_3 = \begin{bmatrix} 0.31(1 - e^{-6.22\tau}) + 6.00\tau e^{-6.22\tau} \\ 0.62(1 - e^{-4.74\tau}) + 11.39\tau e^{-4.74\tau} \\ 3.52\tau e^{-6.02\tau} - 1 \\ 0.16(1 - e^{-7.11\tau}) + 4.67\tau e^{-7.11\tau} \end{bmatrix}, \mathbf{F}_4 = \begin{bmatrix} 0.69(1 - e^{-6.22\tau}) + 15.16\tau e^{-6.22\tau} \\ 0.38(1 - e^{-4.74\tau}) + 5.47\tau e^{-4.74\tau} \\ 0.19(1 - e^{-6.02\tau}) + 3.38\tau e^{-6.02\tau} \\ 10.91\tau e^{-7.11\tau} - 1 \end{bmatrix}$$

then

$$\sum_j f_{ij} = \begin{bmatrix} 31.08\tau e^{-6.22\tau} - e^{-6.22\tau} \\ 21.02\tau e^{-4.74\tau} - e^{-4.74\tau} \\ 9.35\tau e^{-6.02\tau} - e^{-6.02\tau} \\ 30.81\tau e^{-7.11\tau} - e^{-7.11\tau} \end{bmatrix}$$

and the first τ such that $\sum_i f_{ij} = 0$ is $\tau^* = 0.0414$, so stability is ensured for $\tau < \tau^*$ with the same argument as in the Example 4.3.15. This is an improvement of the bound in (4.41) by almost an order of 10.

Example 4.3.18. [A simulation example] Consider the weighted network

$$L = \begin{bmatrix} 0.2 & -0.2 & 0 \\ -0.1 & 0.1 & 0 \\ -0.4 & 0 & 0.4 \end{bmatrix}$$

and the distribution of the processing and propagation delays

$$\Sigma(\nu) = \begin{bmatrix} 0 & \frac{\nu}{2} & 0 \\ \frac{\nu}{6} & 0 & 0 \\ \frac{\nu}{10} & \frac{\nu}{7} & 0 \end{bmatrix}, \; T(\tau) = \begin{bmatrix} 0 & \frac{\tau}{10} & 0 \\ \frac{7\tau}{10} & 0 & 0 \\ \tau & \frac{\tau}{6} & 0 \end{bmatrix}$$

Then

$$e^{-Lt}L = \begin{bmatrix} 0.2e^{-0.3t} & -0.2e^{-0.3t} & 0 \\ -0.1e^{-0.3t} & 0.1e^{-0.3t} & 0 \\ -1.2e^{-0.4t} + 0.8e^{-0.3t} & 0.8e^{-0.4t} - 0.8e^{-0.3t} & 0.4e^{-0.4t} \end{bmatrix}$$

and

$$G(\theta, \tau) = \begin{bmatrix} \frac{0.2}{0.3-\theta} & \frac{0.2}{0.3-\theta} & 0 \\ \frac{0.1}{0.3-\theta} & \frac{0.1}{0.3-\theta} & 0 \\ \frac{0.8}{0.3-\theta} - \frac{0.4}{0.4-\theta} & \frac{0.8}{0.3-\theta} - \frac{0.8}{0.4-\theta} & \frac{0.4}{0.4-\theta} \end{bmatrix} e^{\theta\tau}.$$

For $\tau = 2, \nu = 0.4$ the matrix $F(0.2)$ is calculated to be

$$F(0.2) = \begin{bmatrix} -0.539 & 0.441 & 0 \\ 0.221 & -0.934 & 0 \\ 6.198 & 0.175 & -0.952 \end{bmatrix}$$

and then we can pick $q_1 = 1, q_2 = 1, q_3 \in (0, 0.145)$ to apply Theorem 4.3.4. The simulation results are depicted in Fig. 4.3 (see captions for details).

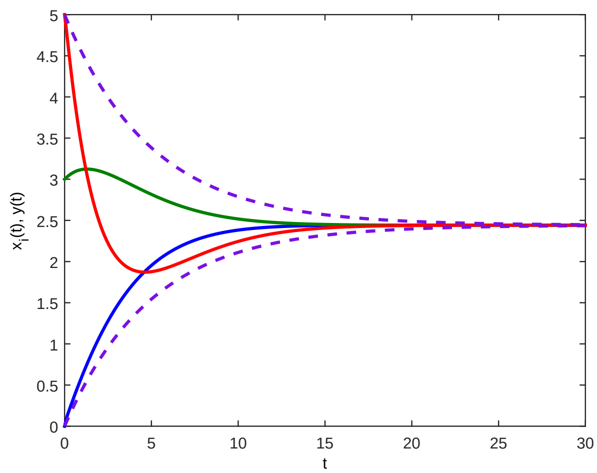

Figure 4.3: Numerical investigations for Example 4.3.18. The initial conditions are $\phi_1(t) = \sin(8t) + 3t$, $\phi_2(t) = 3\sin(800t) + 3$, $\phi_3(t) = \sin(4t) + 5$ and correspond to the solutions of the system x_1 (blue), x_2 (green) and x_3 (red). The rate function $y(t) = |y_0|e^{-0.2t}$ is the dashed line. The simulation was done with the use of the `ddesd` function in `MATLAB`.

4.3.1.1 Remarks on the FPT Method

We developed a Lyapunov-free argument to the study the stability of a linear distributed consensus system with multiple delays. Our main goal was to establish explicit estimates on the rate of convergence of the solutions, as functions of the system's parameters.

We studied the dynamics of system (4.24) by combining two forms of its solution. The first form, (4.28), is a perturbation of the un-delayed system. This way we exploit the valuable kernel e^{-Lt} that describes the global dynamics of the distributed algorithm. The second one, (4.30), characterizes the dynamical behavior of the system in a local manner. It expresses the rate at which each agent converges to the weighted sum of the delayed states of its neighboring agents. This representation illustrates the dissipative convex averaging of the algorithm but sheds no light on the global dynamics. Due to the processing delays, (4.30) is, in turn, a perturbation with respect to the model with just propagation delays.

Given the initial conditions $\boldsymbol{\phi}$ and the system parameters, we considered the unique solution of (4.24) on $[-\max\{\tau, \nu\}, \tau]$ and defined a space of functions each member of which agrees on $[-\max\{\tau, \nu\}, \tau]$. This extension of the solution is due to the way we combined the two solution operators (4.28) and (4.30) and has no effect in the study of the stability of solutions. Next, each member of this function space converges exponentially fast to a constant value k according to a metric with weights $\mathbf{q}$ and exponent rate θ. Based on the solution operator (4.31), we adapted our metric space so that the operator maps this space into itself. In particular, the rate cannot be faster than the local and global exponential rates d_i and $\Re\{\lambda\}$ respectively. Also due to (4.28), the convergence point has a closed form, explicitly defined from the system parameters and the initial data. Another point of interest is that only if k is finite, the conclusion that our metric space is complete and this is guaranteed by Assumption 4.3.2. An assumption which puts an upper bound on the difference between the propagation and processing delays.

We turn our attention to the metric function ρ as defined in (2.10). This is a generalized weighted metric. The supremum over t is necessary for our space to be complete. We chose to work on the maximum over $i \in [N]$ metric because we are dealing with integral equations of a primarily asymmetric system. We introduced the weights $q_i > 0$ as a design feature in an attempt to capture the geometry of the state space of solutions as it depends on a_{ij} ν_i^j and τ_j^i. This idea is motivated by Lyapunov's first method in the stability analysis of the general linear system $\dot{x} = Ax$, where one the investigator is asked to derive the classical Lyapunov matrix P in the Lyapunov equation. This way we essentially transformed the contraction problem into a system of linear inequalities problem and applied existence theorems from Convex Analysis in order to ensure that we can find an appropriate metric to make our operator a contraction one. This is achieved with the application of Theorem 2.7.1.

Advantages. Using fixed one needs not to encounter the difficult task of finding a Lyapunov function. As it is already stated, the derivation of a Lyapunov function for these systems is limited to the case of no processing delays or with increased connectivity and they shed no light on the rate of convergence.

On the other hand, the solution operator is usually obtained with a stability in variation technique and it is an essentially metric-free process. Metrics are used in the step to prove that the operator is a contraction and they are of course of utmost importance. The whole process allows for the investigator to attain an overall control of the dynamic behavior of the system. Additionally, in every step of the argument important information are extracted such as the consensus point and the rate of convergence. The latter one is a function that depends explicitly on the systems parameters.

Disadvantages. FPT methods generally appear to provide a lot information and they are able to handle distributed systems in utmost generality. At the same time, they ask for a lot, in terms of computations and analysis. For the sake of justice, we should mention a number of drawbacks on account of this approach.

One notices that FPT is a lengthy method requiring several different steps, each of which involves tedious algebraic calculations. The deeper we are willing to dig to sharpen our results, the more calculations we are forced to make. In particular, the process of deriving the contraction property can be very painful exactly because the method will take any supplementary information the investigator is willing to provide. This difficulty lies in the heart of distributed algorithms exactly because a global governing kernel does not exist, other than this of the completely undelayed linear time-invariant model which Algebraic Graph Theory provides.

Although our result is stated in utmost generality, we encounter serious difficulties in proving delay independence results when we neglect processing delays. There are different factors that contribute to this situation. The first factor comes from the solution operator. We effectively considered (4.24) as a perturbation to the original un-delayed system. It is only reasonable then to expect delay dependent results (see also [7]). On the other hand, the dimensionality of the problem forced us to make use of the maximum norm. Under this norm one encounters the overall dynamics from the point of view of the local dynamical behavior of an arbitrary agent. This is a serious defect which we managed to partially remedy by introducing the weights $\mathbf{q}$. We tested the results of this approach in the Examples 4.3.11, 4.3.15 and 4.3.16 where we managed to obtain semi-delay independent stability. Unfortunately, $\mathbf{q}$ imposes additional computational complexity to the problem. It is intuitively clear that the more asymmetrical a system is the more difficult its analysis becomes. Example 4.3.17 clearly illustrates that any such asymmetry must be compensated with smaller and smaller maximum allowed delay.

4.4 General Linear Delayed Networks

The drawbacks of the fixed point methods aggravate when we consider more general systems. For this, we follow the approach developed in § 4.2.4. For this we will need Theorem 2.3.4. For fixed $t_0 \in \mathbb{R}$, we consider the initial value problem

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_j a_{ij}(t)\big(x_j(t - \tau_{ij}(t)) - x_i(t)\big), & t \geq t_0 \\ x_i(t) = \phi_i(t), & t \in I_{t_0} \end{cases} \tag{4.43}$$

where $\boldsymbol{\phi} = \big(\phi_1(t), \ldots, \phi_N(t)\big) \in L^1(I_{t_0}, \mathbb{R}^N)$ are given initial data a_{ij} is the strength at which node j affects i and $\tau_{ij}(t)$ the imposed delay between with which i receives the state of j at time t. The solution of (4.43), $\mathbf{x}(t, t_0, \boldsymbol{\phi}) = \big(x_1(t), \ldots x_N(t)\big)$ is an absolutely continuous vector valued function that defined in $[\lambda(t_0), \infty]$ and it takes values in $\mathbb{R}^N$ as in the undelayed case in (3.3) of § 3.2. It is noted that, contrary to the linear time-invariant case, here we neglect any processing delays. We recall the discussion in § 2.1 and especially the notation introduced in the beginning of this chapter.

The assumptions to accompany (4.43) are now stated. Regarding the connectivity weights the assumptions are identical to § 3.2. We put them here again for quick reference.

Assumption 4.4.1. *The connectivity weights a_{ij} are upper bounded, right continuous, non-negative functions of time.*

For the sake of simplicity we exclude the non-uniform cases discussed in § 3.2 from the general theory. We will return in such a scenario investigating an interesting case of unbounded delays for a 2×2 network. We recall the assumption of recurrent connectivity and switching communications in § 3.2 to state the next hypothesis:

Assumption 4.4.2. *For any $t \geq t_0$ there exists $\epsilon > 0$ independent of t such that $a_{ij}(t) \neq 0$ implies that there exist a neighborhood of t, $U_t \subset I_{[t_0,\infty)}$ of length ϵ such that $a_{ij}(s) \geq \underline{f} > 0$ for any $s \in U_t$.*

Assumption 4.4.3. *The delay functions $\tau_{ij} \in C^1([t_0, \infty), \mathbb{R}_+)$ for all $i \neq j$ so that*

$t - \tau(t) = \lambda(t) \to \infty$ *as* $t \to \infty$ *and*

$$\sup_{t \geq t_0} \int_{\lambda(t)}^{t} a_{ij}(s)\, ds < \infty \qquad \forall\, j \in N_i,\ i \in [N].$$

The last condition on Assumption 4.4.3 implies that whenever a_{ij} is bounded from below then $\tau(t)$ is necessarily bounded from above. From Assumption 4.4.2 we conclude that this will be the central case.

Bounds on $\mathbf{x}(t, t_0, \boldsymbol{\phi})$. The following two technical lemmas are instrumental in the analysis to follow for both the solution of (4.43).

Lemma 4.4.4. *Under Assumption 4.4.1, for any* $t_2 \geq t_1 \geq t_0$*, the solution* $\mathbf{x}\big(t_2, t_1, \mathbf{x}_{t_1}\big)$ *of* (4.43) *satisfies*

$$x_i(t_2) \in W_{I_{t_1}, \mathbf{x}_{t_1}}, \forall i \in [N]$$

Proof. Let $t^* \geq t_1$ be the first time that $x_i(t)$ escapes $W_{I_{t_1}, \mathbf{x}}$ say, to the right. Then it must hold both that $x_i(t^*) = \min_{j \in [N], s \in [\lambda(t_1), t_1]} x_j(s)$ and $\dot{x}_i(t^*) > 0$, which is a contradiction in view of the dynamics in (4.43). The same argument can be made for escaping to the left. $\square$

Lemma 4.4.5. *If* $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi})$, $t \geq t_0$ *is the solution of* (4.43) *with the property that* $S_{I_t}\big(\mathbf{x}\big) \to 0$ *as* $t \to \infty$*, then the forward limit set* $\omega(\boldsymbol{\phi})$ *is a singleton with a point in* Δ_I.

Proof. From Lemma 4.4.4 we have that $\omega(\boldsymbol{\phi})$ is non-empty, compact and connected and any element of which must lie in Δ. Since $\lambda(t_0) < \infty$ any point $\boldsymbol{\phi}^{\omega} \in \omega(\boldsymbol{\phi})$ is actually a vector valued function with the property that $\phi_i^{\omega}(s) = \phi_j^{\omega}(s)\ \forall i, j \in [N]$.

It is obvious however that $\mathbf{x}(t, t_0, \boldsymbol{\phi}^\omega) \equiv \boldsymbol{\phi}^\omega(t_0)$ and at the same time a member of $\omega(\boldsymbol{\phi})$. By the uniqueness of solutions it follows that $\boldsymbol{\phi}^\omega$ must be a constant vector valued function in $\mathbb{R}^N \cap \Delta$ and the result follows. $\square$

Lemma 4.4.4 and Lemma 4.4.5 are the time-delayed analogues of Lemma 3.2.2 and in Lemma 3.2.3 of § 3.2, respectively.

The idea here, relies on the elementary observation that for any $t \geq t' \geq t_0$ the solution $\mathbf{x}(t) = \big(x_1(t), \ldots, x_N(t)\big)$ satisfies

$$x_i(t) = e^{-\int_{t'}^t d_i(s)\,ds} x_i(t') + \int_{t'}^t e^{-\int_s^t d_i(w)\,dw} \sum_j a_{ij}(s) x_j(\lambda_{ij}(s))\,ds, \tag{4.44}$$

and the results are expected to reveal exponential type of convergence. We will study the rate at which $S_{I_t}(\mathbf{x})$ contracts by combining (4.44) with Theorem 2.3.4 and Lemma 4.4.4. Revealing the way the solution function contracts over intervals of time, we conclude that the limit point must be in Δ_I, i.e. the solutions contract to a constant, just like the un-delayed case.

Theorem 4.4.6. *Let Assumptions 4.4.1, 4.4.2 and 4.4.3 hold such that* $a_{ij}(t) \neq 0$ *implies* $a_{ij}(t) \geq \underline{f} > 0$ *uniformly in* t *and* $\sup_{t \geq t_0} \tau(t) \leq \tau < \infty$. *If there exists* $B > 0$ *so that for any* $t \geq t_0$ *the graph* $\mathbb{G}_{P_B(t)}$ *is routed-out branching, then unconditional asymptotic consensus for the solution* $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi})$, $t \geq t_0$ *of* (4.43) *is achieved.*

In particular, there exists $k \in W_{I_{t_0}, \boldsymbol{\phi}}$ *such that*

$$\max_i |x_i(t) - k| \leq \frac{S_{I_{t_0}}(\boldsymbol{\phi})}{1 - \underline{\kappa} e^{-\bar{N} a \tau}} e^{-\theta(t - t_0)}$$

where $\theta = -\frac{\ln(1 - \underline{\kappa} e^{-\bar{N} a \tau})}{\sigma B + 2\tau}$, $\underline{\kappa} := \inf_{t \geq t_0} \min_{i,j} \sum_l \min\{p_{il}^t, p_{jl}^t\} \in (0, 1)$ *with* p_{ij}^t *the elements of* $P_B^{(\sigma)}(t)$, $\sigma = l^*([N/2] + 1)$, $\bar{N} = \max_{j \in [N]} N_j$ *is the maximum degree* *over $[N]$ and l^* has the meaning of Remark 2.3.7.*

Theorem 4.4.6 provides explicit estimates on the rate of convergence as a function of the parameters, the connectivity signal and the imposed delays. This is a delay-independent result and in the next section we will show that it can be extended to unbounded delays in view of Assumption 4.4.3.

Proof. Fix $t \geq t_0$. Under the imposed connectivity conditions, Proposition 3.2.5 implies that the stochastic matrix $P_B^{(\sigma)}(t)$ is scrambling. Then by Theorem 2.3.4 and Lemma 4.4.4 we have the estimate:

$$S(\mathbf{x}(t)) \leq (1 - \underline{\kappa}) S_{I_{t-\sigma B-\tau}}(\mathbf{x}) \tag{4.45}$$

for a strictly positive $\underline{\kappa} := \inf_{t \geq t_0} \min_{i,j} \sum_l \min\{p_{il}^t, p_{jl}^t\} < 1$ with p_{ij}^t the elements of $P_B^{(\sigma)}(t)$. As $S_{I_t}(\mathbf{x}) = \max_{i \in [N], s \in [\lambda(t), t]} x_i(s) - \min_{i \in [N], s \in [\lambda(t), t]} x_i(s)$ this leads us to consider $t_1, t_2 \in I_t$ and i, j such that indeed $S_{I_t}(\mathbf{x}) = x_i(t_1) - x_j(t_2)$. Assume without loss of generality that $t_1 \geq t_2$. Then from (4.44) and Lemma 4.4.4 we have

$$\begin{aligned}
&S_{I_t}(\mathbf{x}) = x_i(t_1) - x_j(t_2) \\
&= e^{-\int_{t_1}^{t_2} d_i(s)\,ds} \big(x_i(t_2) - x_j(t_2)\big) - \\
&\qquad\qquad - \int_{t_1}^{t_2} e^{-\int_s^t d_i(w)\,dw} \sum_j a_{ij}(s) \big(x_j(\lambda_{ij}(s)) - x_j(t_2)\big)\,ds \\
&\leq e^{-\int_{t_1}^{t_2} d_i(s)\,ds} S(\mathbf{x}(t_2)) + \big(1 - e^{-\int_{t_1}^{t_2} d_i(s)\,ds}\big) S_{I_{t-2\tau}}(\mathbf{x}) \\
&\leq e^{-\int_{t_1}^{t_2} d_i(s)\,ds} (1 - \underline{\kappa}) S_{I_{t-\sigma B-2\tau}}(\mathbf{x}) + \big(1 - e^{-\int_{t_1}^{t_2} d_i(s)\,ds}\big) S_{I_{t-2\tau}}(\mathbf{x}) \\
&\leq (1 - \underline{\kappa} e^{-\int_{t_1}^{t_2} d_i(s)\,ds}) S_{I_{t-\sigma B-2\tau}}(\mathbf{x}) \\
&\leq (1 - \underline{\kappa} e^{-\bar{N} a \tau}) S_{I_{t-\sigma B-2\tau}}(\mathbf{x})
\end{aligned}$$

Then for any $t \geq t_0 + \sigma B + 2\tau$ there exists $l \geq 1$ such that $t_0 + l(\sigma B + 2\tau) \leq t \leq t_0 + (l+1)(\sigma B + 2\tau)$ and so a recursive argument implies that

$$S_{I_t}(\mathbf{x}) \leq \frac{S_{I_{t_0}}(\boldsymbol{\phi})}{1 - \underline{\kappa} e^{-\bar{N} a \tau}} e^{\frac{\ln(1 - \underline{\kappa} e^{-\bar{N} a \tau})}{\sigma B + 2\tau}(t - t_0)}$$

and the proof is concluded in view of Lemma 4.4.5 which ensures that the limit set is a singleton, say k, and Lemma 4.4.4 which ensures that $k \in W_{t_0, \boldsymbol{\phi}}$. $\square$

4.4.1 A leader-follower with delays

Here we repeat the scenario of a leader-follower topology discussed in § 3.2.1. Let agent 1 with state z_1 to evolution according to the differential equation

$$\dot{z}_1(t) = g\big(t, z_1(t)\big)$$

The network dynamics is

$$i \in [N] : \begin{cases} \dot{z}_1(t) = g\big(t, z_1(t)\big), & t \geq t_0 \\ \dot{z}_i(t) = \sum_j a_{ij}(t)\big(z_j(t - \tau_{ij}(t)) - z_i(t)\big), i \neq 1, & t \geq t_0 \\ z_i(t) = \phi_i(t) & t \in I_{t_0} \end{cases} \tag{4.46}$$

where $I_{t_0} = [t_0 - \tau(t_0), t_0]$ and $\tau(t) = \max_{ij} \tau_{ij}(t)$, the usual notation. Again, the dynamics of the leader's state $z_1(t)$ are free of any interaction with the rest of the group and they satisfy (3.8) which we restate here for quick reference:

$$|z_1(t) - k| \leq \frac{1}{h(t)}$$

for some rate function $h(t)$. We write for $i \neq 1$

$$\dot{z}_i(t) = \sum_{j \neq 1} a_{ij}(t)\big(z_j(t - \tau_{ij}(t)) - z_i(t)\big) + a_{i1}\big(k - z_i(t)\big)$$
$$+ a_{i1}(t)\big(z_j(t - \tau_{ij}(t)) - k\big)$$

The result below, is the delayed alternative of Theorem 3.2.11.

Theorem 4.4.7. *Let the solution* $\mathbf{z} = \mathbf{z}(t, t_0, \boldsymbol{\phi}), t \geq t_0$ *of* (4.46) *and the dynamics of the leader together with* (3.8)*. Assume uniform lower bounds of the connectivity weights and the connectivity conditions of Theorem 4.4.6. If* $\theta > 0$ *is the rate exponent of the convergence of the system with the virtual leader and and if there exists a function* $c(t) : [t_0, \infty) \to (0, \infty)$ *with the properties that*

1. $c(t) \to \infty$ *as* $t \to \infty$

2. $\sup_{t \geq t_0} e^{-\theta(t-t_0)} c(t) < \infty$

3. $\sup_{t \geq t_0} c(t) \int_{t_0}^{t} \frac{e^{-\theta(t-s)}}{1-\kappa e^{-N\alpha\tau}} \max_i \frac{a_{i1}(s)}{h(s-\tau_{i1}(s))}\, ds < \infty$

Then there exists a constant K *such that*

$$||\mathbf{z}(t) - \mathbb{1}k||_\infty \leq \frac{K}{c(t)}. \tag{4.47}$$

Proof. Since (4.43), written as

$$\dot{\mathbf{x}} = -L(t)\mathbf{x}_t$$

has the solution $\mathbf{x}(t, t_0, \boldsymbol{\phi})$ that can be expressed as

$$\mathbf{x}(t, t_0, \boldsymbol{\phi}) = U(t, t_0)\boldsymbol{\phi}$$

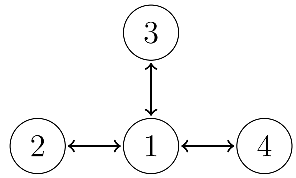

Figure 4.4: The communication topology in Example 7.1.1. The *star* graph.

for an operator $U(t, t_0)\boldsymbol{\phi} = x(t, t_0, \boldsymbol{\phi})$, linear and continuous so that the solution $\mathbf{z}(t, t_0, \boldsymbol{\phi})$ can be written as

$$\boldsymbol{z}(t, t_0, \boldsymbol{\phi}) = T(t, t_0)\boldsymbol{\phi} + \int_{t_0}^{t} T(t, s)\boldsymbol{\eta}(s)\, ds$$

where $\boldsymbol{\eta}(t) = (\eta_1(t), \ldots, \eta_N(t))^T$ with $\eta_i(t) = a_{i1}(t)\big(z(t - \tau_{i1}(t)) - k\big)$ (see also [68], Eq. (3.1.18)). The rest of steps are identical to the proof of Theorem 3.2.11 and are, therefore, omitted. $\square$

4.4.2 Examples and simulations

In this section, we discuss a couple of illustrative examples. The first is a 4×4 linear network with linear switching coupling and bounded delays. The second is a 2×2 time-varying network with static connectivity and unbounded delays. All simulations were carried out in MATLAB with the `ddesd` route.

4.4.2.1 A 4×4 graph

Let a network of $N = 4$ agents with communication weights $a_{ij}(t), i, j = 1, \ldots, 4$. We classify two communication schemes:

$$1 \longleftrightarrow 2 \longleftrightarrow 3 \longleftrightarrow 4$$

Figure 4.5: The communication topology in Example 7.1.2. The *path* graph.

Star topology Depicted in Figure 4.4, in this scheme the adjacency matrix reads

$$A_{\text{star}}(t) = \begin{bmatrix} 0 & a_{12}(t) & a_{13}(t) & a_{14}(t) \\ a_{21}(t) & 0 & 0 & 0 \\ a_{31}(t) & 0 & 0 & 0 \\ a_{41}(t) & 0 & 0 & 0 \end{bmatrix}$$

We assume here the switching (on/off) transmission signal to be defined as follows: There exist $B, \epsilon, \bar{f}, \underline{f} > 0$ such that for any B interval of time, there exists a ϵ-subset so that $\bar{f} \geq a_{ij}(s), a_{ji}(s) \geq \underline{f} > 0$ for $s \in [t, t+\epsilon]$. In this scenario, the analysis is very simple because Theorem 3.2.4 applies: Take $d_i(t) = \sum_j a_{ij}(t)$ and $m > \sup_t \max_{i \in [N]} d_i(t)$. Then $W(t) := mI_{4\times 4} - D(t) + A(t)$ is

$$W(t) = \begin{bmatrix} m - d_1(t) & a_{12}(t) & a_{13}(t) & a_{14}(t) \\ a_{21}(t) & m - d_2(t) & 0 & 0 \\ a_{31}(t) & 0 & m - d_3(t) & 0 \\ a_{41}(t) & 0 & 0 & m - d_4(t) \end{bmatrix}$$

and it is obviously scrambling during some ϵ interval over any B-interval of time. For m large enough the coefficient of ergodicity κ is bounded from below by $\underline{f}$ for $s \in [t, t+\epsilon]$. This implies that for any $t > t_0 + B$, there exists $n \geq 0$ such that $t_0 + nB \leq t \leq t_0 + (n+1)B$ and by Lemma 3.2.2

$$S(\mathbf{x}(t)) \leq S(\mathbf{x}_0) e^{\underline{f}\epsilon} e^{-(k+1)\underline{f}\epsilon} \leq S(\mathbf{x}_0) e^{-\frac{\underline{f}\epsilon}{B}(t-t_0)}.$$

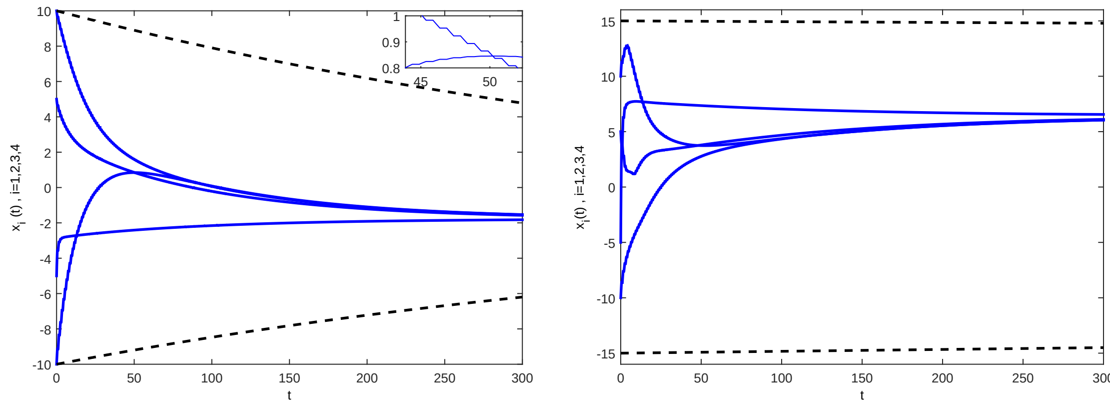

Figure 4.6: Example 7.1.1. (a): The convergence of 4 agents with a star graph topology. The detail on the upper right part is to denote the effect of the switching signal. (b): Convergence under the effect of delays. In both figures the dashed lines depict the theoretical rate estimates. It is remarked that in (b) the estimate is significantly weak.

In the presence of delays, $\tau_{ij}(t) \leq \tau < \infty$ the discussion in Sect. 4.4 applies. We can easily calculate $\underline{\kappa} = \min\{e^{-mB}, \frac{1-e^{-m\epsilon}}{m}\underline{f}\}$ and by Theorem 4.4.6 exponential convergence is guaranteed with rate $\theta = -\frac{\ln(1-\underline{\kappa}e^{-3\bar{f}\tau})}{B+2\tau}$.

As a numerical example take $a_{12}(t) = 0.02u(t)$, $a_{13}(t) = 0.05u(t)$, $a_{14}(t) = 0.03u(t)$, $a_{21}(t) = 0.2(0.01 + e^{-t})u(t)$, $a_{31}(t) = 0.07u(t)$,$a_{41}(t) = 0.06u(t)$ for $u(t) = 2, t \in [n + 1/2, n + 1]$ $n \in \mathbb{N}$ and 0 otherwise. Then $\frac{f\epsilon}{B} = 0.002$ and this is the estimated rate of convergence for the un-delayed system. See Fig. 4.6(a). In the case of delays we set $\tau_{12}(t) = 8 - 0.5\cos(t)$, $\tau_{13}(t) = 3 - 0.2\sin(2t)$, $\tau_{14}(t) = 9$, $\tau_{21}(t) = 10$ $\tau_{31}(t) = 5 - 0.9\sin(t^2)$, $\tau_{41}(t) = 2$ so that $\tau = 10$, $\underline{\kappa} = 0.0019$ and $\theta = 0.00008031$. See Fig. 4.6(b).

Path topology   Depicted in 4.5, the adjacency matrix is

$$A_{\text{path}}(t) = \begin{bmatrix} 0 & a_{12}(t) & 0 & 0 \\ a_{21}(t) & 0 & a_{23}(t) & 0 \\ 0 & a_{32}(t) & 0 & a_{34}(t) \\ 0 & 0 & a_{43}(t) & 0 \end{bmatrix}$$

Now we consider the switching signal to be defined as follows: For all $t \geq 0$ it holds that $a_{ij}(t) \neq 0 \Rightarrow 0 < \underline{f} \leq a_{ij}(t) < \frac{1}{2}$ and also

$$\begin{cases} a_{23}(t) = a_{32}(t) = a_{34}(t) = a_{43}(t) = 0 \ \& \ a_{12}(t), a_{21}(t) \neq 0, t \in [3l\epsilon, (3l+1)\epsilon) \\ a_{12}(t) = a_{21}(t) = a_{34}(t) = a_{43}(t) = 0 \ \& \ a_{23}(t), a_{32}(t) \neq 0, t \in [(3l+1)\epsilon, (3l+2)\epsilon) \\ a_{23}(t) = a_{32}(t) = a_{12}(t) = a_{21}(t) = 0 \ \& \ a_{34}(t), a_{43}(t) \neq 0, t \in [(3l+2)\epsilon, (3l+3)\epsilon) \end{cases}$$

for some fixed $\epsilon > 0$ and $l \in \mathbb{Z}_+$. Here $B = 3\epsilon$, $m = 1$ and

$$C(t,s) = \begin{bmatrix} \bar{d}_1(t,s) & e^{-(t-s)}a_{12}(s) & 0 & 0 \\ e^{-(t-s)}a_{21}(s) & \bar{d}_2(t,s) & e^{-(t-s)}a_{23}(s) & 0 \\ 0 & e^{-(t-s)}a_{32}(s) & \bar{d}_3(t,s) & e^{-(t-s)}a_{34}(s) \\ 0 & 0 & a_{43}(s)e^{-(t-s)} & \bar{d}_4(t,s) \end{bmatrix}$$

where $\bar{d}_i(t,s) = e^{-3\epsilon}\delta(s-(t-3\epsilon)) + e^{-(t-s)}(1-d_i(s))$. This is a non-scrambling matrix so Theorem 3.2.4 is of no use and we need to escalate to Theorem 3.2.9 and especially to Corollary 3.2.10. From this we obtain

$$\begin{aligned} S(\mathbf{x}(t)) \leq S(\mathbf{x}(3(l-1)\epsilon)) &\leq (1-2\underline{f}^2(1-e^{-\epsilon})^2)^{l-1}S(\mathbf{x}(0)) \\ &\leq \frac{S(\mathbf{x}(0))}{1-2\underline{f}^2(1-e^{-\epsilon})^2}e^{-\theta t} \end{aligned}$$

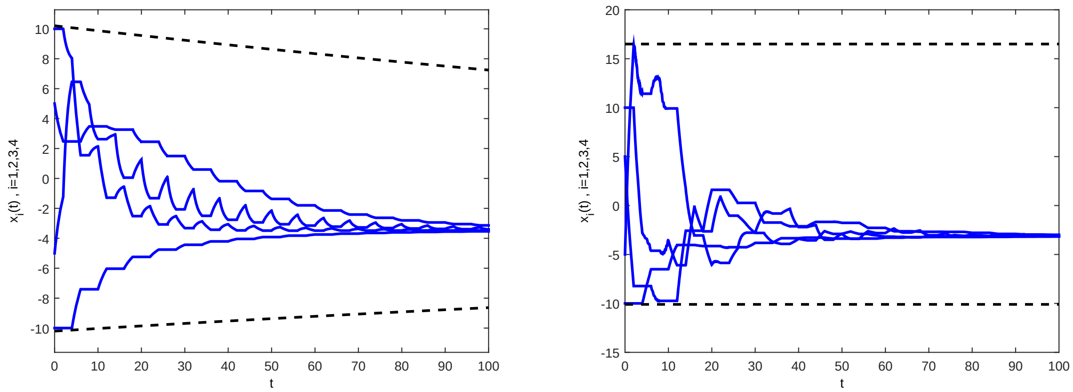

Figure 4.7: Example (a): The convergence of 4 agents with a path graph topology. (b): Convergence under the effect of delays. In both figures, the dashed lines depict the theoretical rate estimate. It is remarked that in (b) the estimate is significantly weak.

where $\theta := \frac{\ln(1-2\underline{f}^2(1-e^{-\epsilon})^2)}{3\epsilon}$. Together with the switching signal, we now consider a common bounded propagation delay $0 \leq \tau(t) \leq \tau < \infty$. We apply Theorem 4.4.6 to estimate the rate of convergence as follows: $\underline{\kappa} = 2\underline{f}^2(1-e^{-\epsilon})^2$, $\sup_t \int_{\lambda(t)}^t d_i(s)\,ds \leq \tau$ so that the rate of convergence with the delay is

$$\theta = \frac{\ln(1-2\alpha^2(1-e^{-\epsilon})^2 e^{-\tau})}{3\epsilon + 2\tau}.$$

As a numerical example, take $\epsilon = 2$, $B = 6$, $\underline{f} = 0.1$ and $a_{12}(t) = 0.2u(t)$, $a_{21}(t) = 0.3u(t)$, $a_{23}(t) = 0.21$, $a_{32}(t) = 0.2u(t)$, $a_{34}(t) = 0.25u(t)$ $a_{43}(t) = 0.1u(t)$ by an appropriate switching function $u(t)$. The rate of convergence is $\theta = 0.00251$, see Fig. 4.7(a). In the presence of the common delay with $\sup_{t\geq 0} \tau(t) = 10$ the rate is $\theta = 2.61 \cdot 10^{-8}$, see Fig. 4.7(b).

4.4.2.2 A 2×2 network with unbounded delays

Fix $t_0 > 0$ and consider the network of two agents to satisfy

$$\begin{cases} \dot{x}_1(t) = \frac{1}{\alpha t}\big(x_2(\beta t) - x_1(t)\big) \\ \\ \dot{x}_2(t) = \frac{1}{\gamma t}\big(x_1(\varepsilon t) - x_2(t)\big), t \geq t_0 \\ \\ \big(x_1(t), x_2(t)\big) = \big(\phi_1(t), \phi_2(t)\big), t \in (\beta t_0, t_0) \end{cases}$$

for some $\alpha, \gamma > 0$ and $\beta, \varepsilon \in (0,1)$. This system lies beyond the theory developed in the preceding sections. In fact, it only takes few elementary, yet tedious, modifications to include systems with unbounded delays. These are, in fact, easily illustrated for $N = 2$. Without loss of generality assume $\beta < \varepsilon$ and $\alpha > \gamma$. Now $\tau_1(t) = (1-\beta)t$ and $\tau_2(t) = (1-\varepsilon)(t)$ so that $\tau(t) = (1-\beta)t$. We work as follows: Firstly, we introduce the rate function $h(t) = \left(\frac{t}{t_0}\right)^\eta$ for $t > t_0$ and $\eta > 0$. It is easy to see that $x_1(t), x_2(t)$ satisfy the system of integral equations:

$$\begin{cases} x_1(t) = t^{-\eta} \int_{\lambda(t)}^t \left[\eta s^{\eta-1} - \frac{s^{\eta-1}}{\alpha} + s^\eta \delta(s - \lambda(t))\right] x_1(s)\, ds + \\ \qquad\qquad + t^{-\eta} \int_{\lambda(t)}^t \frac{s^{\eta-1}}{\alpha} x_2(\beta s)\, ds \\ \\ x_2(t) = t^{-\eta} \int_{\lambda(t)}^t \left[\eta s^{\eta-1} - \frac{s^{\eta-1}}{\gamma} + s^\eta \delta(s - \lambda(t))\right] x_2(s)\, ds + \\ \qquad\qquad + t^{-\eta} \int_{\lambda(t)}^t \frac{s^{\eta-1}}{\gamma} x_1(\varepsilon s)\, ds \end{cases}$$

It is easy to check that for small η the matrix

$$P = \begin{bmatrix} 1 - \frac{\eta}{\alpha}(1-\beta)^\eta & \frac{\eta}{\alpha}(1-\beta)^\eta \\ \\ \frac{\eta}{\gamma}(1-\beta)^\eta & 1 - \frac{\eta}{\gamma}(1-\beta)^\eta \end{bmatrix}$$

is stochastic and obviously scrambling. Then

$$\kappa = \min\left\{1 - \frac{\eta}{\alpha}(1-\beta)^\eta, \frac{\eta}{\gamma}(1-\beta)^\eta\right\} + \min\left\{\frac{\eta}{\alpha}(1-\beta)^\eta, 1 - \frac{\eta}{\gamma}(1-\beta)^\eta\right\} > 0$$

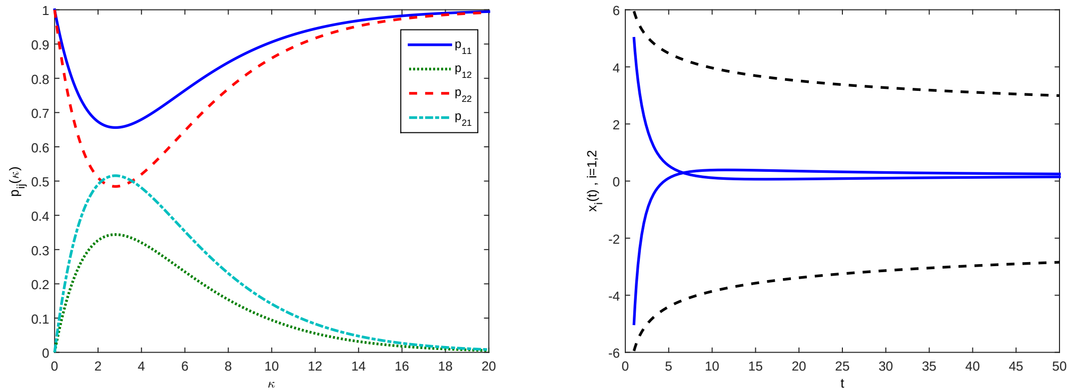

Figure 4.8: Example (a): Graphs of the elements $p_{ij}(\eta)$ for selected values of α, β, γ demonstrating the dependence of κ in η. (b) The convergence of the 2×2 static, time-varying network with unbounded delays.

Then similar analysis as in Theorem 2.3.4 and Lemma 4.4.4 yields:

$$S_{I_t}(\mathbf{x}) \leq (1 - \kappa e^{\frac{\ln \beta}{\alpha}}) S_{I_{\lambda^{(2)}(t)}}(\mathbf{x}) \leq (1 - \kappa \beta^{1/\alpha}) S_{I_{\lambda^{(2)}(t)}}(\mathbf{x})$$

Now as for any t there exists $l \in \mathbb{Z}_+$ such that $\lambda^{(2l)}(t) \leq \frac{t_0}{\beta^2}$ or equivalently $l \geq \frac{\ln(t_0/(t\beta^2))}{2\ln\beta}$. So

$$\max_i |x_i(t) - k| \leq Z S_{t_0}(\boldsymbol{\phi}) \left(\frac{t_0}{t}\right)^{\zeta}$$

where $Z = e^{\eta\beta^{1/\alpha}}$ and $\zeta = -\frac{\eta\beta^{1/\alpha}}{2\ln\beta} > 0$. The rate is sub-exponential exactly because of unbounded delays.

As a numerical example we take $t_0 = 1$, $\alpha = 3$, $\gamma = 2$, $\varepsilon = 0.5$, $\beta = 0.3$. Fig. 4.8(a) depicts the dependence of the elements of P as η varies. The selection of η determines the estimates of κ. If we take $\eta = 1.22$ we obtain $\kappa = 0.655$ and we calculate $Z = 1.953$ and $\zeta = 0.1821$. The simulation of the solution x_1, x_2 is presented in Fig. 4.8(b).

4.5 Supplementary Remarks

The setback of the naive fixed point approach on delayed consensus networks [3, 4, 5, 6, 7] drove me into a deeper study of the dynamic behavior in terms of stability in variation methodologies. The research was cut down to the simplest possible, scalar alternatives of these models, which are nothing but a special case of functions that sustain asymptotically constant solutions. Depending on the sign of the coupling we conclude on the range of the stability bound one is allowed to take for delay, thus unifying a large number of related results. For the case simple time-invariant parameters, almost complete results can be derived. The method of combining solution operators, however, limbs in the case of general linear dynamics [14]. Thus, we resorted to an simple Lyapunov-Razumikhin type of argument for stability in terms of function spaces re-affirming in a much less involved way past results [78]. The price to pay is that the latter estimates, although delay-independent, are significantly weaker than the time invariant ones. In fact, one of the major claims of [14] is that whenever the fixed point methods apply then the derived rate estimates are much stronger than the delay-independent Lyapunov-Razumikhin ones.

Both these methods are developed so as to be applied for the corresponding time-invariant and general linear multi-dimensional networks, respectively, in the last two sections.

The argument developed to prove asymptotic stability with respect to Δ follows the theory developed in § 3.2. Similarly, to the scalar alternative we provided delay-independent results with explicit estimates on the rate of convergence. If the weights are uniformly bounded from below, then it is necessary that the delays are bounded and the rate is exponential. We showed by example how unbounded processing delays can also be considered downgrading the estimates to sub-exponential. Similarly to the scalar case these rates not strong, as the examples suggest. This weakness is partially due to the use of the contraction coefficient as this is proven a weak estimator on its own.

The approach followed for the general linear consensus network with delays uses fundamentally different mathematical tools. What should be noted, though, is that the principle of combining two different forms of the solutions is again heavily exploited. The first form is the integral representation of the solution that outlines the local "convex averaging". This is (4.44), the time-varying equivalent of (4.30) in the case where the latter one is $\sigma = 0$. The second form of the solution is due to the coefficient of ergodicity (applied together with Lemma 4.4.4). This form is a time-varying equivalent of (4.31).

The stability results of this chapter build a theoretical framework for the analysis of the asymptotically constant solutions in a family of linear functional differential equations. They lay the ground for the analysis of nonlinear systems discussed in Chapters 5, 6 as well as the analysis of the application problem in Chapter 7.

Chapter 5: Nonlinear Networks

Distributed cooperative dynamics have been vividly stimulating the attention of the engineering and applied mathematics community for at least the past decade. The latest advances are nowadays past the linearity point. Recent works investigate nonlinear variations of consensus networks supplying the literature with a fruitful of impressive results. In the section to follow we will review a few of the most notable works on nonlinear 1^{st} order consensus schemes in the field.

Nonlinear versions of (0.1) exist in the literature primarily as extensions of the linear scheme, because they preserve its vital qualitative features [81, 82, 83, 84, 85]. In his seminal work [83], Moreau studied the generic system:

$$i \in [N] : \begin{cases} x_i(n+1) = f_i(n, x_1(n), \ldots, x_N(n)), & n \geq n_0, n \in \mathbb{Z}_+ \\ x_i(n_0) = x_i^0, & n = n_0 \end{cases} \tag{5.1}$$

where he built an asymptotic stability argument with respect to Δ, based on set-valued Lyapunov functions. He showed that agreement among agents emerges as $n \to \infty$, on condition that

$$x_i(n+1) \in \mathrm{co}\big\{x_1(n), \ldots, x_N(n)\big\},\ n \geq n_0.$$

The latter equation is telling us that each agent's new state lies in the interior of the set defined as the convex hull of their neighbors' current states. A different approach was adopted, by the same researcher, for the continuous time linear version of the (0.1) (see [51]). In that paper he provided a semi-rigorous proof for the rate of contraction of the spread $S\big(\mathbf{x}(t)\big)$. Finally, a fairly similar type of of distributed consensus is considered in [82].

The nonlinear nature of (5.1) is essentially *sine qua non.* Carathéodory's Theorem assures that

$$x_i(n+1) \in \mathrm{co}\big\{x_1(n), \dots, x_N(n)\big\} \;\Leftrightarrow\; x_i(n+1) = \sum_j a_{ij}(n)x_j(n)$$

where $a_{ij}(n) \geq 0,\ \sum_j a_{ij}(n) \equiv 1$ [37].

An alternative of the linear model is also obtained assuming nonlinear couplings on agents' state difference under passivity conditions. The authors in [81, 84] introduce and study the asymptotic properties of

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_j g_{ij}\big(t, x_j(t) - x_i(t)\big), t \geq t_0 \\ x_i(t_0) = x_i^0 \end{cases} \tag{5.2}$$

with $g_{ij}(t, z)$ being a passive function in z (see Assumption 5.1.1 below). This is an interesting extension of (0.1) as the form $g_{ij}(t, x_i - x_j)$ is quite general. Indeed, it bears great similarities with the prominent Kuramoto model [86] for synchronization of oscillators as well as Krause's opinion dynamics model [53] with $g_{ij}(\mathbf{x}) = a(|x_i - x_j|)(x_j - x_i)$.

In [84] the authors analyze (5.2) with multiple constant propagation delays and a possibly switching communication network. They argue, employing Invariance Principles for functional differential equations, that asymptotic stability to Δ_I is achieved for delays of arbitrary magnitude. The adopted approach does not tackle the issue of the rate of convergence because Invariance Principles can only prove asymptotic convergence for the ordinary or the functional case [69, 87]. The authors note this issue and its importance, suggesting that the question of the rate of convergence for such systems remains open.

A different type of linear extension is presented in [85]. The authors introduce the system

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_j a_{ij}(t)\big(g_{ij}(x_j(t)) - g_{ij}(x_i(t))\big), t \geq t_0 \\ \\ x_i(t_0) = x_i^0 \end{cases} \tag{5.3}$$

as an extension to [45] and prove convergence including both static and switching connectivity conditions. Their work relies on Lyapunov stability and Invariance Principles under a prescribed functional, thus strong assumptions on g_{ij} had to be considered.

All the aforementioned works as well as the models discussed in this thesis so far, study distributed algorithms under the instrumental assumption that the rate of change of the state of an agent $i \in [N]$ strictly depends on the agents current state. This is an assumption that although mathematically important it is over-simplistic for a number of reasons.

In real world applications the agents ability to operate cannot exclusively depend on its current state. Robots have terminals that may take some time to keep processing data after a while, or birds may get tired after maneuvering way beyond their physical abilities. This phenomenon is very important as it affects both the performance and the stability of the network. The mathematical equations now read as functional differential equations of neutral type. To the best of our knowledge, there is no work towards this path in the theory of consensus systems and for good reason. For one, the classical ordinary differential equation theory is no longer applicable and one needs to switch to the theory of functional differential equations [59]. Although the theory of neutral equations has been fully developed, the stability tools are by no means as strong as the ones used in the ordinary (or even the functional case) let alone when we are focused in the Lagrange type stability, the consensus systems enjoy (i.e. stability with respect to a subset of the state space).

Contribution At first, we revisit and discuss (5.2) and (5.3) under our framework. We make the case that, with fairly mild assumptions, these classes of systems can be effectively studied within the theory developed in Chapters 3 and 4 so that explicit estimates about the rate of convergence can be provided.

In addition, we comment on how the autonomous version of (5.2), yields also information on the consensus point in the sense of (4.27) in §4.3. Furthemore, we will explain how the effect of delays in (5.2) may yield the existence of non-trivial periodic solutions and we provide sufficient conditions for asymptotic stability of these solutions in a synchronized manner. We outline these results with an illustrative simulation.

Finally, we introduce and analyze consensus networks of neutral type by means of fixed point theory. We provide sufficient conditions for asymptotic convergence with prescribed rate to an implicitly defined point that is a function of the network topology and the imposed delays.

For the rest of this chapter we recall the discussion in §2.1 and §4.2.2 and the notation used therein.

5.1 The Passivity Hypothesis

The system presented in (5.2) is perhaps the one closest to the linear model. Indeed a simple transformation, known as *direct linearization*, will reveal that we one can readily deal with the linear models. We will work with the delayed versions of these systems because the un-delayed version is only a special case.

A network of N agents exchanges information according to:

$$i \in [N] : \begin{cases} \dot{x}_i(t) = \sum_{j \in N_i} g_{ij}\big(t, x_j(\lambda_{ij}(t)) - x_i(t)\big), & t \geq t_0 \\ \\ x_i(t) = \phi_i(t), & t \in I_{t_0} \end{cases} \tag{5.4}$$

For any $t \geq t_0$ there may or may not exist a connection between j and i. This defines a connectivity regime that can be described by a graph $\mathbb{G}_g(t) = (V, E(t))$ with $(i,j) \in E(t)$ if and only if $g_{ij}(t, \cdot) \neq 0$. The equations assume no self-loops, i.e. $g_{ii} \equiv 0$. The passivity condition for g_{ij} is summarized next.

Assumption 5.1.1. *For any $i, j \in [N] : (i,j) \in E(t)$, $g_{ij}(t,x) : [t_0, \infty) \times \mathbb{R} \to \mathbb{R}$ is continuous in x, right-continuous in t and for $t \geq t_0$ it satisfies the following properties:*

1. $g_{ij}(\cdot, x) : [t_0, \infty) \to [0, g)$ *uniformly in* x,

2. $g_{ij}(t, 0) = 0$ *for any* $t \geq t_0$,

3. $g_{ij}(t, \cdot) \neq 0 \Rightarrow \frac{g_{ij}(t,x)}{x} > 0, \; \forall\; x \neq 0$ *, uniformly in* t,

4. $g_{ij}(t,\cdot) \neq 0 \Rightarrow \lim_{x\to 0} \frac{g_{ij}(t,x)}{x} \in \mathbb{R}_+$ *independent of* t.

The form of g_{ij} sustains the two most crucial features of the linear consensus scheme: The first is that, by construction, g_{ij} are compatible with the previously discussed connectivity regimes (switching connectivity) and the second is the passivity property which makes the solutions to behave in a qualitative identical way, as the following technical Lemma shows:

Lemma 5.1.2. *Let Assumption 5.1.1 hold. For* $t_2 \geq t_1 \geq t_0$ *the solution* $\mathbf{x}(t_2, t_1, \mathbf{x}_{t_1})$ *of* (5.4) *satisfies* $\mathbf{x}(t_2) \in W_{I_{t_1}, \mathbf{x}_{t_1}}$.

Proof. Let $t^* > t_0$ be the first time that the solution $\mathbf{x}(t, t_0, \boldsymbol{\phi})$ of (5.4) escapes $W_{I_{t_1}, \mathbf{x}_{t_1}}$. Then there exists $i \in [N]$ such that

$$x_i(t^*) = \max_{j\in[N]} \max_{s\in I_{t_1}} x_j(s) \qquad \& \qquad \dot{x}_i(t^*) > 0,$$

a contradiction, according to Assumption 5.1.1. Similarly, for the lower bound of $W_{I_{t_1}, \mathbf{x}_{t_1}}$. □

The result of this section is, in fact, so straightforward that it will be stated as a Corollary to Theorem 4.4.6.

Corollary 5.1.3. *Consider the initial value problem* (5.4) *and let Assumption 5.1.1 hold. The solution* $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi}),\ t \geq t_0$ *of* (5.4) *satisfies*

$$x(t) \to \Delta_I \ \ as \ \ t \to \infty$$

exponentially fast if $\mathbb{G}_g(t)$ *satisfies the conditions of Theorem 4.4.6.*

Proof. The passivity assumption obviously ensures that the solution of (5.4) exists in the large. Let $\mathbf{x}(t, t_0, \boldsymbol{\phi}),\ t \geq t_0$ be the fixed solution of (5.4). Define

$$a_{ij}(t) := \frac{g_{ij}(t, x_j(\lambda_{ij}(t)) - x_i(t))}{x_j(\lambda_{ij}(t)) - x_i(t)}$$

and rewrite the initial value problem (5.4) as

$$i \in [N] : \begin{cases} \dot{y}_i(t) = \sum_j a_{ij}(t)\big(y_j(\lambda_{ij}(t)) - y_i(t)\big), & t \geq t_0 \\ y_i(t) = \phi_i(t), & t \in I_{t_0} \end{cases}$$

so that the solutions $\mathbf{y}$ and $\mathbf{x}$ are indistinguishable. Then one can study the behavior of $\mathbf{y}$ to conclude about $\mathbf{x}$. This is known in the literature as direct linearization technique (see [34]). Define $U := [-S_{I_{t_0}}(\boldsymbol{\phi}), S_{I_{t_0}}(\boldsymbol{\phi})]$. Under Assumption 5.1.1 we see that there exists a lower bound $\inf_{t \geq t_0} \min_{i,j \in [N]} \min_{x \in U} \frac{g(t,x)}{x} \geq \underline{f} > 0$ and that $a_{ij}(t)$ satisfies

$$\min_{i,j \in [N]} \min_{x \in U} \frac{g(t,x)}{x} \leq a_{ij}(t) \leq \max_{i,j \in [N]} \max_{x \in U} \frac{g(t,x)}{x}$$

that it is also bounded from above. Theorem 4.4.6 readily applies to prove exponential convergence. $\square$

Explicit estimates can now be provided following Theorem 4.4.6 with the lower and upper bound of a_{ij} defined in the proof of the Corollary 5.1.3. We finally remark that we provided unconditional consensus results assuming that g_{ij} satisfy the passivity properties globally. Since this is may not be usually the case, if g_{ij} are passive in a common, non-empty subset of $\mathbb{R}$, then initial functions ϕ_i that are restricted in this subset guarantee convergence to consensus. This is a conditional type of consensus as it is heavily based on the initial data. A standard example is this when $g_{ij}(x) = \sin(x)$ as this function is passive in $(-\pi, \pi)$ and the results of this section hold if $\phi_i(t) \in (-\frac{\pi}{2}, \frac{\pi}{2}),\ \forall t \in I_{t_0}, i \in [N]$.

5.2 General Nonlinearities

The next category is the one that considers the following scenario: Each agent observes both a nonlinear and delayed view of its neighboring agents' state. Such a model updates dynamically the states of the agents as follows:

$$i \in [N] : \begin{cases} \dot{x}_i = \sum_j g_{ij}\big(t, x_j(\lambda_{ij}(t))\big) - g_{ij}\big(t, x_i(t)\big), & t \geq t_0 \\ x_i(t) = \phi_i(t), & t \in I_{t_0} \end{cases} \tag{5.5}$$

The long-term behavior of the solutions of (5.5) can be treated with similar techniques, under appropriate conditions. Indeed, Assumption 5.2.4 (stated below) makes all the linear theory tools useful, after being slightly adapted. Then stability of the solutions of (5.5) can be proved along the same lines. This will be the subject of §5.2.2. Otherwise, we will resort to stability in variation techniques and heavy assumptions to prove similar results. This will be the case presented right below.

5.2.1 Stability in variation

To the best of our knowledge, there is no appropriate framework for general nonlinear consensus systems. If we are dealing with a system similar to (5.5) but we have no information on g_{ij} other than a conventional growth estimate. The analysis may proceed with a stability in variation argument. Here we will provide such a result for the ordinary system, i.e.

$$\lambda_{ij}(t) \equiv t \tag{5.6}$$

Next, we assume the following condition on the g_{ij} functions:

Assumption 5.2.1. *For any $t \geq t_0$, $i, j \in [N]$ there exist non-negative integrable functions $a_{ij}(t, x)$ and $k_{ij}(t)$ such that for all $x_1, x_2 \in \mathbb{R}$*

$$\left|\big(a_{ij}(t, x_1) - g_{ij}(t, x_1)\big) - \big(a_{ij}(t, x_2) - g_{ij}(t, x_2)\big)\right| \leq k_{ij}(t)|x_1 - x_2|.$$

We understand that our purpose is to approximate (5.5) with the linear system (3.3). In addition, the connectivity scheme for can be relaxed to match the Type II connectivity that characterizes Theorem 3.2.9.

Theorem 5.2.2. *Let Assumptions 3.2.1, 3.2.7, 3.2.8 and 5.2.1 hold. Suppose that there is a rate function $w(t)$ such that*

$$\sup_{t \geq t_0} w(t)h(t, t_0) < \infty \qquad \& \qquad \sup_{t \geq t_0} w(t) \int_{t_0}^{t} h(t, s)\frac{k(s)}{w(s)}\, ds \leq \beta < 1,$$

where $h(t, t_0)$ is the rate function with which the linear system (4.43) *in §4.4, Chapter 4 converges to Δ and $k(t) = \max_{i,j} \sum_l k_{il}(t) + k_{jl}(t)$. If $h(\cdot, t_0) \in L^1_{[t_0,\infty)}$ then the solution $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi})$, $t \geq t_0$ of* (5.5) *satisfies*

$$\mathbf{x}(t) \to \Delta, \quad t \to \infty$$

as fast as $1/w(t)$.

Proof. We set $\tilde{\mathbf{G}}(t, \mathbf{x}) = \big[\tilde{g}_1(t, \mathbf{x}), \ldots, \tilde{g}_N(t, \mathbf{x})\big]^T$ where

$$\tilde{\mathbf{G}}(t, \mathbf{x}) := \left[\sum_j a_{1j}(t, x_j) - g_{1j}(t, x_1), \ldots, \sum_j a_{Nj}(t, x_j) - g_{Nj}(t, x_N)\right]^T$$

We add and sub-tract (5.5) and write it in vector form as follows:

$$\dot{\mathbf{x}} = -L(t)\mathbf{x} + \tilde{\mathbf{G}}(t, \mathbf{x})$$

The classic variation of constants formula implies

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}^0 + \int_{t_0}^{t} \Phi(t, s)\tilde{\mathbf{G}}(s, \mathbf{x}(s))\, ds \tag{5.7}$$

Now, $\mathbf{y}(t) = \Phi(t, t_0)\mathbf{x}^0$ is the solution of (3.3) for which we know that under the imposed conditions that have $\Phi(t, t_0)\mathbf{x}^0 \to \boldsymbol{\alpha}^T\mathbf{x}^0$ for some $\boldsymbol{\alpha} \in \mathbb{R}^N$ with the properties that $\alpha_i \geq 0$, $\sum_i \alpha_i = 1$. In terms of the semi-norm $S(\cdot)$, there exists a bounded rate function $h(t, t_0)$ such that $h \in L^1_{[t_0,\infty)}$ for any fixed $t_0 \in \mathbb{R}$:

$$S\big(\mathbf{y}(t)\big) = S\big(\Phi(t, t_0)\mathbf{x}^0\big) \leq h(t, t_0)S(\mathbf{x}^0) \tag{5.8}$$

Under Assumption 5.2.1 we obtain $S\big(\tilde{\mathbf{G}}(t, \mathbf{x})\big) \leq k(t)S\big(\mathbf{x}\big)$ where

$$k(t) = \max_{i,j} \sum_l k_{il}(t) + k_{jl}(t) \tag{5.9}$$

Now we will express a variation of constants formula for $S\big(\mathbf{x}(t)\big)$ following the proof of Theorem 1.3.1 in [68]. Consider the spread $S\big(\mathbf{y}(t, s, \mathbf{x}(s))\big)$ for $t_0 \leq s \leq t$ so that for all s, the right Dini derivative yields

$$\begin{aligned} \frac{d}{ds}S\big(\mathbf{y}(t, s, \mathbf{x}(s))\big) &= S_{\mathbf{y}}\big(\mathbf{y}(t, s, \mathbf{x}(s))\big)^T \left[\frac{\partial}{\partial t_0}\mathbf{y}(t, s, \mathbf{x}(s)) + \frac{\partial}{\partial \mathbf{y}^0}\mathbf{y}(t, s, \mathbf{x}(s))\dot{\mathbf{x}}(s)\right] \\ &= S_{\mathbf{y}}\big(\mathbf{y}(t, s, \mathbf{x}(s))\big)^T \Phi(t, s)\tilde{\mathbf{G}}(s, \mathbf{x}(s)) \end{aligned}$$

from the basic property of the transition matrix $\frac{\partial}{\partial s}\mathbf{y}(t, s, \mathbf{x}(s)) = \Phi(t, s)L(s)\mathbf{x}(s)$

$$\frac{\partial}{\partial \mathbf{y}^0}\mathbf{y}(t, s, \mathbf{x}(s))\dot{\mathbf{x}}(s) = \Phi(t, s)\big(-L(s)\mathbf{x}(s) + \tilde{\mathbf{G}}(s, \mathbf{x}(s))\big)$$

Integrating from t_0 to t we finally obtain

$$S\big(\mathbf{x}(t)\big) = S\big(\mathbf{y}(t,t_0,\mathbf{x}^0)\big) + \int_{t_0}^{t} S_{\mathbf{y}}\big(\mathbf{y}(t,s,\mathbf{x}(s))\big)^T \Phi(t,s)\tilde{\mathbf{G}}(s,\mathbf{x}(s))\, ds$$

For any fixed t, s a bound for the integrand can be calculated to be:

$$\begin{aligned} S_{\mathbf{y}}\big(\mathbf{y}(t,s,\mathbf{x}(s))\big)^T \Phi(t,s)\tilde{\mathbf{G}}(s,\mathbf{x}(s)) &\le \max_{h,h'} \sum_j \big(\phi_{hj}(t,s) - \phi_{h'j}(t,s)\big)\tilde{g}_j\big(s,\mathbf{x}(s)\big) \\ &\le h(t,s)k(s)S\big(\mathbf{x}(s)\big) \end{aligned}$$

according to Eqs. (5.8) and (5.9). Consequently,

$$S\big(\mathbf{x}(t)\big) \le h(t,t_0)S(\mathbf{x}^0) + \int_{t_0}^{t} h(t,s)k(s)S\big(\mathbf{x}(s)\big)\, ds$$

This inequality implies that $S\big(\mathbf{x}(t)\big)$ is in fact bounded from above by $q(t)$, which satisfies the integral equation

$$q(t) = h(t,t_0)q(t_0) + \int_{t_0}^{t} h(t,s)k(s)q(s)\, ds, \qquad q(t_0) = S(\mathbf{x}^0) \tag{5.10}$$

We study the stability of (5.10) with respect to zero via a fixed point theory argument. Recall the discussion in §2.5 and consider the space

$$\mathbb{M} = \{z \in C^0[(t_0,\infty),\mathbb{R}] : z(t_0) = S(\mathbf{x}^0),\ \sup_{t\ge t_0} w(t)|z(t)| < \infty\}$$

which together with the weighted metric $\rho(y_1, y_2) = \sup_{t\ge t_0} w(t)|z_1(t) - z_2(t)|$ constitute a weighted complete metric space [34]. In this space we will apply Theorem 2.5.6 as follows: Define the operator

$$(\mathcal{Q}z)(t) := \begin{cases} S(\mathbf{x}^0), & t = t_0 \\ h(t,t_0)S(\mathbf{x}^0) + \int_{t_0}^{t} h(t,s)k(s)z(s)\, ds, & t \ge t_0 \end{cases}$$

and note that under for any $z \in \mathbb{M}$, $(\mathcal{Q}z)(t) \to 0$ as the first term vanishes by the imposed conditions and the second term vanishes as the convolution of an L^1 function with a function that goes to zero. The same holds for the weighted quantity $w(t)|(\mathcal{Q}z)(t)|$ in view of the imposed conditions. It is, finally, easy to see that $\mathcal{Q}$ is a contraction in $(\mathbb{M}, \rho)$ since

$$\rho(\mathcal{Q}z_1, \mathcal{Q}z_2) \leq \sup_{t \geq t_0} w(t) \int_{t_0}^{t} h(t,s) \frac{k(s)}{w(s)} \, ds \rho(y_1, y_2) \leq \alpha \rho(z_1, z_2).$$

Theorem 2.5.6 then ensures that $\mathcal{Q}$ attains a unique fixed point in $\mathbb{M}$ and the proof is concluded with argumentation similar to this of Lemma 3.2.3. □

Example 5.2.3. Consider the system

$$\begin{cases} \dot{x} = g(t,x) - g(t,y) \\ \\ \dot{y} = f(t,y) - f(t,x) \end{cases}$$

Theorem 5.2.2 ensures convergence to Δ if one can find positive numbers $a, b > 0$ such that

$$\sup_{t} \sup_{x,y} \frac{|a + g'(t,x)| + |b + f'(t,y)|}{a+b} < 1$$

Take, for instance,

$$g(t,x) = -k\big(2 + \sin(t)\big)x + \frac{x^2}{1+x^2}, \qquad f(t,x) = 2x + \cos(t)\frac{x^3}{1+x^2}$$

It can be easily verified that the above condition reads $\frac{|a-k|+b+2.89}{a+b} < 1$ and it is true for any $a > k > 2.89$. Numerical inspection for values of k is presented in Figure 5.1. We observe that instability occurs for small values of k.

The problem with non-monotonic systems is that there are no mathematical tools to effectively study their solutions. The theory at this point lacks strong

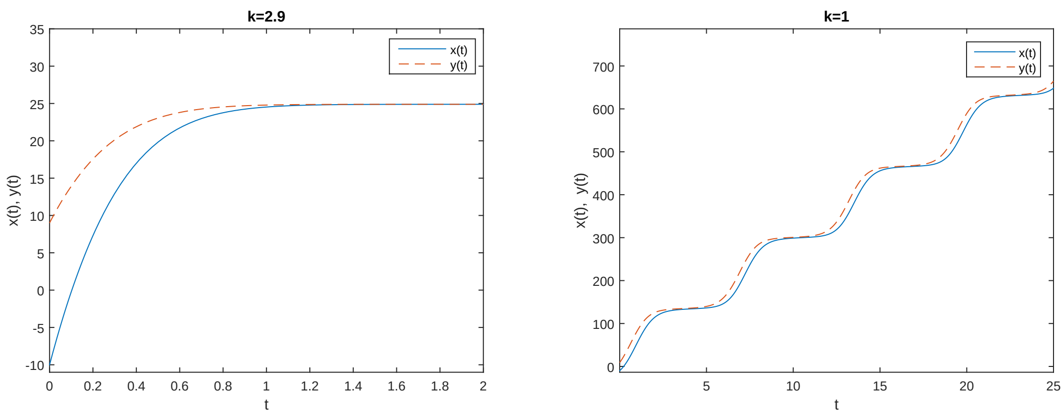

Figure 5.1: Example 5.2.3. Stability for large k and instability for small. The simulation is run with the `dde23` routine in `MATLAB`.

results. Indeed if $g_{ij} \in C^1$ then the mathematical toolbox provides frameworks only after distinguishing between the $\frac{\partial}{\partial x} g_{ij} > 0$ and $\frac{\partial}{\partial x} g_{ij} < 0$ [18]. The general case cannot be easily handled and that is why we resorted to stability in variation.

5.2.2 Monotonic Dynamics

The just mentioned approach is too restrictive because, heuristically speaking, we allowed "too much freedom" on the way g_{ij} is allowed to vary. However, the simple type of g_{ij}, mentioned just above, makes the dynamics of (5.5) not only to mimic those of the linear case but also it permits its solutions to be analyzed in a mathematically tractable way. In particular, we will ask the following *monotonicity* condition

Assumption 5.2.4. *For any* $t \geq t_0$ *and for any compact, connected* $M \subset \mathbb{R}$, $\exists\ \underline{c} < \bar{c} \in \mathbb{R}_+$ *that depend on* M *and integrable functions* $L_{ij} = L_{ij}^M, U_{ij} = U_{ij}^M$:

$[t_0, \infty) \to [\underline{c}, \bar{c}]$ *such that*

$$L_{ij}(t) \le \frac{g_{ij}(t,x) - g_{ij}(t,y)}{x-y} \le U_{ij}(t), \qquad \forall x, y \in M.$$

This assumption characterizes the functions g_{ij} and it is another generalization of the linear (3.2). A straightforward observation is that Assumption 5.2.4 implies uniform boundedness, hence existence in the large for the solutions of (5.5).

Lemma 5.2.5. *Let Assumption 5.2.4 hold. For any $t_2 \ge t_1 \ge t_0$, the solution $\mathbf{x}(t_2, t_1, \mathbf{x}_{t_1}) = \big(x_1(t_2), \dots, x_N(t_2)\big)$ of* (5.5) *satisfies $x_i(t_2) \in W_{I_{t_1}, \mathbf{x}_{t_1}}$.*

Proof of Lemma 5.2.5. Let $t^* \ge t_1$ to be the first time such that for some $i \in [N]$, $x_i(t^*) = \max_i \max_{s \in I_{t_1}} x_i(s)$ and $\dot{x}_i(t^*) > 0$. Then i attains the maximum state over the rest of the agents. By Assumption 5.2.4

$$g_{ij}\big(t^*, x_j(t^* - \tau_{ij}(t^*))\big) - g_{ij}\big(t^*, x_i(t^*)\big) \le L_{ij}^M(t^*)\big(x_j(t^* - \tau_{ij}(t^*)) - x_i(t^*)\big) \le 0$$

for all $j \in N_i$ so that $\dot{x}_i(t^*) \le 0$ and this is a contradiction in view of (5.5). A similar argument can be made for the first time the solution escapes $W_{I_{t_1}, \mathbf{x}_{t_1}}$ to the left. $\square$

By Lemma 5.2.5 one can set the parameters L_{ij}^M, U_{ij}^M based on the initial data $\boldsymbol{\phi}$ and in particular taking $M = W_{I_{t_0}, \boldsymbol{\phi}}$. Next we characterize the necessary connectivity regime which for the sake of simplicity it is considered of Type I:

Assumption 5.2.6. $\forall\, t \ge 0$ *the network is static and* $\exists\, i^* \in [N]$*:* $g_{ii^*}(t, \cdot) \ne 0$.

Assumption 5.2.7. $\forall i, j \in [N]$, $\tau_{ij}(t) \in C^1([t_0, \infty), [0, \tau])$ *for some* $\tau < \infty$.

Pick $m > (N-1)\bar{c}$, $B > 0$ and set

$$\underline{\kappa} := \min\left\{\frac{\underline{c}(1-e^{-mB})}{m}, e^{-mB}\right\}, \qquad \theta := \frac{\ln(1-e^{-\underline{\kappa}\bar{N}\bar{c}\tau})}{2\tau+B}.$$

Theorem 5.2.8. *Under Assumptions 5.2.4, 5.2.6 and 5.2.7, the solution* $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi})$, $t \geq t_0$ *of* (5.5) *satisfies*

$$\max_{i\in[N]} |x_i(t) - k| \leq \frac{S_{I_{t_0}}(\boldsymbol{\phi})}{1-\underline{\kappa}e^{-\bar{N}\bar{c}\tau}} e^{-\theta(t-t_0)} \tag{5.11}$$

for some $k \in W_{I_{t_0},\boldsymbol{\phi}}$.

Proof. Consider the solution $\mathbf{x}(t, t_0, \boldsymbol{\phi}) = (x_1, \ldots, x_N)$ of (5.5). From Assumption 5.2.4 we set $M = W_{I_{t_0},\boldsymbol{\phi}}$ so that $L_{ij}(t) = L_{ij}^M(t), U_{ij}(t) = U_{ij}^M(t) \geq 0$ are well-defined. Observe that

$$\sum_j \tilde{a}_{ij}(t)\big(x_j(\lambda_{ij}(t)) - x_i(t)\big) \leq \dot{x}_i(t) \leq \sum_j \tilde{b}_{ij}(t)\big(x_j(\lambda_{ij}(t)) - x_i(t)\big) \tag{5.12}$$

where

$$\tilde{a}_{ij}(t) := \begin{cases} U_{ij}(t), & x_j(\lambda_{ij}(t)) > x_i(t) \\ L_{ij}(t), & x_j(\lambda_{ij}(t)) < x_i(t) \end{cases}$$

and

$$\tilde{b}_{ij}(t) := \begin{cases} L_{ij}(t), & x_j(\lambda_{ij}(t)) > x_i(t) \\ U_{ij}(t), & x_j(\lambda_{ij}(t)) < x_i(t) \end{cases}$$

Observe that from (5.12), $x_i(t)$ satisfies

$$\sum_{j=1}^N \int_{t-B}^t a_{ij}(t,s) x_j\big(\lambda_{ij}(s)\big)\, ds \leq x_i(t) \leq \sum_{j=1}^N \int_{t-B}^t b_{ij}(t,s) x_j\big(\lambda_{ij}(s)\big)\, ds \tag{5.13}$$

where

$$a_{ij}(t,s) := \begin{cases} e^{-m(t-s)}\tilde{a}_{ij}(s), & i \neq j \\ e^{-mB}\delta(s-(t-B)) + e^{-m(t-s)}\big(m - \sum_j \tilde{a}_{ij}(s)\big), & i = j \end{cases}$$

$$b_{ij}(t,s) := \begin{cases} e^{-m(t-s)}\tilde{b}_{ij}(s), & i \neq j \\ \\ e^{-mB}\delta(s-(t-B)) + e^{-m(t-s)}\big(m - \sum_j \tilde{b}_{ij}(s)\big), & i = j \end{cases}$$

and $\delta(\cdot)$ is the delta function. In vector form (5.13) reads

$$\int_{t-B}^{t} A(t,s)\mathbf{x}_{(s)}\,ds \leq \mathbf{x}(t) \leq \int_{t-B}^{t} B(t,s)\mathbf{x}_{(s)}\,ds \tag{5.14}$$

where $A(t,s) := [a_{ij}(t,s)]$, $B(t,s) := [b_{ij}(t,s)]$ so that the inequalities hold element-wise and $\mathbf{x}_{(s)}$ is the vector assembled by the delayed states of $\mathbf{x}$ as they occur from (5.13). Next, it can be easily shown that $\int_{t-B}^{t} A(t,s)\,ds$ and $\int_{t-B}^{t} B(t,s)\,ds$ are stochastic matrices with the properties that $\forall t_1, t_2 \in [t-B, t]$

1. $\sum_j \int_{t_1}^{t_2} a_{ij}(t,s)\,ds = \sum_j \int_{t_1}^{t_2} b_{ij}(t,s)\,ds \equiv const.$

2. $\sum_j \int_{t-B}^{t} a_{ij}(t,s)\,ds = \sum_j \int_{t-B}^{t} b_{ij}(t,s)\,ds \equiv 1$

For $h, h' \in [N]$ we have from (5.13)

$$x_h(t) - x_{h'}(t) \leq \sum_j \int_{t-B}^{t} \big(b_{hj}(t,s) - a_{h'j}(t,s)\big)x_j(s)\,ds \tag{5.15}$$

Now, for fixed $t > B$, we consider the partition $t_0 < t_1 < t_2 < \cdots < t_l$ where $t_0 = t - B$ and $t_l = t$ such that for any $[t_{k-1}, t_k]$, $b_{hj}(t,s) - a_{h'j}(t,s)$, $s \in [t_{k-1}, t_k]$ does not change sign. Within this interval we apply Theorem 2.3.5 to obtain:

$$\int_{t_{k-1}}^{t_k} \big(b_{hj}(t,s) - a_{h'j}(t,s)\big)x_j(s)\,ds = \int_{t_{k-1}}^{t_k} \big(b_{hj}(t,s) - a_{h'j}(t,s)\big)\,ds x_j(s_j^*)$$

for some $s_j^* \in [t_{k-1}, t_k]$. Now we take:

a. $j' : u_{j'}^k := \int_{t_{k-1}}^{t_k} \big(b_{hj'}(t,s) - a_{h'j'}(t,s)\big)\,ds > 0$

b. $j'' : u_{j''}^k := \int_{t_{k-1}}^{t_k} \big(b_{hj''}(t,s) - a_{h'j''}(t,s)\big)\,ds < 0$

From Property 1, we have that $\sum_j u_j^k \equiv 0$ for any $k \geq 1$ and we set

$$\theta_k = \sum_{j'} u_{j'}^k = \sum_{j'} |u_{j'}^k| = -\sum_{j''} u_{j''}^k = \sum_{j''} |u_{j''}^k| = \frac{1}{2}\sum_j |u_j^k|$$

so that $\theta_k > 0$. Then

$$\begin{aligned}
&x_h(t) - x_{h'}(t) \leq \\
&\leq \sum_{k\geq 1}\sum_{j=1}^{N}\int_{t_{k-1}}^{t_k}\big(b_{hj}(t,s) - a_{h'j}(t,s)\big)x_j(s)\,ds = \sum_{k\geq 1}\sum_{j=1}^{N} u_j(k)x_j(s_j^*(k)) \\
&\leq \sum_{k\geq 1}\theta_k\left(\frac{\sum_{j'}|u_{j'}|x_{j'}(s_{j'}(k))}{\theta_k} - \frac{\sum_{j''}|u_{j''}|x_{j'}(s_{j'}(k))}{\theta_k}\right) \\
&\leq \left[\sum_{k\geq 1}\theta_k\right]\big(\max_{k,i} x_i(s_i(k)) - \min_{k,i} x_i(s_i(k))\big) \\
&\leq \frac{1}{2}\max_{h,h'}\sum_j\int_{t-B}^{t}|b_{h'j}(t,s) - a_{h'j}(t,s)|\,ds\big(\max_{k,i} x_i(s_i(k)) - \min_{k,i} x_i(s_i(k))\big)
\end{aligned}$$

In view of Lemma 5.2.5, the identity $|x-y| = x+y-2\min\{x,y\}$ and the fact that $\sum_j\int_{t-B}^{t} b_{ij}(t,s)\,ds = \sum_j\int_{t-B}^{t} a_{ij}(t,s)\,ds \equiv 1$, $i \in [N]$, (Property 2) we have that the above inequality holds for arbitrary $h, h' \in [N]$. Consequently, we obtain the estimate:

$$S\big(\mathbf{x}(t)\big) \leq \big(1-\kappa(t)\big)S_{I_{t-B-2\tau}}(\mathbf{x}) \tag{5.16}$$

For $\underline{\kappa}(t) := \min_{h,h'}\sum_j \min\big\{\int_{t-B}^{t} b_{hj}(t,s)\,ds, \int_{t-B}^{t} a_{h'j}(t,s)\,ds\big\}$. As the matrices $\int_{t-B}^{t} A(t,s)\,ds, \int_{t-B}^{t} B(t,s)\,ds$ are stochastic, it can be easily shown that $\rho \leq 1$. Also, by the uniform condition $L_{ij}(t) \neq 0 \Rightarrow L_{ij}(t) \geq c$ as imposed by Assumptions 5.2.4 and 5.2.6 we obtain the positive uniform lower bound:

$$\sup_{t\geq t_0+B}\underline{\kappa}(t) \geq \underline{\kappa} = \min\left\{\frac{c(1-e^{-mB})}{m}, e^{-mB}\right\} > 0$$

observe that $c < m$ and hence $\rho \in (0,1)$. Next, for any $t \geq t_0 + B + \tau$ consider the set $W_{I_t,\mathbf{x}_t}$ together with the spread S_{I_t} let $i \in [N]$ be the agent the state that achieves the maximum over all states throughout $I_t = [t - \tau(t), t]$ and $j \in [N]$ be the agent with respective minimum value. Denote these states by with $x_i(t')$ and $x_j(t'')$, with $t', t'' \in I_t$ respectively. Let $t' > t''$. Then $S_{I_t}(\mathbf{x}) = x_i(t') - x_j(t'')$ so that

$$
\begin{aligned}
& x_i(t') - x_j(t'') \leq \\
& \leq e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw} x_i(t'') - x_j(t'') + \int_{t''}^{t'} e^{-\int_s^{t'} \tilde{b}_{ij}(w)\,dw} \sum_j b_{ij}(s) x_j(\lambda_{ij}(s))\,ds \\
& \leq e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw} \big(x_i(t'') - x_j(t'')\big) + \\
& \qquad\qquad + \int_{t''}^{t'} e^{-\int_s^{t'} \tilde{b}_{ij}(w)\,dw} \sum_j b_{ij}(s) \big(x_j(\lambda_{ij}(s)) - x_i(t'')\big)\,ds \\
& \leq e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw} S\big(\mathbf{x}(t'')\big) + \big(1 - e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw}\big) S_{I_{t-B-2\tau}}(\mathbf{x}) \\
& \leq e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw} (1-\rho) S_{I_{t-B-2\tau}}(\mathbf{x}) + \big(1 - e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw}\big) S_{I_{t-B-2\tau}}(\mathbf{x}) \\
& \leq \big(1 - \underline{\kappa} e^{-\int_{t''}^{t'} \sum_j \tilde{b}_{ij}(w)\,dw}\big) S_{I_{t-B-2\tau}}(\mathbf{x}) \\
& \leq \big(1 - \underline{\kappa} e^{-\bar{N} c \tau}\big) S_{I_{t-B-2\tau}}(\mathbf{x})
\end{aligned}
$$

in view of (5.14) and (5.16). The same bound is achieved if $t' \geq t''$ where one must begin by taking the left part of the double inequality of (5.14). Consequently, following the same recursive argument as in Theorem 4.4.6 we obtain the desired estimate (5.11). Finally, we follow the same argumentation as in Lemma 4.4.5 so that the forward limit set must consist of a singleton which by Lemma 5.2.5 is in $W_{I_{t_0},\boldsymbol{\phi}}$ concluding the proof. $\square$

A first remark on the proof is that, contrary to the method developed in the previous sections, the base of which is Theorem 2.3.1, here we followed the steps of the original argument of Markov [22] on the effect of the averaging property of stochastic matrices. That proof can be found in [26]. reverse statement would yield a conditional consensus result that is a direct consequence of Theorem 5.2.8 and it is stated without proof.

Corollary 5.2.9. *Let Assumption 5.2.6 be true. If one can find $M \subset \mathbb{R}$ such that Assumption 5.2.4 is true as well, then for $\phi_i \in C^0([-\tau, 0], M)$, $\forall i \in [N]$, the solution of* (5.5) *exhibits exponential asymptotic consensus with estimates as in Theorem 5.2.8.*

On the consensus point   If $g_{ij}(t, x) \equiv g_{ij}(x)$ and $\tau_{ij}(t) \equiv \tau_{ij}$ then (5.5) turns to an autonomous system and we can characterize the limit point k under a supplementary symmetry assumption:

Assumption 5.2.10. $\forall i \neq j \in [N]$, $g_{ij} \in C^1(\mathbb{R}, \mathbb{R})$ *with the property* $g'_{ij} = g'_{ji}$.

For the main result we will need the next technical lemma:

Lemma 5.2.11. *Set* $I = [-\tau, 0]$ *and let Assumption 5.2.10 hold. Given* $\phi_i \in C^0(I, \mathbb{R})$, $i = 1, \dots, N$ *then* $\exists\, !\ k \in W_{I,\boldsymbol{\phi}}$ *to satisfy*

$$k = \sum_i \alpha_i \phi_i(0) + \sum_{i,j} \beta_i \bigg(\int_{-\tau_{ij}}^{0} g_{ij}(\phi_j(s))\, ds - \tau_{ij} g_{ij}(k) \bigg)$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^N$ *are non-negative vectors with* $\sum_i \alpha_i = \sum_i \beta_i = 1$.

Proof. Define the function $J : W_{I,\boldsymbol{\phi}} \to \mathbb{R}$

$$J(k) := k - \sum_i \alpha_i \phi_i(0) - \sum_i \beta_i \sum_{j \neq i} \int_{-\tau}^{0} g_{ij}(\phi_j(s))\, ds + \sum_i \beta_i \sum_{j \neq i} g_{ij}(k)$$

We begin by excluding the trivial cases. This is for $W_{I,\boldsymbol{\phi}}$ being a singleton, i.e. $W_{I,\boldsymbol{\phi}} = \{k\}$ and hence $\phi_i \equiv k$ and automatically $J \equiv 0$. If $W_{I,\boldsymbol{\phi}}$ is not a singleton, then we take

$$k_1 := \min_{i \in V} \min_{s \in [-\tau, 0]} \phi_i(s) < \max_{i \in V} \max_{s \in [-\tau, 0]} \phi_i(s) =: k_2$$

by continuity of ϕ_i and g_{ij} we conclude that

$$k_1 \leq \phi_i(0) \;\;\&\;\; g_{ij}(k) \leq g_{ij}(\phi_i(s))$$

but with some i, j such that $g_{ij}(k) < g_{ij}(\phi(s))$ for some $s \in I \subset [-\tau, 0]$. Consequently, $J(k_1) < 0$ and similar analysis will yield $J(k_2) > 0$. Then by the elementary theorem of Bolzano there exists $k \in W$ such that $J(k) = 0$. The uniqueness of k follows from the monotonicity of J. Indeed,

$$J' = 1 + \sum_{i,j} \beta_i g'_{ij}(k) > 0.$$

□

Theorem 5.2.12. *Consider the system* (5.5) *with* $g_{ij}(t, x) = g_{ij}(x)$ *and* $\tau_{ij}(t) \equiv \tau_{ij}$*. Let Assumptions 5.2.4 5.2.6 and 5.2.10 hold. The solution* $\mathbf{x} = \mathbf{x}(t, t_0, \boldsymbol{\phi})$*,* $t \geq t_0$ *of* (5.5) *converges to the unique solution of*

$$k = \sum_j \alpha_j \phi_j(0) + \sum_{i,j} \frac{1}{N} \int_{-\tau_{ij}}^{0} g_{ij}\big(\phi_j(s)\big)\, ds - \sum_{i,j} \frac{1}{N} \tau_{ij} g_{ij}(k)$$

exponentially fast with rates dictated by Theorem 5.2.8.

Proof. It suffices to show that $\mathbf{x}$ indeed converges to a point satisfying the aforementioned nonlinear algebraic equation. We rewrite (5.5) as follows:

$$\dot{x}_i(t) = \sum_{j \in N_i} g_{ij}(x_j) - g_{ij}(x_i) - \frac{d}{dt} \int_{t-\tau_{ij}}^{t} \big[g_{ij}\big(x_j(s)\big) - g_{ij}(k)\big]\, ds$$

Consider the solution $\mathbf{y}(t,0,\boldsymbol{\phi}^0)$ of

$$i\in[N]:\begin{cases}\dot{y}_i=\sum_{j\in N_i}g_{ij}(y_j)-g_{ij}(y_i), t\geq 0\\ \\ y_i(0)=\phi_i^0\end{cases}\tag{5.17}$$

Set

$$\mathbf{G}(\mathbf{y}):=\begin{bmatrix}\sum_j g_{1j}(y_j)-g_{1j}(y_1)\\ \sum_j g_{2j}(y_j)-g_{2j}(y_2)\\ \vdots\\ \sum_j g_{Nj}(y_j)-g_{Nj}(y_N)\end{bmatrix}$$

so that in vector form (5.17) reads,

$$\dot{\mathbf{y}}=\mathbf{G}(\mathbf{y}),\qquad \mathbf{y}(0)=\boldsymbol{\phi}^0$$

For $t\geq s\geq 0$, $\mathbf{V}(s):=\mathbf{y}\big(t,s,\mathbf{x}(s)\big)$ and differentiate with respect to s

$$\begin{aligned}\frac{d}{ds}\mathbf{V}(s)&=\frac{\partial\mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial s}+\frac{\partial\mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial\boldsymbol{\xi}}\dot{\mathbf{x}}(s)\\ &=\frac{\partial\mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial\boldsymbol{\xi}}\Big(\dot{\mathbf{x}}(s)-\mathbf{G}(\mathbf{x}(s))\Big)\\ &=-\frac{\partial\mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial\boldsymbol{\xi}}\frac{d}{ds}\begin{bmatrix}\sum_j\int_{t-\tau_{1j}}^{t}\big[g_{1j}(x_j(s))-g_{1j}(k)\big]\,ds\\ \sum_j\int_{t-\tau_{2j}}^{t}\big[g_{2j}(x_j(s))-g_{2j}(k)\big]\,ds\\ \vdots\\ \sum_j\int_{t-\tau_{Nj}}^{t}\big[g_{Nj}(x_j(s))-g_{Nj}(k)\big]\,ds\end{bmatrix}\\ &=-\frac{\partial\mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial\boldsymbol{\xi}}\frac{d}{ds}\mathbf{H}(\mathbf{x}_s)\end{aligned}$$

Note that $\frac{\partial\mathbf{y}(t,s,\mathbf{x}(s))}{\partial\boldsymbol{\xi}}$ is the principal matrix solution of the following linear non-autonomous system

$$\dot{\mathbf{z}}=\mathbf{G}'\big(\mathbf{y}(t,s,\mathbf{x}(s))\big)\mathbf{z}.$$

This is a consensus network with symmetric weights. From Proposition 2.2.1 and Remark 2.2.1.1 we deduce that *regardless* $\mathbf{y}(t,s,\mathbf{x}(s))$, $\mathbf{z}(t) \to \frac{\mathbb{1}\mathbb{1}^T}{N}\mathbf{z}^0$ exponentially fast. Consequently, $\frac{\partial \mathbf{y}(t,s,\mathbf{x}(s))}{\partial \boldsymbol{\xi}}$ satisfies

$$\left| \frac{\partial \mathbf{y}(t,s,\mathbf{x}(s))}{\partial \boldsymbol{\xi}} - \frac{\mathbb{1}\mathbb{1}^T}{N}\mathbf{z}^0 \right| \leq Re^{-r(t-s)} \tag{5.18}$$

for some $R, r > 0$ that depend on the connectivity weights g'_{ij} and the norm. Next, we integrate from 0 to t to obtain the following expression for the solution of (5.5)

$$\mathbf{x}(t,0,\boldsymbol{\phi}) = \mathbf{y}(t,0,\boldsymbol{\phi}^0) - \int_0^t \frac{\partial \mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial \boldsymbol{\xi}} \frac{d}{ds}\mathbf{H}(\mathbf{x}_s)\, ds$$

Next, integration by parts and change of the order of integration yields

$$\begin{aligned}
\mathbf{x}(t) &= \mathbf{y}(t,0,\boldsymbol{\phi}^0) - \mathbf{H}(\mathbf{x}_t) + \frac{\partial \mathbf{y}\big(t,0,\mathbf{x}^0\big)}{\partial \boldsymbol{\xi}}\mathbf{H}(\boldsymbol{\phi}) + \int_0^t \frac{d}{ds}\left[\frac{\partial \mathbf{y}\big(t,s,\mathbf{x}(s)\big)}{\partial \boldsymbol{\xi}}\right]\mathbf{H}(\mathbf{x}_s)\, ds \\
&= \mathbf{y}(t,0,\boldsymbol{\phi}^0) - \mathbf{H}(\mathbf{x}_t) + \frac{\partial \mathbf{y}\big(t,0,\mathbf{x}^0\big)}{\partial \boldsymbol{\xi}}\mathbf{H}(\boldsymbol{\phi}) + \\
&+ \sum_{i,j}\Bigg\{ \int_{-\tau_{ij}}^0 \left[\left(\frac{\partial \mathbf{y}\big(t,w+\tau_{ij},\boldsymbol{\phi}(w+\tau_{ij})\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{y}\big(t,0,\boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}}\right)\mathbf{H}(\boldsymbol{\phi}(w))\right]_{ij} dw + \\
&+ \int_0^{t-\tau_{ij}} \left[\left(\frac{\partial \mathbf{y}\big(t,w+\tau_{ij},\mathbf{x}(w+\tau_{ij})\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{y}\big(t,w,\mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right)\mathbf{H}(\mathbf{x}(w))\right]_{ij} dw \\
&+ \int_{t-\tau_{ij}}^t \left[\left(I_{N\times N} - \frac{\partial \mathbf{y}\big(t,w,\mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right)\mathbf{H}(\mathbf{x}(w))\right]_{ij} dw \Bigg\}
\end{aligned}$$

As $t \to \infty$

$$\mathbf{y}(t,0,\boldsymbol{\phi}^0) \to \mathbb{1}\sum_j \alpha_j \phi_j(0)$$

in view of Theorem 5.2.8, for some $\alpha_j \geq 0$ such that $\sum_j \alpha_j = 1$. Also since $\mathbf{x}(t) \to \mathbb{1}k$ again in view of Theorem 5.2.8

$$\mathbf{H}(\mathbf{x}_t) \to 0 \qquad \& \qquad \int_{t-\tau_{ij}}^t \left(I_{N\times N} - \frac{\partial \mathbf{y}\big(t,w,\mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right)\mathbf{H}(\mathbf{x}(w))\, dw \to 0$$

Next, in view of (5.18)

$$\frac{\partial \mathbf{y}\big(t,w+\tau_{ij},\boldsymbol{\phi}(w+\tau_{ij})\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{y}\big(t,0,\boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}} \to 0$$

and

$$
\begin{aligned}
&\int_0^{t-\tau_{ij}} \left(\frac{\partial \mathbf{y}\big(t, w+\tau_{ij}, \mathbf{x}(w+\tau_{ij})\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{y}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}} \right) \mathbf{H}(\mathbf{x}(w))\, dw = \\
&\int_0^{t-\tau_{ij}} \left(\frac{\partial \mathbf{y}\big(t, w+\tau_{ij}, \mathbf{x}(w+\tau_{ij})\big)}{\partial \boldsymbol{\xi}} - \frac{\mathbb{1}\mathbb{1}^T}{N} \right) \mathbf{H}(\mathbf{x}(w))\, dw - \\
&\int_0^{t-\tau_{ij}} \left(\frac{\partial \mathbf{y}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}} - \frac{\mathbb{1}\mathbb{1}^T}{N} \right) \mathbf{H}(\mathbf{x}(w))\, dw.
\end{aligned}
$$

From (5.18) we deduce that both of these integrals asymptotically vanish because they are convolutions of L^1 functions, i.e. $\frac{\partial \mathbf{y}(t,w,\mathbf{x}(w))}{\partial \boldsymbol{\xi}} - \frac{\mathbb{1}\mathbb{1}^T}{N}$, with a function that goes to zero, i.e. $\mathbf{H}(\mathbf{x}(w))$. Thus, in the limit $t = \infty$ we are left with

$$
\mathbb{1}k = \mathbb{1} \sum_j \alpha_j \phi_j(0) + \mathbb{1} \sum_{i,j} \frac{1}{N} \int_{-\tau_{ij}}^0 g_{ij}\big(\phi_j(s)\big)\, ds - \mathbb{1} \sum_{i,j} \frac{1}{N} \int_{-\tau_{ij}}^0 g_{ij}(k)\, ds
$$

which, by Proposition 5.2.11, we know that it attains a unique solution in $W_{I_{t_0}, \boldsymbol{\phi}}$. $\square$

5.2.3 Periodic synchronized solutions

The effect of delays in co-operative systems may cause more complex behavior than simply weakening the rate of convergence to consensus in the networks we considered so far. Indeed, whenever the dynamics are nonlinear and the type of delays is distributed, there seems to be a possibility of a periodic solution. We are mainly interested in the following type of periodic solutions.

Definition 5.2.13. A function $\mathbf{y}(t) \in \mathbb{R}^N$ is *synchronized* if it is bounded and it satisfies $S(\mathbf{y}(t)) \equiv 0$.

This is an extended concept of agreement that basically accepts consensus solutions along a non-trivial orbit. Fix $t_0 \in \mathbb{R}$ and $\tau > 0$ and consider the initial value problem

$$i \in [N] : \begin{cases} \dot{x}_i(t) = -\sum_j g_{ij}\big(t, t, x_i(t)\big) + \\ \qquad + \sum_j \int_{-\tau}^{0} g_{ij}\big(t, s, x_j(t+s)\big) p(t,s)\, ds, & t \geq t_0 \\ x_i(t) = \phi_i(t), & t \in [t_0 - \tau, t_0] \end{cases} \tag{5.19}$$

where $p \in C^0([t_0, \infty) \times [-\tau, 0], \mathbb{R}_+)$ has the property

$$\int_{-\tau}^{0} p(t,s)\, ds = 1, \qquad t \geq 0 \tag{5.20}$$

This network is a significant variation of the ones studied so far. The structural hypothesis so far imposes the following condition: Agent i receives the signal with the state x_j from agent j with a coupling weight which suffers from no processing delay. This condition would makes sense only if the particular rate is a parameter controlled exclusively by i. Otherwise, if the information on the coupling rate is also transmitted from j should suffer from delays. We will show here that if this is the case, periodic solutions can occur. We study the generic scenario where uncertainty is put in an interval of possible delays. This is typically expressed via a distribution function that smoothly weights the different possible delays.

We open the discussion with an unorthodox result. The following theorem comments on the non-existence of non-trivial periodic synchronized solutions.

Theorem 5.2.14. *Consider the solution* $\mathbf{x} = \mathbf{x}(t, 0, \boldsymbol{\phi})$, $t \geq 0$ *of* (5.19). *If* $\mathbf{x}$ *is synchronized and periodic with period* T *and either of the following two conditions holds*

A: $\forall i \neq j$, $g_{ij}(t, s, x) = g_{ij}(s, x)$ *is continuous in* s *and* x *with the property that*

$g_{ij}(s+T, x) = g_{ij}(s, x)$, $\tau = T$ and $p(t, s) = \delta(s+\tau)$, $\forall t$ for $\delta(\cdot)$ to be the delta function,

B: $\forall i \neq j$, $g_{ij}(t, s, x) = a_{ij}(t)x$ with the property that $a_{ij}(t+T) = a_{ij}(t)$ for any t and $p(t, s) = \delta(s+\tau(t))$, $\forall t$ for $\delta(\cdot)$ to be the delta function and $\tau(t)$ is T-periodic with $\tau(t) \leq \bar{\tau}$,

then $\mathbf{x}$ is constant.

Proof. Let the initial data $\boldsymbol{\phi} \in \Delta$ such that for any $t \geq 0$ the solution $\mathbf{x}(t)$ satisfies $x_i(t) = x_j(t)$ for any $i \neq j$ and $\mathbf{x}(t) = \mathbf{x}(t+T)$. Along this solution pick an arbitrary $i \in [N]$ and observe that the solution satisfies

$$\dot{x}_i(t) = \sum_j \int_{-\tau}^{0} g_{ij}\big(t, s, x_i(t+s)\big)p(t, s)\, ds - g_{ij}\big(t, t, x_i(t)\big) \tag{5.21}$$

Condition A: (5.21) reads

$$\dot{x}_i(t) = \sum_j g_{ij}\big(t-T, x_i(t-T)\big) - g_{ij}\big(t, x_i(t)\big) = -\frac{d}{dt}\sum_j \int_{t-T}^{t} g_{ij}\big(s, x_i(s)\big)\, ds$$

then

$$x_i(t) = \phi_i(0) + \int_{-T}^{0} \sum_j g_{ij}\big(s, \phi_i(s)\big)\, ds - \int_{t-T}^{t} \sum_j g_{ij}\big(s, x_i(s)\big)\, ds$$

if $x_i(t) = x_i(t+T)$ then this implies that the integral of $\sum_j g_{ij}\big(s, x_i(s)\big)$ over any $T-$ interval is constant, hence $x_i(t) \equiv \phi_i(0)$ and it is constant.

Condition B:

If $\mathbf{x}$ is not constant, the $\mathbf{y} := \dot{\mathbf{x}}$ is not constant. In such case (5.21) reads

$$y_i(t) = \sum_j a_{ij}(t) \int_{t-\tau(t)}^{t} y_i(s)\, ds \tag{5.22}$$

so that $y_i(t+T) = y_i(t)$ implies that the integral of y_i over $[t, t+T]$ is constant. Then $y_i(t)$ must have the form $k\sum_j a_{ij}(t)$. Substituting this form to (5.22), we obtain

$$k\sum_j a_{ij}(t) = \sum_j a_{ij}(t)k\int_{t-\tau(t)}^{t}\sum_l a_{il}(s)\,ds. \tag{5.23}$$

From the above necessary condition we have the cases:

a. $k = 0$ or $\sum_j a_{ij}(t) \equiv 0$ so that $\mathbf{y} \equiv 0$ and $x_i(t) \equiv \phi_i(0)$, i.e. $\mathbf{x}$ must be a constant,

b. $\int_{t-\tau(t)}^{t}\sum_j a_{ij}(s)\,ds \equiv 1$ and this leads to $x_i(t) = \phi_i(0) + k\int_0^t \sum_j a_{ij}(s)\,ds$ being unbounded, i.e. a contradiction. □

The above result suggests that constant solutions are more or less the canon for the majority the delayed consensus systems. As the objective is to introduce a case of non-trivial periodic synchronized solution as a result of the delay in consensus systems, exhaustive simulations have suggested the following modification of the consensus networks considered so far.

Delay induced synchronization For $T > 0$ we consider the following initial value problem

$$i \in [N] : \begin{cases} \dot{x}_i(t) = -\sum_j g_{ij}\big(t, x_i(t)\big) + \\ \qquad + \sum_j \int_{t-T}^{t} g_{ij}\big(s, x_j(s)\big)p(s-t)\,ds, & t \geq 0 \\ x_i(t) = \phi_i(t), & t \in [-T, 0] \end{cases} \tag{5.24}$$

where p is a distributed delay satisfying (5.20) with $\tau = T$. The conditions we are imposing on g_{ij} are significantly harder than the ones considered so far.

Assumption 5.2.15. *$\forall i, j \in [N],\ t \geq 0,\ x \in \mathbb{R}$, the following properties hold:*

(i.) $g_{ij}(t,x) > 0$*, uniformly in* t*, if and only if* $j \in N_i$ *and zero otherwise.*

(ii.) $g_{ij}(t,x) = g_{ji}(t,x)$ *&* $g_{ij}(t+T,x) = g_{ij}(t,x)$*,*

(iii.) $\frac{\partial}{\partial x} g_{ij}(t,x) \in C^0\big([0,\infty) \times \mathbb{R}, [\underline{K}, \bar{K}]\big)$ *for* $j \in N_i$ *and some* $0 < \underline{K} \leq \bar{K} < \infty$ *that are independent of* t

(iv.) $\frac{\partial}{\partial x} g_{ij}(t,x) = \frac{\partial}{\partial x} g_{ji}(t,x)$*.*

(v.) $\sum_j g_{ij}(t,x)$ *is independent of* i*.*

Assumption 5.2.16. *The connectivity graph is static and with increased connectivity: i.e. there exists* $j \in [N] : g_{ij} \neq 0$ *for all* $i \in [N] \backslash \{j\}$*.*

Recalling the discussion in Chapter 3 this is a Type I static connectivity.

Proposition 5.2.17. *Let Assumption 5.2.15 hold. If*

$$\bar{K} \int_{-T}^{0} p(s)(-s)\, ds < 1$$

there exists a unique synchronized periodic solution of (5.24) *with period* T*. The solution is constant only if there is* k *such that*

$$\sum_j g_{ij}(t,k) = \int_{-T}^{0} p(s) \sum_j g_{ij}(t+s,k)\, ds$$

Proof. We begin with the second statement. If $\mathbf{x}$ is synchronized and $T-$periodic then $\mathbf{x}(t) = \big(x_1(t), \ldots, x_N(t)\big) = \big(\zeta(t), \ldots, \zeta(t)\big)$ for some appropriate function. By Assumption 5.2.15($v.$),

$$\dot{\zeta}(t) = -\sum_j g_{ij}\big(t, \zeta(t)\big) + \sum_j \int_{t-T}^{t} g_{ij}\big(s, \zeta(s)\big) p(s-t)\, ds$$

independent of i. If $\zeta(t) \equiv k$ for some $k \in \mathbb{R}$, then the last equation reads

$$\begin{aligned} 0 &= -\sum_j g_{ij}\big(t,k\big) + \sum_j \int_{t-T}^{t} g_{ij}\big(s,k\big)p(s-t)\,ds \\ &= -\sum_j g_{ij}\big(t,k\big) + \int_{-T}^{0} p(s) \sum_j g_{ij}(t+s,k)\,ds \end{aligned}$$

We proceed with proving the existence and uniqueness of a periodic solution $\mathbf{x}(t) = \mathbb{1}\zeta(t)$. Adding and subtracting $g_{ij}\big(t, x_j(t)\big)$ we arrive at the following equivalent functional differential equation for x_i

$$\dot{x}_i(t) = \sum_j g_{ij}\big(t, x_j(t)\big) - g_{ij}\big(t, x_i(t)\big) - \frac{d}{dt}\int_{-T}^{0} p(s) \int_{t+s}^{t} g_{ij}\big(w, x_j(w)\big)\,dwds$$

Next we follow the steps of the proof of Theorem 5.2.12. We set $l(s) = \mathbf{x}\big(t, s, \mathbf{z}(s)\big)$, differentiate with respect to s and integrate from 0 to t, in the sense of we express the solution $\mathbf{x}(t) = \mathbf{x}(t, 0, \boldsymbol{\phi})$ of (5.24) as follows:

$$\begin{aligned}
&\mathbf{x}(t) = \\
&= \mathbf{z}(t, 0, \boldsymbol{\phi}^0) - \int_0^t \frac{\partial \mathbf{z}\big(t, s, \mathbf{x}(s)\big)}{\partial \boldsymbol{\xi}} \frac{d}{ds} \int_{-T}^0 p(q) \int_{s+q}^s \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdqds \\
&= \mathbf{z}(t, 0, \boldsymbol{\phi}^0) - \int_{-T}^0 p(q) \int_{t+q}^t \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq \\
&+ \frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}} \int_{-T}^0 p(q) \int_q^0 \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dwdq \\
&+ \int_0^t \frac{d}{ds}\left[\frac{\partial \mathbf{z}\big(t, s, \mathbf{x}(s)\big)}{\partial \boldsymbol{\xi}}\right] \int_{-T}^0 p(q) \int_{s+q}^s \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdqds \\
&= \mathbf{z}(t, 0, \boldsymbol{\phi}^0) - \int_{-T}^0 p(q) \int_{t+q}^t \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq \\
&+ \frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}} \int_{-T}^0 p(q) \int_q^0 \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dwdq \\
&+ \int_{-T}^0 p(q) \int_q^0 \left[\frac{\partial \mathbf{z}\big(t, w-q, \mathbf{x}(w-q)\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}}\right] \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dwdq \\
&+ \int_{-T}^0 p(q) \int_0^{t+q} \left[\frac{\partial \mathbf{z}\big(t, w-q, \mathbf{x}(w-q)\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{z}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right] \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq \\
&+ \int_{-T}^0 p(q) \int_{t+q}^t \left[I_{N\times N} - \frac{\partial \mathbf{z}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right] \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq \\
&= \mathbf{z}(t, 0, \boldsymbol{\phi}^0) + \frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}} \int_{-T}^0 p(q) \int_q^0 \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dwdq \\
&+ \int_{-T}^0 p(q) \int_q^0 \left[\frac{\partial \mathbf{z}\big(t, w-q, \mathbf{x}(w-q)\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}}\right] \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dwdq \\
&+ \int_{-T}^0 p(q) \int_0^{t+q} \left[\frac{\partial \mathbf{z}\big(t, w-q, \mathbf{x}(w-q)\big)}{\partial \boldsymbol{\xi}} - \frac{\partial \mathbf{z}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}}\right] \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq \\
&- \int_{-T}^0 p(q) \int_{t+q}^t \frac{\partial \mathbf{z}\big(t, w, \mathbf{x}(w)\big)}{\partial \boldsymbol{\xi}} \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dwdq
\end{aligned} \tag{5.25}$$

Let $\big(\mathbb{S}, |\cdot|\big)$ be the Banach space of continuous T-periodic synchronized functions. Define the operator $\mathcal{Q} : \mathbb{S} \to \mathbb{B}$ as follows:

$$(\mathcal{Q}\mathbf{x})(t) = \mathbf{x}_{(5.25)}(t)$$

where $\mathbf{x}_{(5.25)}(t)$ is the right hand-side of (5.25) as it was expressed above. Interestingly enough, $\mathcal{Q}$, restricted to $\mathbb{S}$, becomes particularly simplified. Indeed, for $\mathbf{x} \in \mathbb{S}$ under Assumption 5.2.15:

1. $\mathbf{z}(t, 0, \boldsymbol{\phi}^0) \equiv \boldsymbol{\phi}^0$,
2. $\mathbf{G}\big(w, \mathbf{x}(w)\big) \in \Delta$ for any fixed $w \geq -T$.

Consequently, the principal matrix $\frac{\partial \mathbf{z}}{\partial \boldsymbol{\xi}}$ acts on Δ and thus has no effect on any such element of Δ. So for instance,

$$\frac{\partial \mathbf{z}\big(t, 0, \boldsymbol{\phi}^0\big)}{\partial \boldsymbol{\xi}} \int_{-T}^{0} p(q) \int_{q}^{0} \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dw dq \equiv \int_{-T}^{0} p(q) \int_{q}^{0} \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dw dq$$

Additionally, the last three integrals are identically equal to zero. What is left then, is

$$(\mathcal{Q}\mathbf{x})(t) = \boldsymbol{\phi}^0 + \int_{-T}^{0} p(q) \int_{q}^{0} \mathbf{G}\big(w, \boldsymbol{\phi}(w)\big)\, dw dq - \int_{-T}^{0} p(q) \int_{t+q}^{t} \mathbf{G}\big(w, \mathbf{x}(w)\big)\, dw dq$$

and it is easy to see that $(\mathcal{Q}\mathbf{x})(t+T) = (\mathcal{Q}\mathbf{x})(t)$.

Finally, under the stated condition $\mathcal{Q}$ becomes a contraction under the metric $\rho(\mathbf{x}, \mathbf{y}) = \sup_t \max_i |x_i(t) - y_i(t)|$ and Theorem 2.5.6 applies to prove the existence and uniqueness of a fixed point in $\mathbb{S}$, concluding the proof. $\square$

Proposition 5.2.17 states a sufficient condition for existence and uniqueness of a periodic synchronized solution. We will see now that this condition actually suffices for the local asymptotic stability of $\mathbb{1}\zeta$.

Theorem 5.2.18. *Let Assumptions 5.2.15 and 5.2.16 hold. The synchronized solution $\mathbb{1}\zeta(t)$ of Proposition 5.2.17 is locally exponentially stable if*

$$\sup_{t\geq 0}\int_{-T}^{0}\int_{t+s}^{t}\left[\sum_j \frac{\partial g_{ij}(w,\zeta(w))}{\partial x}\right] dwp(s)\,ds < 1.$$

Proof. We will make a first variation orbital stability argument. Assumption 5.2.15 imply that the right hand-side of (5.24) has continuous first order partial derivatives globally. Let $\mathbb{1}\zeta(t)$ be the T-periodic solution of (5.24) defined from Proposition 5.2.17 and $\mathbf{x}(t,0,\boldsymbol{\phi})$ a solution of (5.24) so that $\boldsymbol{\phi}$ is in the vicinity of $\mathbb{1}\zeta$. For $t\geq 0$, set $z(t)=\max_i|x_i(t)-\zeta(t)|$ and take $\frac{d}{dt}$ to be the right Dini derivative. If $x_i(t)-\zeta(t)\geq 0$ then the Taylor theorem implies that

$$\begin{aligned}
&\frac{dt}{dt}z(t) = \dot{x}_i(t)-\dot{\zeta}(t) = \\
&= -\sum_j\Big(g_{ij}\big(t,x_i(t)\big)-g_{ij}\big(t,\zeta(t)\big)\Big)+\\
&\qquad\qquad +\sum_j\int_{t-T}^{t}\Big(g_{ij}\big(g(s,x_j(s)\big)-g_{ij}\big(s,\zeta(s)\big)\Big)p(s-t)\,ds\\
&= -\sum_j\frac{\partial g_{ij}\big(t,\zeta(t)\big)}{\partial x}z(t)+\\
&\qquad\qquad +\sum_j\int_{t-T}^{t}\left(\frac{\partial g_{ij}\big(s,\zeta(s)\big)}{\partial x}\big(x_j(s)-\zeta(s)\big)\right)p(s-t)\,ds+o(|z|)\\
&\leq -\sum_j\frac{\partial g_{ij}\big(t,\zeta(t)\big)}{\partial x}z(t)+\int_{t-T}^{t}\sum_j\frac{\partial g_{ij}\big(s,\zeta(s)\big)}{\partial x}z(s)p(s-t)\,ds+o(|z|).
\end{aligned}$$

A similar argumentation for $x_i(t)-\zeta(t)<0$ yields the same upper bound for $\dot{z}(t)$. Consequently, for initial data near the periodic orbit we omit the higher order terms $o(|z|)$ and observe that $\dot{z}(t)\leq\dot{q}(t)$ for

$$\dot{q}(t)=-\sum_j\frac{\partial g_{ij}\big(t,\zeta(t)\big)}{\partial x}q(t)+\int_{t-T}^{t}\sum_j\frac{\partial g_{ij}\big(s,\zeta(s)\big)}{\partial x}q(s)p(s-t)\,ds$$

In view of (5.20) we write

$$\frac{d}{dt}q(t) = -\frac{d}{dt}\int_{-T}^{0}\int_{t+s}^{t}\Big[\sum_j \frac{\partial g_{ij}\big(w,\zeta(w)\big)}{\partial x}\Big]q(w)\,dwp(s)\,ds$$

and this will yield

$$q(t) = -\int_{-T}^{0}\int_{t+s}^{t}\Big[\sum_j \frac{\partial g_{ij}\big(w,\zeta(w)\big)}{\partial x}\Big]q(w)\,dwp(s)\,ds + q_0 \tag{5.26}$$

with $q_0 = q(0) + \int_{-T}^{0}\int_{t+s}^{t}\big[\sum_j \frac{\partial g_{ij}(w,\zeta(w))}{\partial x}\big]q(w)\,dwp(s)\,ds$. For $z(\cdot)$ as defined above and $\chi > 0$ we consider the functional space

$$\mathbb{V} = \{v(t) \in C^0([-T,\infty),\mathbb{R}) : v(s) = z(s)|_{s\in[-T,0]}\ \&\ \sup_{t\geq 0} e^{\chi t}|q(t)| < \infty\}$$

which, together with the weighted metric, $\rho(v_1,v_2) = \sup_t e^{\chi t}|v_1(t) - v_2(t)|$ constitutes a complete metric space (see Proposition 2.5.7 of §2.5 in Chapter 2 or [34]). Define the mapping $\mathcal{Q} : \mathbb{V} \to \mathbb{B}$

$$(\mathcal{Q}v) = \begin{cases} z(t), & t \in [-T,0] \\ q_{(5.26)}(t), & t \geq 0 \end{cases}$$

where $q_{(5.26)}(t)$ stands for the right hand-side of (5.26). It is easy to see that $\mathcal{Q} : \mathbb{V} \to \mathbb{V}$ and under the imposed condition one can pick $\chi > 0$ small enough so that

$$\sup_{t\geq 0} e^{\chi t}\int_{-T}^{0}\int_{t+s}^{t}\Big[\sum_j \frac{\partial g_{ij}(w,\zeta(w))}{\partial x}\Big]e^{-\chi w}\,dwp(s)\,ds < 1.$$

Then $\mathcal{Q}$ becomes a contraction in $\mathbb{V}$ and Theorem 2.5.6 applies to ensure a unique fixed point. So $q(t)$ converges to 0 exponentially fast and so does $\zeta(t)$. □

Example 5.2.19. Consider the 3×3 network

$$\dot{x}_i = -\sum_{j=1}^{3} a_{ij}(t)g_{ij}\big(x_i(t)\big) + \int_{t-1}^{t}\sum_{j=1}^{3} a_{ij}(s)g_{ij}\big(x_j(s)\big)p(s)\,ds$$

with

$$G(x) = \bar{g} \begin{bmatrix} 0 & 0.01x + \frac{x^3}{1+x^2} & 3x + \sin^2(x) \\ 0.01x + \frac{x^3}{1+x^2} & 0 & 3x + \sin^2(x) \\ 3x + \sin^2(x) & \frac{x^2}{1+x^2} & 0 \end{bmatrix}$$

for some $\bar{g} > 0$ and

$$A(t) = \begin{bmatrix} 0 & 2 + \sin(2\pi t) & 3 + \sin(4\pi t) \\ 2 + \sin(2\pi t) & 0 & 3 + \sin(4\pi t) \\ 2 + \sin(2\pi t) & 3 + \sin(4\pi t) & 0 \end{bmatrix}.$$

It can be easily verified that the system satisfies Assumptions 5.2.15 and 5.2.16. Moreover, if $p(s) \equiv 1$, (5.20) is satisfied, as well. Choosing $\bar{g} < \frac{1}{6.14}$ both the conditions of Proposition 5.2.17 and Theorem 5.2.18 hold and this means that there is a unique solution that is exponentially stable with rate $\chi = 0.0015$. See Figure 5.2 for a numerical calculation of the solution. A simulation study of the solutions of the network reveals that the upper bound of $\bar{g}$ is conservative. Indeed the monotonicity of g_{ij} suggests that periodic solutions exist and are asymptotically stable for arbitrary values of $\bar{g}$ and arbitrary initial data.

5.3 Networks Of Neutral Type

A finite population of autonomous agents is connected over a linear time invariant communication network with the corresponding graph to be sufficiently connected as discussed in §2.2. Based on this (nominal) system we consider its neutral variation. Our aim is to establish sufficient conditions for asymptotic convergence to a constant value via a stability in variation argument and application of Theorem

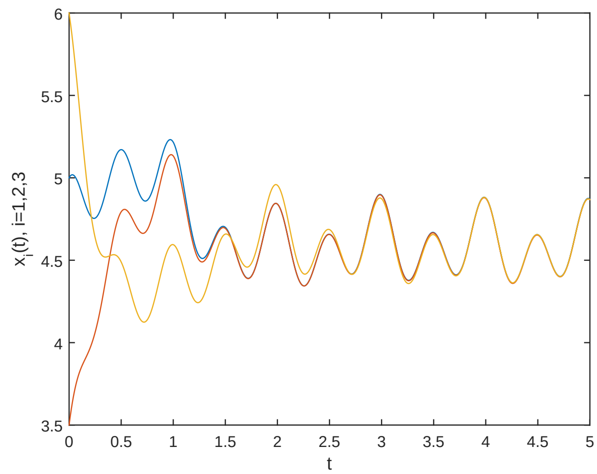

Figure 5.2: Simulation of Example 5.2.19 with the `ddesd` routine in `MATLAB`.

2.5.9. More specifically, a network of $N < \infty$ autonomous agents evolves its state $\mathbf{x} = (x_1, \ldots, x_N)^T$ according to

$$i \in [N] : \begin{cases} \frac{d}{dt}\big(x_i(t) + \int_{-\tau}^{0} f_i(x_i(t+s))p_i(s)\, ds\big) = \sum_j a_{ij}\big(x_j(t) - x_i(t)\big), t \geq 0 \\ x_i(t) = \phi_i(t), t = [-\tau, 0] \end{cases}$$

for

$$\int_{-\tau}^{0} p_i(s)\, ds \equiv 1.$$

The latter condition models a delay, the uncertainty of which implies the integrable distribution function $p_i : [-r, 0] \to \mathbb{R}$. Equivalently

$$\frac{d}{dt}\left(x_i(t) + \int_{-\tau}^{0} \tilde{f}_i(x_i(t+s))p_i(s)\, ds\right) = \sum_j a_{ij}\big(x_j(t) - x_i(t)\big)$$

where $\tilde{f}_i(x_i(t+s)) = f_i(x_i(t+s)) - f_i(k)$ for some real constant k to play the consensus point and it is to be determined below. All in all, we arrive at the following initial value problem

$$\begin{cases} \frac{d}{dt}\left(\mathbf{x} + \int_{t-\tau}^{t} \tilde{\mathbf{F}}\big(\mathbf{x}(q), \mathbf{p}(q-t)\big)\, dq\right) = -L\mathbf{x},\ t \geq 0 \\ \\ \mathbf{x}(s) = \boldsymbol{\phi}(s), -\tau \leq t \leq 0. \end{cases} \tag{5.27}$$

At the moment, we know very little about the solution $\mathbf{x} = \mathbf{x}(t, 0, \boldsymbol{\phi})$. We don't know if it exists or if it is unique. The analysis starts the same way as so far. We impose the necessary set of Assumptions that will come in hand.

Assumption 5.3.1. *The communication graph is routed-out branching.*

This is a necessary and sufficient condition for convergence of the ordinary model and all the properties of the Laplacian discussed in the previous section to hold. From this Assumption and Proposition 2.2.1 we know that

$$|e^{-Lt}L|_1 = Ke^{-\Re(\lambda)t} \tag{5.28}$$

for some $K > 0$ that essentially depends on the norm (here we use the 1-norm thus $|\cdot|_1$ denotes the corresponding induced matrix norm) and $\Re\{\lambda\} > 0$ is the second smallest real part of the eigenvalues of L. The left eigenvector associated with the zero eigenvalue is $\mathbf{c}$ and it satisfies $\mathbf{c}^T L = 0$ and $\sum_i c_i = 1$.

Next, we will use a condition on the nonlinear neutral terms f_i, the most reasonable of which is a global Lipschitz condition.

Assumption 5.3.2. *For every $i = 1, \ldots, N$, $f_i : \mathbb{R} \to \mathbb{R}$ is integrable and there exist $U_i \in \mathbb{R}_+$ such that*

$$|f_i(x) - f_i(y)| \leq U_i|x - y|, \quad \forall x, y \in \mathbb{R}.$$

Although reasonable, this assumption is at the same time very restrictive. Possible extensions and relaxations are to be discussed in the discussion section, below. For the moment we keep in mind that such a condition at least ensures uniqueness of a solution [59]. The existence (in the large) property is yet to be established together with stability.

Assumption 5.3.3.

$$\left(1 + \frac{K}{\Re(\lambda) - \gamma}\right) \max_i \left\{ U_i \int_{-\tau}^{0} |p_i(q)|\, dq \right\} < 1.$$

Theorem 5.3.4. *Let Assumptions 5.3.1, 5.3.2 and 5.3.3 hold. If in addition* $\sum_i c_i U_i < 1$ *then the solution* $\mathbf{x}(t) = \mathbf{x}(t, 0, \boldsymbol{\phi})$ *of* (5.27) *satisfies*

$$||\mathbf{x}(t) - \mathbb{1}k||_1 \leq Ce^{-\gamma t}$$

where k *is the unique solution of* (5.31), $C > ||\boldsymbol{\phi}(0) - \mathbb{1}k||_1$ *is some finite constant and* $\gamma < \Re(\lambda)$ *small that explicitly depends on the systems parameters.*

The proof of this result is an application of Theorem 2.5.9. For this we need some preparatory steps:

Derivation of the solution operator   Using the standard variations of constants formula we see that the solution $\mathbf{x}$ of (5.27) satisfies

$$\begin{aligned} \mathbf{x}(t) &= e^{-Lt}\boldsymbol{\phi}(0) - \int_0^t e^{-L(t-s)} \frac{d}{ds} \int_{s-\tau}^{s} \tilde{\mathbf{F}}\big(\mathbf{x}(q), \mathbf{p}(q-s)\big)\, dq ds \\ &= e^{-Lt}\tilde{\boldsymbol{\phi}} - \int_{t-\tau}^{t} \tilde{\mathbf{F}}\big(\mathbf{x}(q), \mathbf{p}(q-t)\big)\, dq + \int_0^t e^{-L(t-s)} L \int_{s-\tau}^{s} \tilde{\mathbf{F}}\big(\mathbf{x}(q), \mathbf{p}(q-s)\big)\, dq ds \end{aligned} \tag{5.29}$$

where $\tilde{\boldsymbol{\phi}} = \boldsymbol{\phi}(0) + \int_{-\tau}^{0} \tilde{\mathbf{F}}\big(\boldsymbol{\phi}(q), \mathbf{p}(q)\big)\, dq$.

The space of solutions Let $C^0 = C([-\tau, \infty), \mathbb{R}^N)$ be a subspace an appropriate Banach space $(\mathbb{B}, |\cdot|)$ of bounded continuous functions. This set is defined in $[-\tau, \infty)$ and takes values in $\mathbb{R}^N$. For fixed $\boldsymbol{\phi} \in C^0([-\tau, 0], \mathbb{R}^N)$, $k \in \mathbb{R}$, $C > 0$ and $\gamma > 0$ we define the following set

$$\mathbb{M} = \left\{ \mathbf{z} \in C^0 : \mathbf{z}(s) = \boldsymbol{\phi}(s)|_{s\in[-\tau,0]},\ \sup_{t\geq 0} e^{\gamma t} \left|\left|\mathbf{z}(t) - \mathbb{1}k\right|\right|_1 \leq C \right\} \tag{5.30}$$

This paragraph ends with the following technical result:

Lemma 5.3.5. *The set* $\mathbb{M}$ *as defined in* (5.30) *is closed, convex and non-empty, if* $C \geq ||\boldsymbol{\phi}(0) - \mathbb{1}k||_1$.

Proof. The set is obviously closed because it is constructed to contain all of its limit points. $\mathbb{M}$ is also convex: For any pair $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{M}$, $\mathbf{z}_3 := \beta\mathbf{z}_1 + (1-\beta)\mathbf{z}_2$ is also a member of $\mathbb{M}$ for any $\beta \in [0, 1]$. Indeed $\mathbf{z}_3(t) = \beta\boldsymbol{\phi}(t) + (1-\beta)\boldsymbol{\phi}(t) = \boldsymbol{\phi}(t)$ for $-\tau \leq t \leq 0$ and for $\mathbf{z}_3(t) - \mathbb{1}k = \beta(\mathbf{z}_1(t) - \mathbb{1}k) + (1-\beta)(\mathbf{z}_2(t) - \mathbb{1}k)$ it holds that $e^{\gamma t}||\mathbf{z}_3(t) - \mathbb{1}k||_1 \leq \beta C + (1-\beta)C = C$. Finally, under the imposed condition, the function

$$\mathbf{z}(t) = \begin{cases} \boldsymbol{\phi}(t), & t \in [-T, 0] \\ \mathbb{1}k + (\boldsymbol{\phi}(0) - \mathbb{1}k)e^{-\gamma t}, & t \geq 0 \end{cases}$$

is a member of $\mathbb{M}$ so the set is not empty. $\square$

The consensus point

Lemma 5.3.6. *Let Assumptions 5.3.2 and 5.3.1 hold. If*

$$\sum_i c_i U_i < 1$$

then there exists a unique $k \in \mathbb{R}$ such that

$$k = \sum_i c_i \left(\phi_i(0) + \int_{-\tau}^{0} \big(f_i(\phi_i(q)) - f_i(k)\big) p_i(q)\, dq \right) \tag{5.31}$$

Proof. Consider the complete metric space $(\mathbb{R}, \rho)$ where $\rho(x, y) = |x - y|$ is the standard distance between two points on the line. Then it is easily seen that for the operator $F(k) = \sum_i c_i\big(\phi_i(0) + \int_{-\tau}^{0} \big(f_i(\phi_i(-q)) - f_i(k)\big)p_i(q)\, dq\big)$ that maps $\mathbb{R}$ into itself

$$\rho(F(k_1), F(k_2)) \leq \left(\sum_i c_i U_i \right) \rho(k_1, k_2),\ \ \forall k_1, k_2 \in \mathbb{R}$$

and the result follows from Theorem 2.5.6. □

Now that we have obtained these easy yet significant results we can proceed to the full proof of Theorem 5.3.4.

Proof of Theorem 5.3.4. The proof of our main result is based on Theorem 2.5.9. Having defined k and established a first estimate of C we are ready to further elaborate on our solution space and the solution operator. The first step is to show that the function $\mathcal{P} : \mathbb{M} \to \mathbb{R}$

$$(\mathcal{P}\mathbf{z})(t) = \begin{cases} \boldsymbol{\phi}(t), & -\tau \leq t \leq 0 \\ \\ \mathbf{z}_{(5.29)}(t), & t \geq 0 \end{cases} \tag{5.32}$$

is under conditions an operator $\mathcal{P} : \mathbb{M} \to \mathbb{M}$. The first step towards this is to examine $\lim_{t\to\infty}(\mathcal{P}\mathbf{z})(t)$. Indeed we see that

$$e^{-Lt}\tilde{\boldsymbol{\phi}} \to \mathbb{1}\mathbf{c}^T\tilde{\boldsymbol{\phi}} = \mathbb{1}\mathbf{c}^T\tilde{\boldsymbol{\phi}}(k)$$

but

$$\int_{t-\tau}^{t} \tilde{\mathbf{F}}\big(\mathbf{z}(q), \mathbf{p}(q-t)\big)\, dq \to 0$$

and

$$\int_0^t e^{-L(t-s)} L \int_{s-\tau}^{s} \tilde{\mathbf{F}}\big(\mathbf{z}(q), \mathbf{p}(q-s)\big)\, dq ds \to 0$$

exactly because $e^{-Lt}L$ is an L^1 function and the last equation is justified as it is the convolution of an L^1 function with a function that goes to zero. Finally,

$$\lim_{t\to\infty} (\mathcal{P}\mathbf{z})(t) = \mathbb{1} \sum_i c_i \bigg(\phi_i(0) + \int_{-\tau}^{0} \big(f_i(\phi_i(-q)) - f_i(k) \big)\, dq \bigg)$$

so that if k is defined as in (5.31) we conclude on

$$(\mathcal{P}\mathbf{z})(t) \to \mathbb{1}k$$

Now we define two operators $\mathcal{A}, \mathcal{B} : \mathbb{M} \to \mathbb{B}$ as follows:

$$\begin{aligned}(\mathcal{A}\mathbf{z})(t) &= \int_0^t e^{-L(t-s)} L \int_{s-\tau}^{s} \tilde{\mathbf{F}}\big(\mathbf{z}(q), \mathbf{p}(q-s)\big)\, dq ds \\ (\mathcal{B}\mathbf{z})(t) &= e^{-Lt}\tilde{\boldsymbol{\phi}} - \int_{t-\tau}^{t} \tilde{\mathbf{F}}\big(\mathbf{x}(q), \mathbf{p}(q-t)\big)\, dq\end{aligned}$$

We now proceed to check the conditions of Theorem 2.5.9 one by one:

Condition (i) Let $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{M}$. Then $(\mathcal{A}\mathbf{z}_1)(t) + (\mathcal{B}\mathbf{z}_2)(t)$ behaves in the limit exactly as $(\mathcal{P}\mathbf{z})(t)$, simply because it is only the time-varying (state-independent) part that contributes to the limit point. Hence, since $\mathbf{z}_1 \equiv \mathbf{z}_2$ in $[-\tau, 0]$, it is only left to prove the convergence estimate: Note that if $\gamma < \Re(\lambda)$ simple calculations yield

$$\sup_t e^{\gamma t} ||(\mathcal{A}\mathbf{z}_1)(t)||_1 \leq \frac{K \max_i U_i \int_{-\tau}^{0} e^{-\gamma q} p_i(q)\, dq}{\Re\{\lambda\} - \gamma} C$$

and

$$\sup_t e^{\gamma t} ||(\mathcal{B}\mathbf{z}_2)(t)||_1 \leq ||\tilde{\boldsymbol{\phi}}||_1 + \max_i U_i \int_{-\tau}^{0} e^{-\gamma q} p_i(q)\, dq C$$

so from Assumption 5.3.3 the first condition of Krasnoselskii's Theorem is satisfied as this way it is always possible to pick C large enough so that

$$\sup_t e^{\gamma t} ||(\mathcal{A}\mathbf{z}_1)(t)|| + \sup_t e^{\gamma t} ||(\mathcal{B}\mathbf{z}_2)(t)||_1 \leq C.$$

In fact it suffices to pick $C > \max\{||\boldsymbol{\phi}(0) - \mathbb{1}k||, D\}$ where

$$D = \frac{||\tilde{\boldsymbol{\phi}}||_1}{1 - \left(\frac{K}{\Re\{\lambda\} - \gamma} + 1\right) \max_i U_i \int_{-\tau}^{0} e^{-\gamma q} p_i(q)\, dq} > 0$$

for γ small enough again in view of Assumption 5.3.3.

Condition (ii) We note that $\mathcal{A}\mathbb{M}$ is a subset of $\mathbb{B}$ since it maps $\mathbb{M}$ to a subset of functions which vanish to zero as fast as $e^{-\gamma t}$, in view of $\gamma < Re\{\lambda\}$. It really suffices to show that $\mathcal{A}\mathbb{M}$ is equicontinuous because then it follows that it is continuous with respect to the supremum norm in $\mathbb{B}$. The former can be shown by differentiating $(\mathcal{A}\mathbf{z})(t)$ with respect to t:

$$\frac{d}{dt}(\mathcal{A}\mathbf{z})(t) = \int_{-\tau}^{0} \tilde{\mathbf{F}}\big(\mathbf{z}(t+q), \mathbf{p}(q)\big)\, dq - \int_0^t e^{-L(t-s)} L^2 \int_{-\tau}^{0} \tilde{\mathbf{F}}\big(\mathbf{z}(s+q), \mathbf{p}(q)\big)\, dq ds$$

and it is only a tedious algebraic exercise to show that for $\mathbf{z} \in \mathbb{M}$

$$\sup_t \left|\left| \frac{d}{dt}(\mathcal{A}\mathbf{z})(t) \right|\right|_1 \leq C'$$

a constant that is independent of the element $\mathbf{z}$ and depends only on $\mathbb{M}$. This uniform condition implies equi-continuity and hence Proposition 2.5.4 applies to show that $\mathcal{A}$ is a compact map that is also continuous.

Condition (iii) Now, since $\mathbb{M}$ is a closed subset of $\mathbb{B}$ it also constitutes a (complete) metric space under the weighted metric

$$\rho(\mathbf{z}_1, \mathbf{z}_1) = \sup_t e^{\gamma t} ||\mathbf{z}_1(t) - \mathbf{z}_2(t)||_1$$

Then

$$\rho\big(\mathcal{B}\mathbf{z}_1, \mathcal{B}\mathbf{z}_2\big) \leq \left[\max_i U_i \int_{-\tau}^{0} e^{-\gamma q} |p_i(q)| \, dq\right] \rho(\mathbf{z}_1, \mathbf{z}_1)$$

which is automatically a contraction in view Assumption 5.3.3. Then every condition of Theorem 2.5.9 is satisfied, hence $\mathcal{P} = \mathcal{A} + \mathcal{B}$ has as fixed point in $\mathbb{M}$. $\square$

We would like comment on conditions as imposed by the Assumption 5.3.2. Taking a look in (5.31) we are tempted to consider for a moment a linear version of $f_i(x) = -|U_i|x$. Then the consensus point is

$$k = \frac{\sum_i c_i \phi_i(0) - \sum_i c_i |U_i| \int_{-r}^{0} \phi(q) p_i(q) \, ds}{1 - \sum_i c_i |U_i|}$$

and this creates instability $k = \infty$ at values of $|K_i|$ close to 1. We conclude that, so long as, we are searching for asymptotic consensus solutions the smallness on U_i is not unnecessarily strict.

Finally, Assumption (5.3.3) is undoubtedly the hardest one as it imposes additional restricting conditions on both U_i and τ, which is the result of the stability in variation argument. It is shown however that such strict conditions occur very regularly in the literature and examples can be constructed that justify them for the sake of stabilization of solutions (see also [59]).

5.4 Supplementary Remarks

We approached a number of significant extensions of linear consensus networks. Whenever the nonlinear systems incorporate features that support the cooperative behavior among agents, such as the passivity or the monotonic hypotheses, the linear theory machinery applies with minor modifications.

The passive systems bear the greatest resemblance with the linear ones while the general nonlinear ones, presented in §5.2 are separated into two categories. Our analysis distinguishes the co-operative (monotonic) mode from non-cooperative (nonmotonotnic) one and convergence results were provided for both. The monotonic mode is characterized by the non-negative first derivative the functions $g_{ij}(t,x)$ with respect to x and it is a straightforward generalization of the linear case and can be analyzed with the use of the same fundamental tools from the non-negative matrix theory. Although a simplified increased connectivity regime was considered, the results can be easily extended to the case of simple/recurrent connectivity as follows. The double inequality in (5.14) can be naturally extended as

$$\mathbf{x}(t) \geq \int_{t-Q}^{t} A(t,t_1) \int_{t_1-Q}^{t_1} A(t_1,t_2) \cdots \int_{t_{\sigma-1}-Q}^{t_{\sigma-1}} A(t_{\sigma-1},t_\sigma)\mathbf{x}_{t_{1,\sigma}}\, dt_\sigma dt_{\sigma-1} \cdots dt_1$$

$$\mathbf{x}(t) \leq \int_{t-Q}^{t} B(t,t_1) \int_{t_1-Q}^{t_1} B(t_1,t_2) \cdots \int_{t_{\sigma-1}-Q}^{t_{\sigma-1}} B(t_{\sigma-1},t_\sigma)\mathbf{x}_{t_{1,\sigma}}\, dt_\sigma dt_{\sigma-1} \cdots dt_1$$

since $A(t,s), B(t,s) \geq 0$. The notation $\mathbf{x}_{t_{1,\sigma}}$ follows the one in the proof of Theorem 5.2.8. The analysis continues for the lower and upper product of matrices, each of which can be shown to be stochastic.

The autonomous case also provides information on the consensus point. The monotonicity condition implies that the consensus point serves as a unique solution of a nonlinear algebraic equation.

The competitive model is studied without delays, through a stability in variation and a fixed point argument. The results rely on the "smallness" of the nonlinear effect and are essentially considered as a perturbation to a monotonic linear system. Simulations suggest that instability occurs when the non-cooperative mode prevails. In the particular 2×2 example, the cooperative agent tries to follow the non-cooperative one. Stability then can only occur when the rate at which the former agent cooperates is faster than the rate the latter agent diverges.

In the next section we pointed out that although the standard type in synchronization solutions of consensus systems is the constant solutions, in an interesting turn of events we showed that for the special type of distributed delays, nonlinearity provides an alternative: the existence and asymptotic stability of synchronized yet non-constant periodic solutions. We proved the existence and uniqueness of a periodic solution with a fixed point theorem approach and its local stability using the standard variational approach.

For the solution of neutral networks, we developed a fixed point theory argument based on a combination of the contraction mapping principle and Krasnoselskii's result on perturbed operators. The derived sufficient conditions are based on the smallness of the delays and/or the Lipschitz constants. The linearity of the problem produced very elegant results that characterize both the convergence, the rate as well as the consensus point. Our primary goal was to open the subject of neutral distributed networks. We claim that proving simple convergence to a common constant is the first step yet less exciting phenomenon a research might come across. Theorem 5.3.4 is a combination of conservative imposed conditions and the over-simplistic static communication network. We conjecture that the methods developed in this work can be adapted to the study of networks with more realistic and more interesting neutral components or networks with time varying or nonlinear couplings or even delayed state arguments. Then one could investigate the existence (and perhaps stability) of more interesting asymptotic phenomena such as periodic or chaotic solutions.

Another serious difficulty is the derivation of the solution operator. Although the method of variation of parameters in dynamical systems is very popular and well-studied over the years [68], experience has proved that regardless if we are to follow a Lyapunov or a Fixed Point Method, stability problems are to be studied on a case by case basis. In our model, the nominal system, exhibits extraordinary robust stability, but the derivation of the solution operator to a useful form required integration by parts step. It is not clear how one could proceed for example if the nominal system, involved propagation delays or if the nominal system was nonlinear. In the latter case we would be forced to use a nonlinear variation of parameters formula, that would result in excessive technical problems.

Chapter 6: Flocking Networks

In a series of papers, [52, 88] Felipe Cucker and Steve Smale introduced an interesting communication scheme for speed alignment among a finite population of birds. In its elementary version the model reads

$$i \in [N] : \begin{cases} \dot{x}_i = u_i \\ \dot{u}_i = \sum_j a_{ij}(\mathbf{x})(u_j - u_i) \\ x_i(0) = x_i^0, \ u_i(0) = u_i^0, \ \text{given} \end{cases} \tag{6.1}$$

where x_i denotes the position of the bird i and u_i denotes its speed. Therefore, a bird $i \in [N]$ is an autonomous agent the state of which is adequately defined with the pair (x_i, u_i). Without loss of generality we will assume $x_i, u_i \in \mathbb{R}$ so that the state vectors

$$\mathbf{x} = (x_1, \ldots, x_N) \quad \text{and} \quad \mathbf{u} = (u_1, \ldots, u_N) \quad \in \mathbb{R}^N.$$

The communication rate is assumed to have the explicit form

$$a_{ij}(\mathbf{x}) = a_{ij}^{CS}(\mathbf{x}) = \frac{K}{(\alpha^2 + |x_i - x_j|^\beta)}. \tag{6.2}$$

The idea behind (6.2) is based on the intuitively reasonable assumption that the further the bird j is from bird i the weaker the effect on one another should be.

Therefore, the interest concentrates on the solutions $(\mathbf{x}, \mathbf{u})$ that satisfy the so-called *asymptotic flocking* condition, first introduced in [89]:

Definition 6.0.1. The asymptotic flocking condition for the solution $(\mathbf{x}, \mathbf{u})$ is

$$|u_i(t) - u_j(t)| \to 0 \qquad \text{and} \qquad \sup_t |x_i(t) - x_j(t)| < \infty$$

The goal is to derive sufficient conditions for asymptotic flocking. These will typically be an expression that involves both the system's parameters and the initial configuration $\mathbf{x}^0, \mathbf{u}^0$. If such a condition is true, then Definition 6.0.1 implies that the flock of birds will align its speed while at the same time it remains in shape. The first result on the model is presented below:

Theorem 6.0.2. *Let the system* (6.1) *with* a_{ij} *as defined in* (6.2) *with its solution* $(\mathbf{x}, \mathbf{u})$. *Assume that one of the three hypotheses holds*

1. $\beta < 1/2$.

2. $\beta = 1/2$ *and* $\frac{1}{2}\sum_{i \neq j}(u_i^0 - u_j^0)^2 < (\frac{K}{3N})^2/8$.

3. $\beta > 1/2$ *and*

$$\left[\left(\frac{1}{2\beta}\right)^{\frac{1}{2\beta-1}} - \left(\frac{1}{2\beta}\right)^{\frac{2\beta}{2\beta-1}}\right] > \sum_{i \neq j}(x_i^0 - x_j^0)^2 + \alpha^2.$$

Then $(\mathbf{x}, \mathbf{u})$ *satisfies the flocking condition according to Definition 6.0.1.*

Proof. See [52, 88]. □

These systems are known in the literature as 2^{nd} order consensus models and they have attracted enormous attention from the Applied Mathematics community

[90, 91, 92, 93, 94, 89, 55], to name a few. It was, in fact, very soon realized that the model could accept a number of non-trivial improvements. Firstly, the idea behind (6.2) is due to an early work of the authors on language evolution [95] and the proof of Theorem 6.0.2 is based on this particular form of communication weights. This is a restrictive condition that was remedied with the introduction of a Lyapunov functional in [89] that produces clear and elegant sufficient conditions for asymptotic flocking with generic symmetric connectivity weights. On the other hand, recent results support the claim that birds make use of topological rather than euclidean metrics as they attempt to synchronize their velocity [96]. Based on this assumption, in [97] a significant improvement of (6.1) is proposed by considering the coupling

$$a_{ij}(\mathbf{x}) = a_{ij}^{MT}(\mathbf{x}) = \frac{\psi(|x_i - x_j|)}{\sum_j \psi(|x_i - x_j|)}, \tag{6.3}$$

for an arbitrary non-negative function $\psi(y)$ such that $\lim_{y\to\infty} \phi(y) = 0$ just as a_{ij}^{CS} does. This is a state-dependent, non-symmetric coupling that generates a consensus system, the study of which requires the use of the contraction coefficient κ [55]. The authors could not however avoid asking for small couplings such that $\sum_j a_{ij} \equiv 1$ as well as increased (Type I) connectivity.

We conclude this introductory note with reviewing works on stochastic/noisy flocking networks. A stochastic 2^{nd} order consensus family, different from the one discussed in §3.3, is proposed for the 2^{nd} order model in [90, 92, 93]. These models, known as noisy-flocking dynamics, involve additive Brownian noise in the equations and sufficient conditions for consensus in the almost sure sense are derived. In [93] the authors study the following stochastic variation of (6.1)

$$\begin{cases} dx_i = u_i dt \\ du_i = \sum_i a(|x_i - x_j|^2)\big(u_j - u_i\big)dt + \sum_j c_{ij} dB_j, t \geq t_0 \end{cases} \tag{6.4}$$

subject to given initial data x_i^0 and u_i^0, where dB_j stands for white noise induced from the agent j and weighted by c_{ij} when agent i updates its speed. The objective is to obtain conditions for asymptotic flocking in the mean-square sense based on the smallness of the noisy perturbation. Their argument relies on the algebraic properties of the symmetric communication weights (of a_{ij}^{CS} type) so that a Lyapunov stability argument for stochastic stability is developed.

Contribution We apply the results and methods developed in the previous chapter to derive sufficient conditions for asymptotic flocking on a number of old and new flocking networks. The models we are concerned here are the standard second order consensus flocking model with asymmetric coupling rates and simple/switching connectivity. The analysis is based on Theorems 2.3.1, 3.2.9, respectively. After taking a small digression commenting on the discrete time version of our model, we elevate the analysis to completely non-linear flocking networks regarding the monotonic nonlinear first order model of Chapter 5. Next, we comment on delayed versions of the flocking networks applying the Results of Chapter 4 and the discussion is concluded with results concerning the stochastic/noisy versions of flocking were conditions for asymptotic flocking in the almost sure sense and the r^{th} mean.

6.1 Ordinary Flocking

This section is divided into two parts. The first one concerns the classic state-dependent flocking network in (6.1). We essentially revisit this scheme and apply the results of Chapter 3 to derive general sufficient conditions for flocking in the event of Type I or Type II connectivity. We point out that despite the non-linear dependence of the coupling strength on the relative distance, the algorithm remains essentially linear. In the next section we proceed to introduce a non-linear version of (6.1) by applying the results of Chapter 5. We will see that the conditions are fairly similar.

6.1.1 Linear averaging

A finite population of N birds act as a group of autonomous agents and exchange information according to

$$i \in [N] : \begin{cases} \dot{x}_i(t) = u_i(t) \\ \dot{u}_i(t) = \sum_{j \in N_i} a_{ij}(t, \mathbf{x})\big(u_j(t) - u_i(t)\big) \\ x_i(0) = x_i^0, u_i(0) = u_i^0, \text{ given.} \end{cases} \tag{6.5}$$

In this work, for a solution $(\mathbf{x}, \mathbf{u})$ of (6.5) we assume

$$a_{ij}\big(t, \mathbf{x}(t)\big) \neq 0 \Rightarrow a_{ij}\big(t, \mathbf{x}(t)\big) \geq f\big(S(\mathbf{x}(t))\big) \tag{6.6}$$

for some positive non-increasing function $f(\cdot)$ with the property that

$$f(y) \to 0 \text{ as } y \to \infty \tag{6.7}$$

and

$$\sup_{t,y} a_{ij}(t,y) \le a < \infty. \tag{6.8}$$

The objective is to derive sufficient conditions so that the solution $(\mathbf{x}, \mathbf{u})$ of (6.5) exhibits asymptotic flocking in the sense of Definition. 6.0.1. The problem clearly fits the framework developed in Ch. 3. Firstly, Lemmas 3.2.2 and 3.2.3, apply directly to ensure boundedness of the $\mathbf{u}$, therefore existence in the large, as well as, that the first part of Definition 6.0.1 implies $\mathbf{u}(t) \to \Delta$ as $t \to \infty$. Secondly, the derivation of sufficient conditions for asymptotic flocking will be based on the contraction rates obtained in §3.2. These rates effectively depend on the different types of connectivity schemes that were thoroughly discussed in §3.1. We restate these schemes here for quick reference but also slightly modified for the purposes of the particular analysis.

1. TYPE I & static: Here there is always one bird i to affect the rest of flock, all the time. There is no switching or failing signals. If a connection between two birds exists for some $t \ge t_0$, then it does so for all times.

2. TYPE II & static: This is a relaxed, simple routed-out branching connectivity condition. While no switching/failing signaling is assumed, there is no central agent (bird) to affect the rest of the flock at the same time. This is a static and decentralized scheme.

3. TYPE II & switching: This is the mildest connectivity assumption where we allow switching connectivity. We recall the Assumptions 3.2.7, 3.2.8 and the Remark 2.3.7 with the notation used there.

Finally from (6.8) we define $m := \max_i \sup_{t,y} \sum_j a_{ij}(t,y) < \infty$ and we are now ready to state the first result of this chapter.

Theorem 6.1.1. *Consider* (6.5) *with conditions* (6.6), (6.7) *and* (6.8) *to be true and let* $\big(\mathbf{x},\mathbf{u}\big)$ *be its solution.*

1. *TYPE I & static: The solution exhibits asymptotic flocking if*

$$S(\mathbf{u}^0) < \int_{S(\mathbf{x}^0)}^{\infty} f(w)\,dw \tag{6.9}$$

2. *TYPE II & static: The solution exhibits asymptotic flocking if*

$$S(\mathbf{u}^0) < \frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}\gamma B} \int_{P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^{\gamma}(s)\,ds \tag{6.10}$$

where $P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0} = \max\big\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S(\mathbf{u}^0)\gamma B|\big\}$.

3. *TYPE II & switching. Let Assumptions 3.2.7 and 3.2.8 hold. The solution exhibits asymptotic flocking if*

$$S(\mathbf{u}^0) < \frac{(1-e^{-m\epsilon})^{\sigma}}{m^{\sigma}\sigma B} \int_{P^{\sigma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^{\sigma}(s)\,ds \tag{6.11}$$

where $\sigma = l^*([N/2]+1)$, l^* *with the meaning of Remark 2.3.7 and* $\epsilon > 0$ *with the meaning of Assumption 3.2.8.*

Proof. We begin with the first connectivity condition, where there is at least one agent affecting the rest of the group. We follow the same path as in Theorem 3.2.4 for $\mathbf{u}$ and show that

$$\frac{d}{dt}S\big(\mathbf{u}(t)\big) \leq -f\big(S\big(\mathbf{x}(t)\big)\big)S\big(\mathbf{u}(t)\big) \Rightarrow S\big(\mathbf{u}(t)\big) \leq e^{-\int_0^t f(S(\mathbf{x}(w)))\,dw}S\big(\mathbf{u}^0\big)$$

so that asymptotic flocking will occur with exponential rate of convergence if $S\big(\mathbf{x}(t)\big) \leq r$ for some $r > 0$. For this, we follow [55] and introduce the functional

$$V_1(\mathbf{x}, \mathbf{u}) = S(\mathbf{u}) + \int_0^{S(\mathbf{x})} f(w)\, dw \tag{6.12}$$

so that along a solution of (6.5) $\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ we have

$$\frac{d}{dt}V_1(t) = \frac{d}{dt}V_1\big(\mathbf{x}(t), \mathbf{u}(t)\big) \leq -f\big(S\big(\mathbf{x}(t)\big)\big)S\big(\mathbf{u}(t)\big) + f\big(S\big(\mathbf{x}(t)\big)\big)S\big(\mathbf{u}(t)\big) = 0$$

so that $V_1(t) \leq V_1(0)$. This is equivalent to

$$S(\mathbf{u}(t)) + \int_0^{S(\mathbf{x}(t))} f(w)\, dw \leq S(\mathbf{u}^0) + \int_0^{S(\mathbf{x}^0)} f(w)\, dw$$

From the imposed condition (6.9) on the initial data we deduce that there exists r' such that

$$S(\mathbf{u}^0) = \int_{S(\mathbf{x}^0)}^{r'} f(w)\, dw$$

so that $S\big(\mathbf{x}(t)\big) \geq S\big(\mathbf{x}^0\big)$. Now,

$$0 \leq S\big(\mathbf{u}(t)\big) \leq \int_{S(\mathbf{x}^0)}^{r'} f(w)\, dw - \int_{S(\mathbf{x}^0)}^{S(\mathbf{x}(t))} f(w)\, dw$$

which makes sense if $S\big(\mathbf{x}(t)\big) \leq r'$. Pick $r = \max\{r', S(\mathbf{x}^0)\}$ to conclude that condition (6.9) ensures that the flock of birds will remain connected, hence they will coordinate their speeds exponentially fast.

For the second part, the flock is static and routed-out branching, hence it is routed-out branching over the interval $[t - B, t]$, for any $t > 0$ and $B > 0$. Let $W(\mathbf{x}(s)) = mI_{N\times N} - D(\mathbf{x}(s)) + A(\mathbf{x}(s))$ and

$$C(t, s) = e^{-mB}\delta(s - (t - B))I_{N\times N} + e^{-m(t-s)}W(\mathbf{x}(s))$$

For the scrambling index $\gamma \geq 1$ of the topological graph $\mathbb{G}_{P(\mathbf{x}(t))}$ (which is independent of time) $P_B^{(\gamma)}\big(\mathbf{x}(t)\big)$ is stochastic from Proposition 3.2.5 and has the same scrambling index as $P\big(\mathbf{x}(t)\big)$. Since the corresponding graph $\mathbb{G}_W$ is independent of time, so will be the scrambling index γ. Following the proof of Theorem 3.2.9

$$\begin{aligned} S\big(\mathbf{u}(t)\big) &\leq \kappa\big(P_B^{(\gamma)}(\mathbf{x}(t))\big)S(\mathbf{u}\big(t-\gamma B)\big) \\ &\leq \big(1 - cf^{\gamma}(S(\mathbf{x}(t)))\big)S\big(\mathbf{u}(t-\gamma B)\big) \end{aligned} \tag{6.13}$$

with $c := \frac{(1-e^{-mB})^{\gamma}}{m^{\gamma}}$ and $S\big(\mathbf{x}(t)\big) \geq r$ for r such that $f(r) \leq \frac{me^{-mB}}{1-e^{-mB}}$. Define

$$V_2(\mathbf{x},\mathbf{u}) = \int_{t-\gamma B}^{t} S\big(\mathbf{u}(s)\big)\,ds + c\int_0^{S(\mathbf{x})} f^{\gamma}(s)\,ds.$$

The derivative of $\dot{V}_2$ along $\big(\mathbf{x}(t),\mathbf{u}(t)\big)$ is

$$\dot{V}_2(t) = S\big(\mathbf{u}(t)\big) - S\big(\mathbf{u}(t-\gamma B)\big) + cf^{\gamma}\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) \leq 0$$

in view of Lemma 3.2.2 (from which it is deduced that $S\big(\mathbf{u}(t)\big) \leq S\big(\mathbf{u}(t-\gamma B)\big), \forall t$ and (6.13). Then for $t \geq \gamma B$ we have $V_2(t) \leq V_2(\gamma B)$ which is equivalent to

$$\int_{t-\gamma B}^{t} S\big(\mathbf{u}(s)\big)\,ds + c\int_0^{S(\mathbf{x}(t))} f^{\gamma}(s)\,ds \leq \int_0^{\gamma B} S\big(\mathbf{u}(s)\big)\,ds + c\int_0^{S(\mathbf{x}(\gamma B))} f^{\gamma}(s)\,ds$$

Let the following condition hold

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)\,ds < c\int_{S(\mathbf{x}(\gamma B))}^{\infty} f^{\gamma}(s)\,ds \tag{6.14}$$

and we pick r' such that

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)\,ds = c\int_{S(\mathbf{x}(\gamma B))}^{r'} f^{\gamma}(s)\,ds$$

then from the last inequality, $S\big(\mathbf{x}(t)\big) \leq S\big(\mathbf{x}(\gamma B)\big)$ implies

$$0 \leq \int_{S(\mathbf{x}(t))}^{r'} f^{\gamma}(s)\,ds \Rightarrow S\big(\mathbf{x}(t)\big) \leq r'$$

so that the flock remains bounded and exponential speed alignment is ensured. Finally, we show that (6.10) implies (6.14). Indeed,

$$\int_0^{\gamma B} S\big(\mathbf{u}(s)\big)\, ds \leq \gamma B S(\mathbf{u}^0)$$

from Lemma 3.2.2. We look for a lower bound of $S\big(\mathbf{x}(t)\big)$. If $S\big(\mathbf{x}(t)\big) \geq S(\mathbf{x}^0)$ from the form of (6.5) the rate at which $S(\mathbf{x}(t)\big)$ may shrink can be deduced from the extreme initial configuration

$$\mathbf{x}^0 = (x^0, 0, \dots, 0) \quad \& \quad \mathbf{u}^0 = (u^0, \dots, 0)$$

with $x^0, u^0 \neq 0$ so that $S(\mathbf{x}^0) = x^0$ and $S(\mathbf{u}^0) = |u^0|$. Neglecting the averaging effect which will inevitably diminish $S\big(\mathbf{u}(t)\big)$, $x^0 < 0$ implies that the first bird at t will have approached (or bypassed) the rest of the group by $-|x^0| + |u^0|t$. All in all, at $t = \gamma B$

$$S\big(\mathbf{x}(\gamma B)\big) \geq \max\big\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S(\mathbf{u}^0)\gamma B|\big\} = P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}$$

so that

$$\int_{S(\mathbf{x}(\gamma B))}^{\infty} f^\gamma(s)\, ds \geq \int_{P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^\gamma(s)\, ds$$

then

$$S(\mathbf{u}^0) < \frac{(1-e^{-mB})^\gamma}{m^\gamma \gamma B} \int_{P^{\gamma,B}_{\mathbf{x}^0,\mathbf{u}^0}}^{\infty} f^\gamma(s)\, ds.$$

The case of switching connectivity is treated as in Theorem 3.2.9 and V_2 is used in a similar way after substituting γ with σ. Then (6.11) substitutes (6.10) to ensure asymptotic flocking and the proof is concluded. □

Remark 6.1.2. In the case of static connectivity, $\gamma = 1$ implies that (6.9) and (6.10) coincide as $B \downarrow 0$.

Remark 6.1.3. With respect to each connectivity scheme, when either

$$\int^{\infty} f(s)\,ds = \infty, \qquad \int^{\infty} f^{\gamma}(s)\,ds = \infty, \qquad \int^{\infty} f^{\sigma}(s)\,ds = \infty,$$

we have *unconditional* asymptotic flocking.

6.1.1.1 A discrete-time version

We take a brief digression into the discrete time case for flocking networks. Discrete time dynamics are normally easier to handle and they are preferred in computational applications. However, discretization of flocking networks such as (6.1) requires very small mesh step, otherwise they may lead to instability [52]. In discrete time (6.1) reads

$$i \in [N] : \begin{cases} x_i(k+\eta) = x_i(k) + \eta v_i(k) \\ v_i(k+\eta) = v_i(k) + \eta \sum_j a_{ij}\big(\mathbf{x}(k)\big)\big(v_j(k) - v_i(k)\big), \\ x_i(0) = x_i^0,\ v_i(0) = v_i^0,\ \text{given} \end{cases} \tag{6.15}$$

for $k \in \mathbb{Z}_+$ and $\eta > 0$ the fixed mesh value. For simplicity, we assume

$$a \geq a_{ij}(\mathbf{x}) \geq f\big(S(\mathbf{x})\big) \tag{6.16}$$

$\forall j \neq i$ for $a < \infty$ and f a non-negative, non-increasing function. To simplify notation we write $\mathbf{x}(k)$ for $\mathbf{x}(kh)$ and the same for $\mathbf{v}$. The next result stems out of Theorem 3.1.5.

Corollary 6.1.4. *Consider* (6.15) *and its solution* $\big(\mathbf{x}(k), \mathbf{v}(k)\big)$*. Let the assumptions of Theorem 3.1.5 hold. Set*

$$C := \limsup_{x\to\infty} x f(x) \leq \infty \quad \textit{and} \quad 0 < \eta < \frac{1}{(N-1)a}$$

If

$$S(\mathbf{v}^0) < C - f\big(S(\mathbf{x}^0)\big)S(\mathbf{x}^0)$$

then $(\mathbf{x}, \mathbf{v})$ *exhibits asymptotic flocking in the sense of Definition 6.0.1.*

Proof. The smallness on the mesh η is imposed to make the second part of (6.15) a well posed consensus algorithm. Then, under the Assumption 3.1.1 of Theorem 3.1.5, the corresponding graph is $\mathbb{G}_{P(\mathbf{x}(k))}$ is scrambling for every k and hence it justifies the contraction bound

$$S\big(\mathbf{v}(k)\big) \leq \big(1 - \eta f\big(S\big(\mathbf{x}(k)\big)\big)\big)S\big(\mathbf{v}(k)\big)$$

Proving that $\sup_t S\big(\mathbf{x}(t)\big) < \infty$ implies that the flock will always remain sufficiently connected so that a_{ij} will be uniformly lower bounded and a direct application of Theorem 3.1.5 suffices to prove velocity alignment. Indeed let k^* be the first time that $S\big(\mathbf{x}(k^*)\big) \geq S\big(\mathbf{x}(k^* - 1)\big)$. Then for the "functional"

$$V(t) = S\big(\mathbf{v}(t)\big) + f\big(S\big(\mathbf{x}(t)\big)\big)S\big(\mathbf{x}(t)\big)$$

we have

$$\begin{aligned}
&V(k^*) - V(k^* - 1) \leq \\
&\leq \big(1 - \eta f\big(S\big(\mathbf{x}(k^* - 1)\big)\big)\big)S\big(\mathbf{v}(k^* - 1)\big) - S\big(\mathbf{v}(k^* - 1)\big) + \\
&\qquad\qquad + f\big(S\big(\mathbf{x}(k^*)\big)\big)S\big(\mathbf{x}(k^*)\big) - f\big(S\big(\mathbf{x}(k^* - 1)\big)\big)S\big(\mathbf{x}(k^* - 1)\big) \\
&\leq -\eta f\big(S\big(\mathbf{x}(k^* - 1)\big)\big)S\big(\mathbf{v}(k^* - 1)\big) + \\
&\qquad\qquad + f\big(S\big(\mathbf{x}(k^* - 1)\big)\big)\big[S\big(\mathbf{x}(k^*)\big) - S\big(\mathbf{x}(k^* - 1)\big)\big] \\
&\leq -\eta f\big(S\big(\mathbf{x}(k^* - 1)\big)\big)S\big(\mathbf{v}(k^* - 1)\big) + \eta f\big(S\big(\mathbf{x}(k^* - 1)S\big(\mathbf{v}(k^* - 1)\big) = 0
\end{aligned}$$

but $V(k^*-1) \leq S(\mathbf{v}^0) + \eta f\big(S(\mathbf{x}^0)\big)S(\mathbf{x}^0) = V(0)$. Consequently, $V(k^*) \leq V(0)$ or equivalently

$$0 \leq S\big(\mathbf{v}(k^*)\big) \leq S(\mathbf{v}^0) + f\big(S(\mathbf{x}^0)\big)S(\mathbf{x}^0) - f\Big(S\big(\mathbf{x}(k^*)\big)\Big)S\big(\mathbf{x}(k^*)\big)$$

choosing $C' < C$ small enough so that $S(\mathbf{v}^0) = C' - S(\mathbf{x}^0)$ and this implies

$$0 \leq C' - f\Big(S\big(\mathbf{x}(k^*)\big)\Big)S\big(\mathbf{x}(k^*)\big)$$

which in turn implies that $S\big(\mathbf{x}(k^*)\big)$ is bounded. □

Remark 6.1.5. If $\lim_x xf(x) = \infty$ then we have asymptotic flocking without any condition on the initial data just like the continuous time case.

The above simple proof provides a generalization of the discrete models known in the literature [52, 88] in the sense that a_{ij} are not assumed to be either symmetric or to have any particular expression. The above result can be generalized to the case of simple and/or switching schemes, (i.e. Theorem 3.1.8). We understand that the smallness condition on η is clearly technical, hence undesirable. This partially explains our persistent preference to the continuous time setting throughout the thesis.

6.1.2 Nonlinear averaging

An interesting variation of (6.1) can come from the results obtained in Chapter 5. In the algorithm to be studied here the speed coordination is achieved through

$$i \in [N] : \begin{cases} \dot{x}_i = u_i \\ \dot{u}_i = \sum_{j \in N_i} a_{ij}(t, \mathbf{x})\big(g_{ij}(u_j) - g_{ij}(u_i)\big), \ t \geq t_0 \\ x_i(t_0) = x_i^0, \ u_i(t_0) = u_i^0, \ \text{given} \end{cases} \tag{6.17}$$

This is a form of (5.5) with $\lambda(t) \equiv t$ and it will satisfy the monotonic conditions of §(5.2.2). For the sake of simplicity, here we will assume here full and static connectivity. More specifically,

Assumption 6.1.6. *The functions* $a_{ij}(t, \mathbf{y}) \in C^0([t_0, \infty) \times \mathbb{R}^N, [0, a])$ *for some* $a < \infty$ *whenver* $i \neq j$ *such that*

$$a_{ij}(t, \mathbf{y}) \geq f\big(S(\mathbf{y})\big)$$

for some integrable non-increasing $f(\cdot)$ *with the property that* $\lim_{z\to\infty} f(z) = 0$.

Assumption 6.1.7. *For fixed* $W \subset \mathbb{R}$ *there exist numbers* $0 < \underline{c} \leq \bar{c}$ *(possibly depending on* W*) such that*

$$\underline{c} \leq \frac{g_{ij}(x) - g_{ij}(y)}{x - y} \leq \bar{c}, \qquad \forall x, y \in W.$$

For the purposes of this section $\kappa := \min_{h,h' \in [N]} \sum_j \min\{p_{hj}^{(1)}, p_{h'j}^{(2)}\}$. Under Assumption 6.1.7, Lemma 5.2.5 directly applies. So for any initial data, $W = [\min_i u_i^0, \max_i u_i^0]$ is an appropriate compact set that characterizes $\underline{c}$ and $\bar{c}$. We recall Theorem 2.3.3 on which the vital contraction rates will be based.

Theorem 6.1.8. *Let the Assumptions 6.1.6 and 6.1.7 hold. The solution* $(\mathbf{x}, \mathbf{u})$ *of* (6.17) *exhibits asymptotic flocking if the initial conditions satisfy:*

$$S(\mathbf{u}^0) < N\underline{c}(\mathbf{u}^0) \int_{S(\mathbf{x}^0)}^{\infty} f(s)\, ds$$

for $\underline{c}(\mathbf{u}^0)$ *in the sense of Assumption 6.1.7.*

Proof. At first, the initial velocity $\mathbf{u}^0$ defines the constants $\underline{c}$ and $\overline{c}$ from Assumption 6.1.7 and Lemma 5.2.5. Next, it can be easily shown that throughout the solution $(\mathbf{x}, \mathbf{u})$

$$\sum_j \underline{b}_{ij}(t)\big(u_j(t) - u_i(t)\big) \le \dot{u}_i(t) \le \sum_j \overline{b}_{ij}(t)\big(u_j(t) - u_i(t)\big)$$

where:

1. $\underline{b}_{ij}(t) = \begin{cases} \underline{c}a_{ij}(t, \mathbf{x}(t)), \text{ if } & u_j(t) \ge u_i(t) \\ \overline{c}a_{ij}(t, \mathbf{x}(t)), \text{ if } & u_j(t) < u_i(t) \end{cases}$

2. $\overline{b}_{ij}(t) = \begin{cases} \overline{c}a_{ij}(t, \mathbf{x}(t)), \text{ if } & u_j(t) \ge u_i(t) \\ \underline{c}a_{ij}(t, \mathbf{x}(t)), \text{ if } & u_j(t) < u_i(t) \end{cases}$

for all $i \in [N]$. Next we pick $m > (N - 1 + \overline{c})a$ such that

$$\overline{P}(t)\mathbf{u}(t) \le e^{-mt}\frac{d}{dt}\big(e^{mt}\mathbf{u}(t)\big) \le \overline{P}(t)\mathbf{u}(t)$$

where $\overline{P}(t) = [\overline{p}_{ij}(t)]$ with $p_{ij}(t) = \overline{b}_{ij}(t)$ and $p_{ii}(t) = \big(m - \sum_j \overline{b}_{ij}(t)\big)$ and similar for $\underline{P}(t)$. We are interested in obtaining an upper bound for $\frac{d}{dt}S(\mathbf{u})$:

$$\begin{aligned}\frac{d}{dt}S(\mathbf{u}) &= \frac{d}{dt}\Big(e^{-mt}S\big(e^{mt}\mathbf{u}\big)\Big)\\ &= -mS(\mathbf{u}) + e^{-mt}\frac{d}{dt}S(e^{mt}\mathbf{u})\\ &\le -mS(\mathbf{u}) + S\bigg(e^{-mt}\frac{d}{dt}(e^{mt}\mathbf{u})\bigg)\\ &\le -mS(\mathbf{u}) + \big(m-\kappa(t)\big)S(\mathbf{u}) = -\kappa(t)S(\mathbf{u})\end{aligned}$$

in view of Theorem 2.3.3. From the choice of m a direct calculation of $\kappa(t)$ yields the lower bound $\kappa(t) > N\underline{c}f\big(S(\mathbf{x}(t))\big)$ so that

$$\frac{d}{dt}S\big(\mathbf{u}(t)\big) \le -N\underline{c}f\big(S(\mathbf{x}(t))\big)S\big(\mathbf{u}(t)\big) \tag{6.18}$$

If $\sup_{t\ge t_0} S\big((t)\big) < S(\mathbf{x}^0)$ then from (6.18) we have

$$\frac{d}{dt}S\big(\mathbf{u}(t)\big) \le -N\underline{c}f\big(S(\mathbf{x}^0)\big)S\big(\mathbf{u}(t)\big) \tag{6.19}$$

i.e. exponential fast speed alignment and therefore flocking. Otherwise, consider the functional

$$W(\mathbf{x},\mathbf{u}) = S(\mathbf{u}) + N\underline{c}\int_0^{S(\mathbf{x})} f(s)\,ds$$

the time-derivative of which along the solution $(\mathbf{x},\mathbf{u})$ yields $\frac{d}{dt}W \le 0$ in view of (6.18). Then from the imposed condition on the initial data there is r^* such that

$$S(\mathbf{u}^0) = N\underline{c}\int_{S(\mathbf{x}^0)}^{r^*} f(s)\,ds \tag{6.20}$$

and since $W(t) \le W(t_0)$ for $t \ge t_0$ we have

$$S\big(\mathbf{u}(t)\big) + N\underline{c}\int_0^{S(\mathbf{x}(t))} f(s)\,ds \le S(\mathbf{u}^0) + N\underline{c}\int_0^{S(\mathbf{x}^0)} f(s)\,ds$$

and substituting from (6.20) we have that

$$0 \le S\big(\mathbf{u}(t)\big) = N\underline{c}\Big(\int_{S(\mathbf{x}^0)}^{r^*} f(s)\,ds + \int_0^{S(\mathbf{x}^0)} f(s)\,ds - \int_0^{S(\mathbf{x}(t))} f(s)\,ds\Big)$$

from which it is deduced that $\int_{S(\mathbf{x}(t))}^{r^*} f(s)\,ds > 0$ and this implies that $S(\mathbf{x}(t)) \le r^*$ a valuable upper bound for the range of the flock so that again from (6.18)

$$\frac{d}{dt}S\big(\mathbf{u}(t)\big) \le -N\underline{c}f(r^*)S\big(\mathbf{u}(t)\big)$$

and that exponential speed alignment implies

$$S(\mathbf{x}(t)) \le S(\mathbf{x}^0) + \frac{S(\mathbf{u}^0)}{N\underline{c}f(r^*)} < \infty,$$

concluding the proof. □

6.2 Delayed Flocking

In this section we apply the results of Chapter 4 and especially the framework developed in §4.4. For this we recall the notation used there. The delayed version of (6.1) to be discussed in this section involves time-varying propagation delays only

$$i \in [N] : \begin{cases} \dot{x}_i(t) = u_i(t) \\ \dot{u}_i(t) = \sum_{j \in N_i} a_{ij}\big(\mathbf{x}(t)\big)\big(u_j(\lambda_{ij}(t)) - u_i(t)\big), & t \ge t_0 \\ x_i(t_0 - \tau(t_0)) = x_i^0, u_i(t) = \phi_i(t), & t \in I_{t_0} \end{cases} \tag{6.21}$$

where $\mathbf{x}^0$ and $\mathbf{u} = \boldsymbol{\phi}$ are sufficient given initial data for the problem to be well-posed as

$$\mathbf{x}(t) = \mathbf{x}^0 + \int_{\lambda(t_0)}^{t} \mathbf{u}(s)\,ds, \quad t \ge \lambda(t_0)$$

The working hypothesis with this setup is that the bird i receives a delayed version of the speed of its neighboring birds but not of the position. This is partially because it is exactly this dynamic variable that it can suffer from delays. We will shortly describe how similar results can be obtain if one imposes delays on the position of the neighbors as well.

The analysis relies on the rate at which $S_{I_t}(\mathbf{u})$ contracts. This is an estimate already obtain in the proof of Theorem 4.4.6.

Theorem 6.2.1. *Consider the solution* $(\mathbf{x}, \mathbf{u})$ *of* (6.21)*. Assume TYPE II & Static network connectivity over the network so that* $a_{ij}(\cdot, \cdot)$ *satisfy conditions* (6.6), (6.7) *and* (6.8)*. If* $\tau = \sup_{t \geq t_0} \max_{i,j} \tau_{ij}(t) < \infty$*, then asymptotic flocking occurs according to Def. 6.0.1, if for some* $B > 0$*, the initial data satisfy*

$$S_{I_{t_0}}(\boldsymbol{\phi}) < \frac{(1 - e^{-mB})^\gamma}{m^\gamma (2\tau + B)} e^{-\bar{N} a \tau} \int_{P^{2\tau+B}_{\mathbf{x}^0, \boldsymbol{\phi}}}^{\infty} f^\gamma(s)\, ds \tag{6.22}$$

where $P^{2\tau+B}_{\mathbf{x}^0, \boldsymbol{\phi}} = \max\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - S_{I_{t_0}}(\boldsymbol{\phi})(2\tau + B)|\}$ *and* $\bar{N} = \max_i |N_i|$*.*

Proof. The contraction estimate from Theorem 4.4.6 is

$$S_{I_t}(\mathbf{x}) \leq \left(1 - \frac{(1 - e^{-mB})^\gamma}{m^\gamma} f^\gamma(S(\mathbf{x}(t))) e^{-N a \tau}\right) S_{I_{t-\gamma B - 2\tau}}(\mathbf{x}) \tag{6.23}$$

At first we note that

$$\frac{d}{dt} S\big(\mathbf{x}(t)\big) \leq S\big(\mathbf{u}(t)\big) \leq S_{I_t}(\mathbf{u}) \leq S_{I_{t'}}(\mathbf{u})$$

for any $t' \geq t$ in view of Lemma 4.4.4 that directly applies. We introduce the functional

$$V(\mathbf{x}, \mathbf{u}) = \int_{t-2\tau-B}^{t} S_{I_s}(\mathbf{u})\, ds + C \int_0^{S(\mathbf{x})} f^\gamma(s)\, ds. \tag{6.24}$$

for $C = \frac{(1-e^{-mB})^\gamma}{m^\gamma} e^{-\bar{N}a\tau}$. We then evaluate it along $\mathbf{x}(t), \mathbf{u}(t)$ and take the time derivative to obtain

$$\frac{d}{dt}V(t) \leq 0 \Rightarrow V(t) \leq V(t_0)$$

or equivalently

$$\begin{aligned} \int_{t-2\tau-B}^{t} S_{I_s}(\mathbf{u})\, ds + C \int_0^{S(\mathbf{x}(t))} f^\gamma(s)\, ds \leq \\ \leq \int_0^{2\tau+B} S_{I_s}(\mathbf{u})\, ds + C \int_0^{S(\mathbf{x}(2\tau+B))} f^\gamma(s)\, ds \end{aligned} \tag{6.25}$$

Assume the condition:

$$\int_0^{2\tau+B} S_{I_s}(\mathbf{u})\, ds < C \int_{S(\mathbf{x}(2\tau+B))}^{\infty} f^\gamma(s)\, ds \tag{6.26}$$

Then we can pick r' such that

$$\int_0^{2\tau+B} S_{I_s}(\mathbf{u})\, ds = C \int_{S(\mathbf{x}(2\tau+B))}^{r'} f^\gamma(s)\, ds \tag{6.27}$$

Substituting (6.27) into (6.25) we get

$$\int_0^{S(\mathbf{x}(t))} f(s)\, ds \leq \int_{S(\mathbf{x}(2\tau+B))}^{r'} f^\gamma(s)\, ds + \int_0^{S(\mathbf{x}(2\tau+B))} f^\gamma(s)\, ds$$

so that if $S(\mathbf{x}(t)) \geq S(\mathbf{x}(2\tau + B))$ then necessarily $S(\mathbf{x}(t)) \leq r'$ which implies

$$\sup_t S(\mathbf{x}(t)) < \infty$$

throughout the solution and the exponentially fast alignment of the flock velocity is achieved. It is only left to show that the imposed (6.22) implies (6.26).

On the one hand ,the left part of the inequality (6.26) is upper bounded by $S_{I_{t_0}}(\boldsymbol{\phi})(2\tau + B)$. On the other hand, unless $S(\mathbf{x}(t)) \leq S(\mathbf{x}^0)$, the rate at which

$S(\mathbf{x}(t))$ may shrink can be deduced from the extreme scenario $\mathbf{x}^0 = (x^0, 0, \ldots, 0)$ with $x^0 < 0$ so that $S(\mathbf{x}^0) = x^0$ and $\boldsymbol{\phi}(s) = (\phi_1(s), 0, \ldots, 0)$, $s \in [-\tau, 0]$. Neglecting the averaging effect which will inevitably diminish $|S_{I_t}(\mathbf{u})|$, $x^0 < 0$ implies that the first agent will have approached or bypassed the rest of the group by $-|x^0| + |u^0|t$. Finally, at $t = 2\tau + B$ the spread of $\mathbf{x}(2\tau + B)$ is lower bounded by

$$S(\mathbf{x}(2\tau + B)) \geq \max\{S(\mathbf{x}^0), |S(\mathbf{x}^0) - |W_{t_0}^{\boldsymbol{\phi}}|(2\tau + B)|\}$$

and the proof is concluded. $\square$

A direct corollary occurs when f^γ is not summable. Indeed, based on the type of connectivity we imposed in Theorem 6.2.1, unconditional, delay-independent asymptotic flocking in the sense of Definition 6.0.1 occurs if

$$\int^\infty f^\gamma(s)\, ds = \infty.$$

6.3 Noisy Flocking

A standard application where a system of differential equations of Itô type occurs is a stochastic perturbation of a deterministic nominal system. In this section we will study two flocking models, both of which are simplifications of a general non-linear system of stochastic differential equations. For the rest of the section we recall the discussion in §2.6.

Consider the set $[N]$ of autonomous agents and fix $t_0 \in \mathbb{R}$, $T \geq t_0$. The two vector valued stochastic processes

$$\mathbf{X}_t = (X_t^{(1)}, \ldots, X_t^{(N)}),\ \mathbf{U}_t = \left(U_t^{(1)}, \ldots, U_t^{(N)}\right)$$

stand for the positions and the velocities of the members of the flock. In fact, $(\mathbf{X}_t, \mathbf{U}_t)$ are considered to be the solution of the system of Itô stochastic differential equations

$$i \in [N] : \begin{cases} dX_t^{(i)} = U_t^{(i)} dt \\ dU_t^{(i)} = \sum_j a_{ij}(t, \mathbf{X}_t)\big(U_t^{(j)} - U_t^{(i)}\big)dt + g_{ij}(t, \mathbf{X}_t, \mathbf{U}_t)dB_t^{(ij)} \end{cases} \tag{6.28}$$

for $t \in [t_0, T]$, subject to initial data

$$i \in [N] : X_{t_0}^{(i)} = X_i^0, U_{t_0}^{(i)} = U_i^0.$$

Equivalently, they are the solution of

$$\begin{aligned} d\mathbf{X}_t &= \mathbf{U}_t dt \\ d\mathbf{U}_t &= -L(t, \mathbf{X}_t)\mathbf{U}_t dt + \sum_{i=1}^{N} G_i(t, \mathbf{X}_t, \mathbf{U}_t) d\mathbf{B}_t^{(i)} \end{aligned} \tag{6.29}$$

for $t \in [t_0, T]$, subject to initial data

$$\mathbf{X}_{t_0} = \mathbf{X}^0, \mathbf{U}_{t_0} = \mathbf{U}^0.$$

Provided $\mathbf{X}_t, \mathbf{U}_t$ are $\mathcal{U}_t$-measurable and $L(\mathbf{X}, t)\mathbf{U}_t \in L_N^1(0, T)$, $G \in L_{N\times N}^2(0, T)$ the processes $\mathbf{X}_t, \mathbf{U}_t$ satisfy

$$\begin{cases} \mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^t \mathbf{U}_s\, ds \\ \mathbf{U}_t = \mathbf{U}^0 - \int_{t_0}^t L\big(s, \mathbf{X}_s\big)\mathbf{U}_s\, ds + \sum_{i=1}^N \int_{t_0}^t G_i(s, \mathbf{X}_s, \mathbf{U}_s) d\mathbf{B}_s^{(i)} \end{cases} \quad a.s. \tag{6.30}$$

Now

$$\mathbf{B}^{(i)}(\cdot) = \big(B^{(i1)}(\cdot), B^{(i2)}(\cdot), \ldots, B^{(iN)}(\cdot)\big) \tag{6.31}$$

is an N-dimensional Brownian motion and $\mathbf{X}^0, \mathbf{U}^0$ are two N-dimensional random variables independent of $\mathbf{B}(\cdot)$. Since we analyze the asymptotic behavior of solutions, we are essentially interested in the collection $\{(\mathbf{X}_t, \mathbf{U}_t)\}_{t\geq t_0}$ as solution of the above system of SDE's. We adapt Definition 6.0.1 in the present stochastic setting:

Definition 6.3.1. The system (6.28) exhibits *asymptotic strong stochastic* flocking if and only if the position-velocity processes $X_t^{(i)}, U_t^{(i)}$, $i \in [N]$ satisfy the conditions

$$\lim_{t\to\infty} \big|U_t^{(i)} - U_t^{(j)}\big| = 0, \quad a.s. \qquad \text{and} \qquad \sup_{t\geq t_0} \big|X_t^{(i)} - X_t^{(j)}\big| < \infty, \quad a.s.$$

Furthemore, (6.29) exhibits *asymptotic strong stochastic* flocking in the mean square sense if the aforementioned processes converge accordingly. Our aim is to prove convergence results for the following two simplified stochastic networks:

1. Time invariant flocking with state-independent multiple diffusions:

$$i \in [N] \begin{cases} dX_t^{(i)} = U_t^{(i)} dt \\ dU_t^{(i)} = \sum_j a_{ij}\big(U_t^{(j)} - U_t^{(i)}\big)dt + g_{ij}(t)dB_t^{(j)} \end{cases} \tag{6.32}$$

2. Time varying flocking model with state-dependent stochastic disturbance and uniform time-varying diffusion coefficient:

$$i \in [N] : \begin{cases} dX_t^{(i)} = U_t^{(i)} dt \\ dU_t^{(i)} = \sum_j a_{ij}(t)\big(U_t^{(j)} - U_t^{(i)}\big)dt + g(t)\big(U_t^{(j)} - U_t^{(i)}\big)dB_t^{(j)} \end{cases} \tag{6.33}$$

6.3.1 LTI flocking with noise

We begin with the study of (6.32) subject to initial data $\mathbf{X}^0, \mathbf{U}^0$. In the absence of noise, (6.32) reduces to (2.3).

Assumption 6.3.2. *The associated graph $\mathbb{G}_A$ of the matrix $A = [a_{ij}]$, is routed-out branching.*

We recall the discussion in §2.2. From Assumption 6.3.2 we there exists a unique normalized left eigenvector of the Laplacian matrix with respect to the zero eigenvalue, $\mathbf{c} \in \mathbb{R}^N$. The solution of (2.3) with initial data $\mathbf{U}^0$ is $e^{-Lt}\mathbf{U}^0$ and it satisfies

$$|e^{-Lt}\mathbf{u}^0 - \mathbb{1}\mathbf{c}^T\mathbf{u}^0| \leq Ke^{-\Re\{\lambda\}t}$$

for some $K > 0$ that depends both on the norm $|\cdot|$, the parameters a_{ij} and $\Re\{\lambda\} > 0$, i.e. the second smallest real part of the eigenvalues of L.

Proposition 6.3.3. *The solution* $\big(\mathbf{X}_t, \mathbf{U}_t\big)$ *of* (6.32) *satisfies*

$$\mathbf{X}_t = \mathbf{X}^0 + \int_{t_0}^{t} \mathbf{U}_s \, ds$$

$$\mathbf{U}_t = e^{-L(t-t_0)}\mathbf{U}^0 + \int_{t_0}^{t} e^{-L(t-s)}G(s)d\mathbf{B}_s$$

for $t \in [t_0, T]$.

Proof. The form of $\mathbf{X}_t$ is the definition of the process so we will only prove the expression of $\mathbf{U}_t$. Define the process

$$\mathbf{V}_t := \mathbf{U}^0 + \int_{t_0}^{t} e^{L(s-t_0)}G(s)d\mathbf{B}_s$$

the differential of which is $d\mathbf{V}_t = e^{L(t-t_0)}G(t)d\mathbf{B}_t$. We will use Itô's product rule to calculate the differential of $e^{-L(t-t_0)}\mathbf{V}_t$ which is identical to $\mathbf{U}_t$:

$$d\big(e^{-L(t-t_0)}\mathbf{V}_t\big) = G(t)d\mathbf{B}_t - Le^{-L(t-t_0)}\mathbf{V}_t dt = -L\mathbf{U}_t dt + G(t)d\mathbf{B}_t$$

and the result follows. □

We see that in this simple case, the solution $\mathbf{U}_t$ is expressed in closed form. Asymptotic stochastic flocking is determined by the behavior of the local martingales $\int_{t_0}^{t} g_{ij}(s)dB_s^j$ as $t \to \infty$.

Theorem 6.3.4. *Let Assumption 6.3.2 hold. If* $\mathbb{E}[(\mathbf{U}^0)^2], \mathbb{E}[(\mathbf{X}^0)^2] < \infty$ *and for any* $i, j \in [N]$, $g_{ij}(\cdot)$ *satisfy*

$$\lim_{t\to\infty} \int_{t_0}^{t} g_{ij}^2(s)\, ds < \infty \qquad \text{and} \qquad \int_{t}^{\infty} g_{ij}^2(s)\, ds \in L^1_{[t_0,\infty]}$$

then asymptotic stochastic flocking occurs in the sense of Definition 6.3.1.

In particular, the agents align their speed around the $\mathcal{U}_\infty$*-measurable random variable*

$$k := \mathbf{c}^T \mathbf{U}^0 + \sum_{i,j} \int_{t_0}^{\infty} c_i g_{ij}(s) dB_s^{(j)}.$$

and they exhibit asymptotic stochastic flocking in the almost sure and in the mean square sense.

Proof. At first, we clarify that k is well-defined since $\int_{t_0}^{\infty} g_{ij}(s) dB_s^{(j)}$ is almost surely finite exactly because the first imposed condition on g_{ij} yields almost sure finiteness by the Martingale Convergene Theorem [33]. Next,

$$\mathbf{U}_t - \mathbb{1}k = \left(e^{-L(t-t_0)} - \mathbb{1}\mathbf{c}^T\right)\mathbf{U}^0 + \int_{t_0}^{t} \left(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T\right) G(s) d\mathbf{B}_s + \mathbb{1}\mathbf{c}^T \int_{t}^{\infty} G(s) d\mathbf{B}_s$$

From the Properties of Itô's integral and the Cauchy-Schwarz inequality we obtain the following bound:

$$\begin{aligned}
&\mathbb{E}\left[||\mathbf{U}_t - \mathbb{1}k||_2^2\right] \leq \\
&\leq K^2 e^{-2\Re\{\lambda\}(t-t_0)} \mathbb{E}[||\mathbf{U}^0||_2^2] + \\
&+ \mathbb{E}\left[\left(\int_{t_0}^{t} \left(e^{-L(t-s)} - \mathbb{1}\mathbf{c}^T\right) G(s) d\mathbf{B}_s\right)^2\right] + \mathbb{E}\left[\left(\int_{t}^{\infty} \mathbb{1}\mathbf{c}^T G(s) d\mathbf{B}_s\right)^2\right] \qquad (6.34)\\
&\leq K^2 e^{-2\Re\{\lambda\}(t-t_0)} \mathbb{E}[||\mathbf{U}^0||_2^2] + \\
&+ \sum_{i,j} \int_{t_0}^{t} K^2 e^{-2\Re\{\lambda\}(t-s)} g_{ij}^2(s)\, ds + \sum_{i,j} \int_{t}^{\infty} c_i^2 g_{ij}^2(s)\, ds
\end{aligned}$$

By assumption $g_{ij}^2(\cdot)$ vanishes. Next, $\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]$ is bounded from above by three terms, each of which converges to zero as $t \to \infty$: the first, after Assumption 6.3.2, the third by the imposed condition on $g_{ij}(s)$'s and the second as a convolution of an L^1 function with a function that goes to zero.

Then the random variable $\mathbf{U}_t$ converges asymptotically to Δ in the mean square sense. To prove almost sure speed coordination we first see that from the Chebyshev inequality for any, $\varepsilon > 0$

$$\mathbb{P}\big(|U_t^{(i)} - U_t^{(j)}| \geq \varepsilon\big) \leq \frac{1}{\varepsilon^2}\mathbb{E}\big[|U_t^{(i)} - U_t^{(j)}|^2\big] \leq \frac{1}{\varepsilon^2}\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]$$

it is an easy exercise to show that all of the terms that bound $\mathbb{E}\big[||\mathbf{U}_t - \mathbb{1}k||_2^2\big]$ from above in (6.34) are integrable over $[t_0, \infty)$ (the second term can be proved by a simple change in the order of integration). Then because $\mathbb{P}\big(|U_t^{(i)} - U_t^{(j)}|\big)$ is summable, almost sure convergence to $\mathbb{1}k \in \Delta$ follows (see Theorem 4(c) of §7.2 in [32]).

Finally,

$$|X_t^{(i)} - X_t^{(j)}| \leq |X_{t_0}^{(i)} - X_{t_0}^{(j)}| + \int_{t_0}^{t} |U_s^{(i)} - U_s^{(j)}|\, ds < \infty \ a.s.$$

and hence $X_t^{(i)} - X_t^{(j)}$ is bounded in probability, therefore it is bounded in the 2^{nd}-mean (see Theorem 4(b) of §7.2 in [32]). □

It is noted that since g_{ij} are deterministic functions, k is a normally distributed random variable with mean $\sum_i c_i \mathbb{E}[U_i^0]$ and variance $\sum_{i,j} c_i^2 \int_{t_0}^{\infty} g_{ij}^2(s)\, ds$.

The results of this section can be trivially generalized to the case of time-varying connectivity weights $a_{ij}(t)$. In this case the kernel $\Phi(t, t_0)$ behaves similarly in the case of Sec. 3.2, whatever the connectivity regime may be.

6.3.2 LTV flocking with noise

Algebraic methods do not apply in general linear systems whereas stability in variation can effectively work in the case of state-independent noise as it was analyzed above. When a version with state-dependent noisy compartments is considered, one would not want to disregard its stabilizing contribution. Given (6.33) subject to the initial data $\mathbf{X}^0, \mathbf{U}^0$ we will derive expressions based on the coefficient of ergodicity.

Assumption 6.3.5. *The functions $a_{ij}(\cdot)$ introduced in* (6.33) *are continuous and uniformly bounded functions of time.*

Assumption 6.3.6. *The adjacency matrix of A_g of the diffusion compartment, corresponds to a complete graph.*

The necessity of Assumption 6.3.6 stems from the fact that the white noise $d\mathbf{B}_t$ as an integrator obeys no rules of monotonicity with respect to its integrand. For simplicity we introduce the notation $S(\mathbf{U}_t^2) = \max_{i,j}(U_t^{(i)} - U_t^{(j)})^2$

Proposition 6.3.7. *Under Assumptions 6.3.5 and 6.3.6, the solution $(\mathbf{X}_t, \mathbf{U}_t)$ of* (6.33) *satisfies*

$$d\big(S(\mathbf{U}_t^2)\big) \leq 2\bigg(g^2\frac{N}{2} - f(t)\bigg)S(\mathbf{U}_t^2)dt - g(t)S(\mathbf{U}_t^2)\sum_l dB_t^{(l)}$$

where $f(t) = \min_{i,j}\sum_l \min\{a_{il}(t), a_{jl}(t)\}$.

Proof. We fix $t \geq t_0$ and we will always consider the elements $i = i_t, j = j_t \in [N]$ that maximize $U_t^{(ij)} := \big(U_t^{(i)} - U_t^{(j)}\big)$. Firstly, we need an expression of the differential $d\big(e^{2mt}(U_t^{(ij)})^2\big)$, For this we compute the differentials $dU_t^{(ij)}$ and $d\big(U_t^{(ij)}\big)^2$ using Itô calculus:

$$dU_t^{(ij)} = \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)dt + g(t)\sum_l U^{(li)}dB_t^{(l)} - g(t)\sum_l U_t^{(lj)}dB_t^{(l)} =$$

$$= \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)dt - g(t)U^{(ij)}\sum_l dB_t^{(l)}$$

$$d\big(U_t^{(ij)}\big)^2 = 2U_t^{(ij)}\big[\sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big) + Ng^2U_t^{(ij)}\big]dt - 2g\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)}$$

Eventually,

$$\begin{aligned}
&d\big(e^{2mt}(U_t^{(ij)})^2\big) = \\
&= 2me^{2mt}(U_t^{(ij)})^2dt + e^{2mt}d\big(U_t^{(ij)}\big)^2 \\
&= 2e^{2mt}U_t^{(ij)}\left[\left(m + \frac{Ng^2}{2}\right)U_t^{(ij)} + \sum_l \big(a_{il}U_t^{(li)} - a_{jl}U_t^{(lj)}\big)\right]dt - \\
&\qquad\qquad - 2ge^{2mt}\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)} \\
&= 2e^{2mt}U_t^{(ij)}\sum_l \big(a_{il} - a_{jl}\big)U_t^l dt - 2ge^{2mt}\big(U_t^{(ij)}\big)^2\sum_l dB_t^{(l)}
\end{aligned}$$

where $a_{ii} = m + \frac{Ng^2}{2} - \sum_l a_{il}$ which is positive for m large enough. At this point we shall focus on $\sum_l(a_{il} - a_{jl})U_t^{(l)}$ for which we notice that $a_{ij} > 0$ and $Q = \sum_l(a_{il} - a_{jl}) = 0$ for all $i, j \in [N]$. Then, if we let $w_l := a_{il} - a_{jl}$, we note that

$$\theta = \sum_{l:w_l>0} w_l = -\sum_{l:w_l<0} w_l$$

so

$$
\begin{aligned}
Q &= \sum_{l:w_l>0} w_l U_t^{(l)} + \sum_{l:w_l<0} w_l U_t^{(l)} \\
&= \sum_{l:w_l>0} w_l U_t^{(l)} - \sum_{l:w_l<0} |w_l| U_t^{(l)} \\
&= \theta\Big(\frac{\sum_{l:w_l>0} w_l U_t^{(l)}}{\theta} - \frac{\sum_{l:w_l<0} |w_l| U_t^{(l)}}{\theta} \Big) \\
&\le \theta\big(\max_l U_t^{(l)} - \min_l U_t^{(l)} \big) \\
&= \theta U_t^{(ij)}
\end{aligned}
$$

then since $\theta \le \frac{1}{2}\max_{i,j}\sum_l |a_{il}-a_{jl}| = m + g^2(t)\frac{N}{2} - \min_{i,j}\sum_l \min\{a_{il}, a_{jl}\}$ we obtain the bound for $f(t) = \min_{i,j}\sum_l \min\{a_{il}(t), a_{jl}(t)\}$

$$
d\big(e^{2mt}(U_t^{(ij)})^2\big) \le 2e^{2mt}\Big(m + \frac{Ng^2(t)}{2} - f(t)\Big)\big(U_t^{(ij)}\big)^2 dt - 2g(t)e^{2mt}\big(U_t^{(ij)}\big)^2 \sum_l dB_t^{(l)}
$$

and finally

$$
\begin{aligned}
d(U_t^{(ij)})^2 &= d\big(e^{-2mt}e^{2mt}(U_t^{(ij)})^2\big) \\
&= e^{-2mt} d\big(e^{2mt}(U_t^{(ij)})^2\big) - 2m(U_t^{(ij)})^2 dt \\
&\le \big(Ng^2(t) - 2f(t)\big)\big(U_t^{(ij)}\big)^2 dt - 2g(t)\big(U_t^{(ij)}\big)^2 \sum_l dB_t^{(l)}
\end{aligned}
$$

as i, j are chosen to be the maximizers of $\big(U_t^{(i)} - U_t^{(j)}\big)^2$ the proof is concluded. $\square$

Now we are ready to prove the flocking result :

Theorem 6.3.8. *Under Assumptions 6.3.5 and 6.3.6, asymptotic stochastic flocking for the system* (6.33) *occurs if*

$$
e^{\int_{t_0}^t Ng^2(s) - 2f(s)\, ds} \in L^1_{[t_0,\infty)}.
$$

Proof. From Proposition 6.3.7 the solution $(\mathbf{X}_t, \mathbf{U}_t)$ satisfies [31]

$$S(\mathbf{U}_t^2) \leq e^{\int_{t_0}^t \left(\frac{N}{2} g^2(s) - 2f(s)\right) ds + \sum_l \int_{t_0}^t g(s) dB_t^{(l)}} S(\mathbf{U}_{t_0}^2), \quad a.s.$$

so that the second moment is calculated from the properties of the martingales $\int_{t_0}^t g(s) dB_s$ as :

$$\mathbb{E}\big[S(\mathbf{U}_t^2)\big] \leq e^{\int_{t_0}^t N g^2(s) - 2f(s)\, ds} \mathbb{E}\big[S(\mathbf{U}_{t_0}^2)\big]$$

and the rest of the proof is identical to this of Theorem 6.3.4. □

We note here that if $g^2 \in L^1_{[t_0,\infty)}$ then the non-summability of f suffices to prove consensus and in addition $e^{\int^t f(s)\, ds} \in L^1_{[t_0,\infty)}$ suffices to prove flocking. Other wise, $\int_{t_0}^\infty g^2(s)\, ds = \infty$ implies that $|\int_{t_0}^t g(s) dB_s|$ behaves asymptotically as

$$\sqrt{2 \int_{t_0}^t g^2(s)\, ds \log\log \int_{t_0}^t g^2(s)\, ds}, \quad a.s.$$

by the iterated logarithm for martingales [32, 33]. With this in mind it is possible that the assumption of the integrability of g^2 can be relaxed.

The two results above are improvements of [93, 90] as we allow weights to be non-symmetric and the diffusion coefficient to be time-varying. In the particular case of time independent weights we also allow minimal connectivity as well as we are able to identify the consensus point of the velocities. Furthermore, one is free to assume non-linear state-dependent connections on condition that they are lower-bounded away from zero. Initial conditions as for example (6.9) that automatically imply a bounded distance among agents as in the deterministic case do, are yet to be established.

6.4 Supplementary Remarks

This chapter provides non-linear flocking convergence results based on the derived rate estimates of Chapters 3, 4 and 5.

For the ordinary models we provide two types of results: The first type concerns a model with convex and linear averaging and the other type a nonlinear monotonic averaging. For the former model we state sufficient conditions for flocking under simple and/or switching connectivity regime. The non-linear averaging, on the other hand, assumes increased connectivity but it can be extended to simple connectivity along the lines discussed in §5.4. The delayed case involves arbitrary time-varying propagation delays in the sense of the general theory for linear delayed networks in §4.4.

All these results involve the use of Lyapunov functionals that combine the time evolution of $S(\mathbf{u})$ and $S(\mathbf{x})$. Asymptotic flocking occurs if the corresponding metrics of the solutions satisfy particular inequality conditions. Based on the growth of solution (or special extreme scenarios) provide sufficient relations on the initial configuration which imply the solutions conditions, i.e. asymptotic flocking.

The case of noisy stochastic flocking is different. New results were provided for time invariant and time varying (state-independent) asymmetric couplings. The former is based on elementary algebraic analysis and its main interest revolves around the consensus point that becomes a random variable with certain distribution. The latter is derived with yet another variation of the proof of the contraction coefficient. Due to purely technical reasons, the coupling weights $a_{ij}(t)$ were considered as state-independent. The main drawback in the noisy flocking dynamics is that we could not drop the uniform lower bound condition, just like the former works [90, 93] could not either. We conjecture that stochastic flocking in the sense of the deterministic models, can only be provided with some probability on the event that some sufficient initial conditions are satisfied.

Chapter 7: An Application In Power Networks

Modern Energy Supply is typically a structure of interconnected power generation plants that independently produce power to serve a load over a common distribution network [98, 99, 16]. Since every power unit produces energy at some cost (see § 7.2), a fundamental power optimization problem, is the determination of the optimal combination of outputs of all generating units to minimize the total cost while satisfying the load demand and operational constraints. This is the very well-known Economic Dispatch Problem (EDP).

Over the past years, many optimization methods for the EDP have been proposed in the literature. The conventional ones include the lambda iteration algorithm, or gradient-based search methods [98, 16] ,whereas modern heuristic optimization techniques are based on operational research and artificial intelligence concepts such as evolutionary algorithms [100, 101, 102, 103, 104, 105], simulated annealing [106, 107] artificial neural networks [108, 109, 110], taboo search [111, 112] and particle swarm optimization techniques [113, 114, 115, 116].

Although the performance and applicability of economic dispatch has been improved by these optimization techniques, it is still essential to maintain a single control center that can access the state of the entire system. Indeed, all the aforementioned algorithms actually require operations at a central computing station that needs to have a priori knowledge of the entire network parameters. This centrally controlled framework may cause some performance limitations in the future power grid.

Since 1990, many electric utilities including both government- and private-owned electric utilities were liberated. This has had profound effects on the operation of electric systems where implemented, most of which is the economic value to the network operator. The EDP is a relevant procedure in the operation of a power system. The deregulation of the electric utilities has, therefore, led to research on a decentralized model of control where utilities, transmission system operators (TSO) and independent power producers (IPP) cooperate and compete using market and other mechanisms [16].

Smart-Grid architecture The next generation of power systems is expected to satisfy high standards of efficiency, resilience and reliability against cyber-attacks or natural disasters, improved integration of renewable energy resources and plug-in hybrid electrical vehicles. Such electrical grids, known as Smart-Grid, are designed to monitor, predict and intelligently respond to the behavior of all electric power suppliers and consumers connected to it in order to deliver such standards [117, 118].

Figure 7 is an abstract illustration of a Smart-Grid architecture. A necessary requirement towards this is the development of advanced control and communication technology both in the physical and the algorithmic layer. In a smart grid environment, the communication and measurement requires a multiagent systems

(MAS) technology [118]. MAS are a computational system in which several agents cooperate to achieve a desired task. The performance of a MAS can be decided by the interactions among various agents. Agents cooperate to achieve more than if they act individually. Increasingly, MAS are the preferential choice for developing distributed systems such as the Smart-Grid [118]. The development of monitoring and measurement in Smart Grid with the use of MAS technology involves a combination of several agents working without human intervention, in collaboration pursuing assigned tasks to achieve the overall goal of the system. The distributed

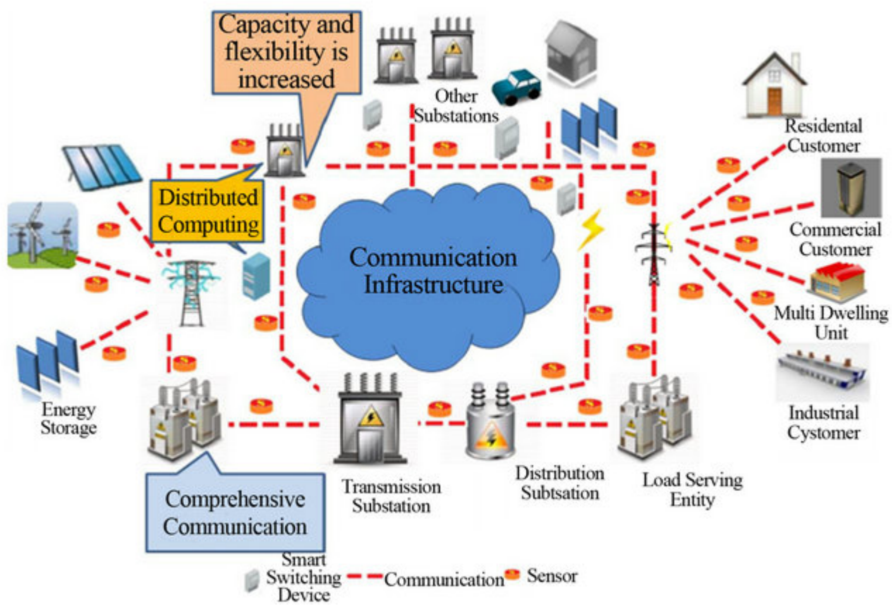

Figure 7.1: A Smart-Grid electric network with multiple communication and control sensors.

solution of the EDP for power systems has been recently introduced in the literature generators [119, 120, 121, 122]. We are particularly interested in [120, 121, 122] that bear the greatest resemblance with the present work. Zhang et al. propose the use of averaging consensus schemes for solving the EDP problem in an identical scenario of power sources, loads and sensors with and without the presence of communication delays.

Contribution The successful and reliable implementation of complex electrical networks such as the Smart-Grids, require advanced measurement and control methods that operate in a distributed way. Recent advances in consensus systems [8, 11, 14] provide results that could be effectively applied to solve the fundamental EDP of an electric network in such a manner. The present work introduces and studies by theory and simulation a distributed solution of the EDP using a variation of lambda iteration [16], under the presence of arbitrary signal propagation delays.

We consider a scenario of several power generators and loads connected to a common transmission network (see Figure 7.2). This is a grid enhanced with autonomous sensors, each of which controls at most one power generator and one power load. The sensors are connected to a common communication network and share certain information. The central characteristic of the communications is that they suffer from multiple time-varying delays. The network topology among the sensor is assumed time varying and complete, that is every agent communicates with each other even via arbitrary delayed signals. The objective is the solution of the EDP under the power generation constraints that each unit must meet and the delayed communication regime between the sensors. In particular every sensor will be designed to receive, process and transmit back information and will act both as a follower and as a leader in multiple computational levels so that the EDP is to be solved in a decentralized manner. The theoretical results will be supported with an illustrative example.

Our work differs on a number of points. At first, their model is primarily discrete and follows simplified average consensus schemes introduced in the early work of [45] both in ordinary and delay form. These systems are too symmetric, hence unrealistic, both in the communication rate and the imposed delays. In particular, the delayed case is treated in too much uniformity: each sensor receives the information from its neighbors under the same delay while it averages all the information with a delayed version of its own information. The working hypothesis is that the system dynamics evolve under both propagation and processing constant delays of identical magnitudes; a fairly unrealistic scenario. Moreover, the proposed algorithms solve the distributed EDP only in part. A leader sensor needs to be chosen so as to control the overall power mismatch and dynamically adjust the incremental cost value.

We propose a decentralized version of the lambda iteration algorithm. Each sensor controls a part of the electric grid, sends and receives information and executes multiple and simultaneous dynamic consensus iterations. This way it learns all the information needed to concur with the optimal operation point values that solve the EDP. Our scenario takes into account the presence and the effect of multiple delays, different for every iteration algorithm. Every sensor serves both as a leader and a follower in the network and in its utmost generality it needs to know the static parameters of the network, i.e. the connectivity weights and the delays each sensor operates under. However this is a knowledge on the communication level, no information on the transmission network is needed as the sensors, through the consensus scheme, learn the information (loads and generator powers) dynamically.

The rate at which the agents learn depends on the communication parameters and of course it plays a vital role in the stability of the system.

7.1 Delayed Consensus Re-stated

For the present section we will need the results of § 4.4 and particularly Theorems 4.4.6 and 4.4.7, the results of which, together with the model, we restate here for quick reference.

A set of N autonomous agents each of which possess a value of interest, say x_i for $i = 1, \ldots, N$, that shares and updates it dynamically so that $\mathbf{x}(t) = \big(x_1(t), \ldots, x_N(t)\big)$ satisfies

$$\begin{cases} \dot{x}_i(t) = \sum_{j \neq i} a_{ij}(t)\big(x_j(t - \tau_{ij}(t)) - x_i(t)\big) \\ x_i(t) = \phi_i(t), t \in I_{t_0} \end{cases} \tag{7.1}$$

where $I_{t_0} = [t_0 - \tau(t_0), t_0]$. The set $W_{I_{t_0}, \boldsymbol{\phi}}$ is

$$[\min_i \min_{s \in I_{t_0}} \phi_i(s) - \min_i \max_{s \in I_{t_0}} \phi_i(s)]$$

and its length is denoted by $S_{I_{t_0}}(\boldsymbol{\phi})$. For the communication network we assume that it is fully connected and static but with time varying weights, i.e. $\forall i \neq j$, $a_{ij}(t) \Rightarrow a_{ij}(t) \in [\underline{\alpha}, \bar{\alpha}]$. Then for every $B > 0$ and $t \geq t_0$,

$$\int_{t-B}^{t} A(s)\, ds$$

has every non-diagonal strictly positive and the matrix

$$P(t) = e^{-mB} I_{N \times N} + \int_{t-B}^{t} e^{-m(t-s)}(mI_{N \times N} - D(s) + A(s))\, ds$$

is stochastic such that

$$\underline{\kappa} = \inf_{t \geq t_0} \min_{i,j} \sum_l \min\{p_{il}(t), p_{jl}(t)\} > N \min\left\{e^{-mB}, \frac{1-e^{-mB}}{m}\underline{\alpha}\right\} \in (0,1).$$

If

1. $\sup_{t\geq 0} \max_{i,j} \tau_{ij}(t) = \tau < \infty$,

2. for every $B > 0$ and all $t \geq t_0$, the matrix $\int_t^{t+B} A(s)ds$ has every non-diagonal element strictly positive,

then Theorem 4.4.6 implies that $\exists k \in W_{I_{t_0},\boldsymbol{\phi}}$ such that

$$\max_i |x_i(t) - k| \leq \frac{S_{I_{t_0}}(\boldsymbol{\phi})}{1 - \underline{\kappa} e^{-(N-1)\bar{\alpha}\tau}} e^{-\theta(t-t_0)}$$

where $\theta = -\frac{\ln(1-\underline{\kappa}e^{-(N-1)\bar{\alpha}\tau})}{B+2\tau} > 0$.

A leader in a consensus network, is an agent that affects the rest of the group, but it cannot be affected by it. In the presence of a leader, say agent 0 with state z_0 to satisfy a generic differential equation $\dot{z}_0(t) = g\big(t, z_0(t)\big)$ modeling possibly internal dynamics under the hypothesis

$$|z_0(t) - \bar{z}| \leq Z e^{-\zeta(t-t_0)} \tag{7.2}$$

for some constants $Z, \zeta > 0$.

The "leader-follower" system can be written as

$$\begin{cases} \dot{z}_0(t) = g\big(t, z_0(t)\big) \\ \dot{z}_i(t) = \sum_{j\neq 0} a_{ij}\big(z_j(t-\tau_{ij}(t)) - z_i(t)\big) + a_{i0}\big(z_0(t-\tau_{i0}(t)) - z_i(t)\big), & t \geq t_0 \\ x_i(t) = \phi(t), & t \in I_{t_0} \end{cases}$$

where $i = 1, \ldots, N$. Then

$$|z_i(t) - \bar{z}| \leq K_1 e^{-\theta(t-t_0)} + K_2 \frac{e^{-\zeta(t-t_0)} - e^{-\theta(t-t_0)}}{\zeta - \theta}$$

where

$$K_1 = \frac{S_{I_{t_0}}(\boldsymbol{\phi})}{1 - \kappa e^{-(N-1)\bar{\alpha}\tau}}, \; K_2 = \frac{2Z\bar{\alpha}e^{\zeta\tau}}{\left(1 - \kappa e^{-(N-1)\bar{\alpha}\tau}\right)}.$$

as Theorem 4.4.7 enforces.

Remark 7.1.1. The aforementioned results hold for simple (recurrent) connectivity regimes, as well. In this case the rate estimates change for the worse and they are beyond the scopes of our work.

The presence of a leader does not alternate the qualitative behavior of the system other than the consensus point. This is the limit point of the leader. It is important to understand that the network eventually synchronizes to a constant value only when the leader converges to this value. Otherwise the followers, although never stop following the leader, they will not synchronize with it.

7.2 The Elementary EDP

A system of N power generating units, connected to a single bus bar serves a received electrical load P_{load} (Figure, 7.2). The input to each unit, shown as F_i represents the cost rate of the unit. The output of each unit P_i is the electrical power generated by that particular unit. The total cost rate of this system is the sum of the costs of each of the individual units. The essential constraint on the operation of this system is that the sum of the output powers must equal the load demand.

$$F_T = F_1 + F_2 + \cdots + F_N$$

$$F_T = \sum_{i=1}^{N} F_i(P_i)$$

$$\phi = 0 = P_{load} - \sum_{i=1}^{N} P_i$$

This is a constrained optimization problem that may be attacked formally using advanced calculus methods that involve the Lagrange function:

$$\mathcal{L} = F_T + \lambda\phi \tag{7.3}$$

The necessary conditions for an extreme value of the objective function result when we take the first derivative of $\mathcal{L}$ with respect to each independent variables and set the derivatives equal to 0:

$$\frac{\partial \mathcal{L}}{\partial P_i} = \frac{dF_i(P_i)}{dP_i} - \lambda = 0 \tag{7.4}$$

Following [16] we will assume that the cost functions $F_i(P_i)$ are smooth and quadratic:

$$F_i(P_i) = \frac{1}{2}\chi_i P_i^2 + \psi_i P_i + \omega_i \tag{7.5}$$

for some strictly positive parameters χ_i, ψ_i, ω_i assumed to be known.

Together with (7.4) we must add the constraint that the sum of the power outputs must be equal to the power demanded by the load. All in all, we have the following systems of equations that it is necessary be satisfied in the optimal operation point:

$$i = 1, \ldots, N : \begin{cases} \frac{dF_i(P_i)}{P_i} = \lambda \\ \sum_i P_i = P_{load} \end{cases} \tag{7.6}$$

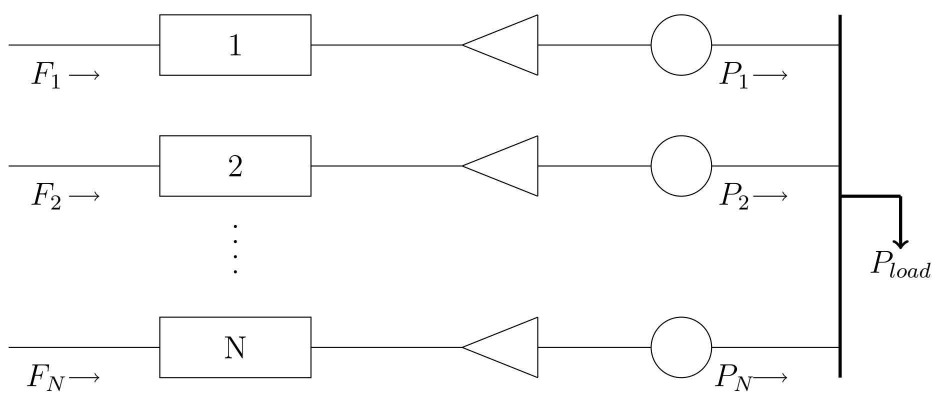

Figure 7.2: N units committed to serve P_{load}. The schematics follow [16]. The rectangular symbolize the boiler fuel input that determines the fuel cost function. The triangle sumbolizes the steam turbine and the circle is the power generator. In the distributed case, a sensor-controller is attached to a part of the network to supervise one power generatore and one load (part of P_{load})) as well as to execute an optimization problem.

We characterize this problem as elementary because important parameters were ignored. In the discussion session we will explain interesting and more realistic generalizations within the theoretical context which is to be developed in the section to follow.

7.3 Distributed Solution of the Loss-less EDP

Assume that at each thermal unit we have a controller that has full access to the parameters of its area of duty (load and generator) but it has limited information for the parameters of the other areas. The latter information is propagated through the communication network and it suffers from delays. In particular each controller i has

1. Instant information of the load in its section $P_{load}^{(i)}$ and transmission of it over the network. Here the i^{th} controller will be a leader of a consensus algorithm responsible to communicate this information to the rest of the controllers. We will adopt the notation $p_{(load,j)}^{(i)}$ the state of the follower $j \neq i$ on this algorithm. All in all, the vector

$$\mathbf{p}_{load}^{(i)} = \left(\ldots, p_{(load,i-1)}^{(i)}, p_{(load,i+1)}^{(i)}, \ldots \right)$$

symbolizes the states the sensor-followers, each of which seeks to learn the load the i^{th} controller transmits. These are all dynamic variables that follow a leader-follower consensus model. All in all we have N such models with possible different communication weights and delays of the type of (7.1) each of which has a leader i with $z_0 \equiv P^{(i)}_{load}$. We will assume that all, but the leader's, initial functions are set to zero. The vital characteristics of such a model are the lower and upper bounds of the communication weights and the maximum imposed delay. These are denoted as $\underline{\alpha}_i, \bar{\alpha}_i, \bar{\tau}_i$. The convergence of this model is guaranteed under a simple connectivity assumption as Remark 7.1.1 suggests. The vital parameters for th

2. Instant information on the power the i^{th} generator produces, denoted by $P^{(i)}_{gen}$. This will be dynamically updated and transmitted over the network satisfying the equation
$$P^{(i)}_{gen}(t) = \frac{\lambda_i(t) - \psi_i}{\chi_i}. \tag{7.7}$$
Here, the i^{th} controller communicates over the network the state $P^{(i)}_{gen}(t)$ with some delay.

3. Delayed received information of the signal $P^{(j)}_{gen}$ from all over the network. The i^{th} controller serves as a receiver of the generated power of the rest of the controllers with some delay.

7.3.1 The incremental cost algorithm

Each sensor chooses to update its state λ_i by averaging it with the rest of the sensors, as follows:

$$\dot{\lambda}_i(t) = \sum_j w_{ij}\big(\lambda_j(t - \tau_{ij}(t)) - \lambda_i(t)\big) + w^b_i\big(P^c_{load,i}(t) - P^c_{gen,i}(t)\big) \tag{7.8}$$

were w_i^b is a coupling positive control constant, $P^c_{load,i}(t)$ is the cumulative information of the load on the network, sensor i has at time t and $P^c_{gen,i}(t)$ is the cumulative information on the produced generator, sensor i has at time t. Using the notation above we deduce that

$$P^c_{load,i}(t) = P^{(i)}_{load} + \mathbb{1}^T \mathbf{p}^{(i)}_{load}(t) \tag{7.9}$$

$$P^c_{gen,i}(t) = P^{(i)}_{gen}(t) + \sum_{j \neq i} P^{(i)}_{gen}(t - \tau_{ij}(t)) \tag{7.10}$$

7.3.2 Analysis

The target of the sensors on a consensus value is $\lambda_i(t) \equiv \lambda_\infty$. From (7.5) we have the fixed point of $P_{gen,i}$

$$P^\infty_{gen,i} = \frac{\lambda_\infty - \psi_i}{\omega_i}.$$

Then the optimal operation point is

$$P^\infty_{gen} = \sum_{l=1}^{N} P^\infty_{gen,l} = P_{load} = \lambda_\infty \sum_{l=1}^{N} \frac{1}{\chi_l} - \sum_{l=1}^{N} \frac{\psi_l}{\chi_l} \tag{7.11}$$

Then limit consensus point, the sensors try to reach is:

$$\lambda_\infty = \frac{P_{load} + \sum_{l=1}^{N} \frac{\psi_l}{\chi_l}}{\sum_{l=1}^{N} \frac{1}{\chi_l}} \tag{7.12}$$

$$\dot{\lambda}_i(t) = \sum_{j \in N_i} w_{ij}\big(\lambda_j(t - \tau_{ij}(t)) - \lambda_i(t)\big) + w_i^b\big(P_{load} - P^c_{gen,i}(t)\big) +$$

$$+ w_i^b\big(P^c_{load,i}(t) - P_{load}\big)$$

Now,

$$
\begin{aligned}
P^c_{load,i}(t) - P^c_{gen,i}(t) &= \frac{\lambda_\infty - \lambda_i(t)}{\chi_i} + \sum_{j\neq i}\left(P^\infty_{gen,j} - P^{(i)}_{gen}(t-\tau_{ij}(t))\right) \\
&= \frac{\lambda_\infty - \lambda_i(t)}{\chi_i} + \sum_{j\neq i}\frac{\lambda_\infty - \lambda_j(t-\tau_{ij}(t))}{\chi_j} \\
&= \left(\sum_j \frac{1}{\chi_j}\right)\big(\lambda_\infty - \lambda_i(t)\big) + \sum_j \frac{1}{\chi_j}(\lambda_i(t) - \lambda_j(t-\tau_{ij}(t))) \\
&= \left(\sum_j \frac{1}{\chi_j}\right)\big(\lambda_\infty - \lambda_i(t)\big) + \sum_j \frac{1}{\chi_j}\big(\lambda_i(t) - \lambda_j(t-\tau_{ij}(t))\big)
\end{aligned}
$$

so the consensus algorithm is written as

$$
\begin{aligned}
\dot{\lambda}_i(t) = &\sum_j \left(w_{ij} - \frac{w^b_i}{\chi_j}\right)\big(\lambda_j(t-\tau_{ij}(t)) - \lambda_i(t)\big) + \\
&\qquad + \left(\sum_j \frac{w^b_i}{\chi_j}\right)\big(\lambda_\infty - \lambda_i(t)\big) + g_i(t)
\end{aligned}
\tag{7.13}
$$

where

$$
g_i(t) = w^b_i\big(P^c_{load,i}(t) - P_{load}\big). \tag{7.14}
$$

Equation (7.13) is a perturbed consensus system with a virtual leader of constant value λ_∞. The weights of the new network $A = [a_{ij}]_{i,j\in\{0,N\}}$

$$
a_{ij}(t) = \begin{cases}
a_{ii} = 0, & i = j \\
0, & i = 0 \\
\sum_{j=1}^N \frac{w^b_i(t)}{\chi_j}, & j = 0 \\
w_{ij}(t) - \frac{w^b_i(t)}{\chi_j}, & i \neq 0, j \neq 0
\end{cases} \tag{7.15}
$$

Then using Theorem 3.2.4 we deduce that $\boldsymbol{\lambda}(t)$ converges to the optimal economic point if there is an $c > 0$ such that

$$
\inf_{t\geq t_0}\left(w_{ij}(t) - \frac{w^b_i}{\chi_j}\right) \geq c > 0 \qquad \forall i \neq j \tag{7.16}
$$

the convergence of the algorithm occurs exponentially fast and the spread of the vector $\boldsymbol{\lambda}$, $\sup_{t\geq 0}(\min_i \lambda_i(t) - \max_i \lambda_i(t))$, is upper bounded by the constants K_1 and K_2 as they were defined in § 7.1 as explicit functions of the systems' parameters and Z is the constant determined accordingly as it is the consensus system under which each sensor communicates (learns) the load of the network acting as a leader (follower). Consequently,

$$|Z| \leq \sum_{i=1}^{N-1} \frac{P_{load}^{(i)}}{1 - \bar{\rho}_i e^{-(N-1)\bar{\alpha}_i \bar{\tau}_i}} \tag{7.17}$$

as it was explained in the beginning of this section.

7.4 A Simulation Example

We will outline the previous analysis with an illustrative example taken from [16] (page 65). Here a network of $N = 3$ units generates power to serve a cumulative load of

$$P_{load} = 850 \text{ MW}$$

Each sensor is set to control part of this load and one generator as follows

1. Sensor 1: 200 MW of load and generates $P_{gen}^{(1)}$. The fuel cost function is

$$F_1(P_{gen}^{(1)}) = 561 + 7.92 P_{gen}^{(1)} + 0.001562 (P_{gen}^{(1)})^2$$

2. Sensor 2: 300 MW of load and generates $P_{gen}^{(2)}$. The fuel cost function is

$$F_2(P_{gen}^{(2)}) = 310 + 7.85 P_{gen}^{(2)} + 0.00194 (P_{gen}^{(2)})^2$$

3. Sensor 3: 350 MW of load and generates $P_{gen}^{(3)}$. The fuel cost function is

$$F_3(P_{gen}^{(3)}) = 78 + 7.97 P_{gen}^{(3)} + 0.00482 (P_{gen}^{(3)})^2$$

The units of F_i are in \$/hr. Using (7.4) and the condition $\sum_{i=1}^{3} P_{gen}^{(i)} = 850$ we derive the economic (consensus) point

$$\lambda_\infty = 9.148\$/\text{MWhr}.$$

For the distributed solution of the above EDP we set $t_0 = 0$ and we consider the primary time-varying communication network

$$A = \begin{bmatrix} 0 & 0.5 & \frac{0.7+t^2}{t^2+1} \\ 1+\cos^2(1+t^2) & 0 & 6(1+e^{-t}) \\ 0.9 & \frac{0.7+t}{0.5+t} & 0 \end{bmatrix}$$

and the delays

$$T = \left[\tau_{ij}(t)\right] = \begin{bmatrix} 0 & 1 & 1 \\ 0.23 & 0 & 0.23 \\ 0.8\cos(16\pi t) & 0.8\cos(16\pi t) & 0 \end{bmatrix}$$

The secondary network involves the dynamic process of $P_{load}^{(i)}$ from $\mathbf{p}_{load}^{(i)}$. In particular:

Sensors 1, 2 learn P_{load}^{3} with the state vector $(p_{load,1}^{(3)}, p_{load,2}^{(3)})$ under the network

$$\begin{cases} \frac{d}{dt} p_{load,1}^{(3)}(t) = 1.2\sin^2(t)\big(P_{load}^{(3)} - p_{load,1}^{(3)}(t)\big) \\ \frac{d}{dt} p_{load,2}^{(3)}(t) = 0.02\big(p_{load,1}^{(3)}(t-0.5) - p_{load,2}^{(3)}(t)\big) \end{cases}$$

Sensors 1, 3 learn P^2_{load} with the state vector $(p^{(2)}_{load,1}, p^{(2)}_{load,3})$ under the network

$$\begin{cases} \frac{d}{dt}p^{(2)}_{load,1}(t) = 0.8\sin^2(2\pi t)\big(p^{(2)}_{load,3}(t-0.23) - p^{(2)}_{load,1}(t)\big) \\ \\ \frac{d}{dt}p^{(2)}_{load,3}(t) = 0.6\sin^2(3\pi t)\big(P^{(2)}_{load} - p^{(2)}_{load,3}(t)\big) \end{cases}$$

Sensors 2, 3 learn P^1_{load} with the state vector $(p^{(1)}_{load,2}, p^{(1)}_{load,3})$ under the network

$$\begin{cases} \frac{d}{dt}p^{(1)}_{load,2}(t) = 0.3\big(p^{(2)}_{load,3}(t-0.5) - p^{(1)}_{load,2}(t)\big) \\ \\ \frac{d}{dt}p^{(1)}_{load,3}(t) = 1.7\big(P^{(1)}_{load} - p^{(1)}_{load,3}(t)\big) \end{cases}$$

Finally, the balance vector

$$\Big(w^b_1, w^b_2, w^b_3\Big)^T$$

is set to a common control constant $w^b_i \equiv w$. It is easy to see that the primary consensus system corresponds to a fully connected communication graph and the secondary consensus systems correspond to a simple connected graph. Also, all the delays are bounded. Then the results of § 7.1 together Remark 7.1.1 apply and the aforementioned analysis holds with numerically calculated parameters $\zeta < 0.02$, $Z \leq 855$, $\underline{\alpha} = 0.35$, $\bar{\alpha} = 6$ and the largest delay is $\tau = 1$ and it is only a simple calculation to K_1 and K_2. The most important criterion however is (7.16). This imposes the smallness condition

$$w < w^* = 0.00194$$

Simulations are provided in Figure 7.3 for control values below and beyond w^*. We observe that whenever w is above this critical value the algorithm slows down or does not converge.

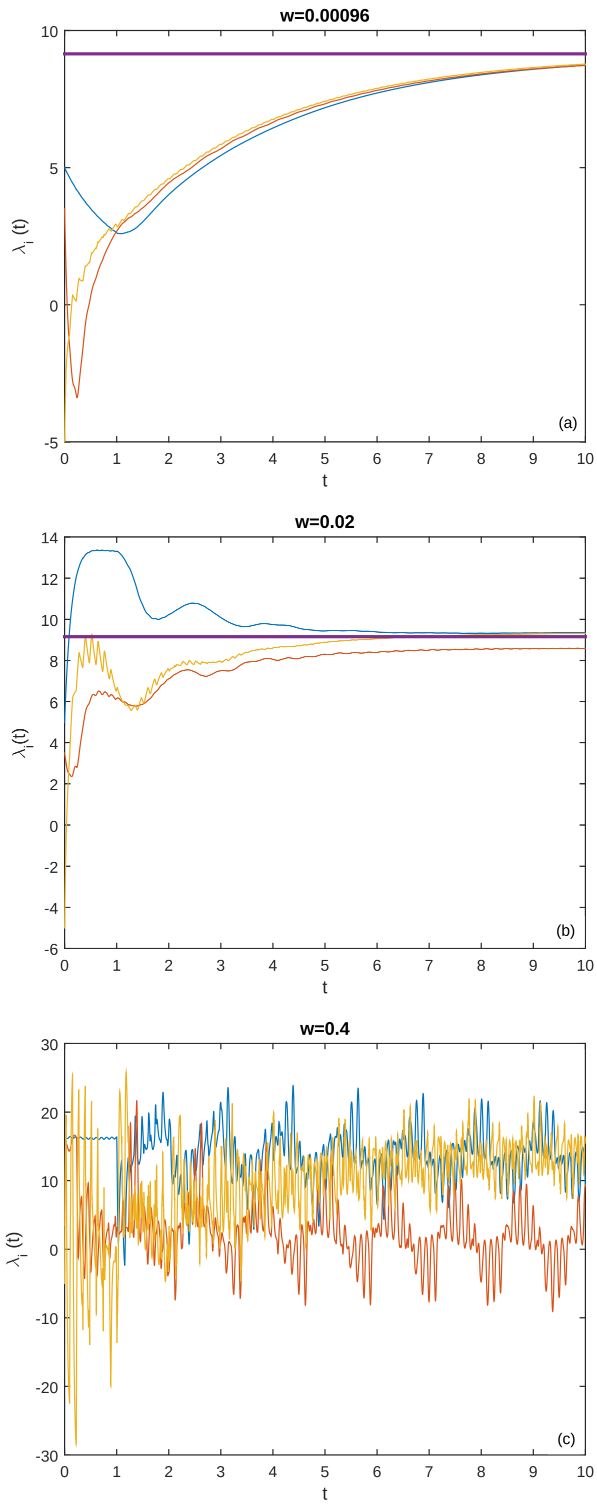

Figure 7.3: Simulations run in `MATLAB` with the `ddesd` routine. The distributed incremental cost solutions $\lambda_i(t)$ appears to converge very fast for small w. (b) As w increases beyond w^* we still see oscillatory and slower convergence. (c) For w large the algorithm does not converge and for even larger values it diverges.

Chapter 8: Discussion- Future Research

The present thesis investigates distributed consensus schemes and develops a unified framework for the study of the stability of their solutions. The discussion spans from discrete and continuous schemes, linear and nonlinear facets, static and switching connectivity regimes drive by signals of deterministic or stochastic nature. The analysis is primarily focused on the stability of solution with respect to the consensus space. In this concluding part we will discuss the main results chapter by chapter from the point of view of posing new questions for future research.

Chapter 3 The purpose of this chapter is to pose a unified perspective on the old problem of linear consensus. We studied the deterministic and the stochastic variations separately. For the first case the objective was to derive the strongest convergence results possible, under the mildest assumptions. This requires clear estimates on the rate of convergence taking into account scenarios of switching/failing signals and essentially asymmetric couplings. The most important novelty is the adaptation of the contraction coefficient to the continuous time framework.

The utilization of the contraction coefficient for the study of consensus dynamic yields simple and concise convergence results without strong assumptions on the coupling weights or the connectivity regime for both the ordinary and the delays case. The rate estimates depend on the coupling weights, the switching signal, the parameter $B > 0$ and the magnitude of the delays. On the positive side, we have a rigorous theory that provides a deep insight in the behavior of these very important family of systems. The framework is flexible enough to be effectively applied to other consensus systems. On the negative side, the rate estimates are very weak, as the simulations clearly suggest.

The following questions occur only naturally: How one should choose the values of $m > 0$ or $B > 0$ (the role of which has been clarified in §3.4) and, eventually, for what type of weights and connections can these rates be improved. Certain types of topological graphs known as *expander graphs* achieve high efficiency when executing the consensus algorithm with sparse connectivity [123]. The main feature is that whereas N increases, the second smallest eigenvalue λ_2 does not decrease. What is (if there exists one) the connection between these communication graphs and the contraction coefficient?

In the stochastic framework we demonstrated that the mildest sufficient connectivity conditions, the recurrent connectivity is essentially reproduced. Although there are counter-examples that show that for milder connectivity, consensus may not be achieved [65], there is no definitive result on the "bifurcation" connectivity regimes between which consensus is always achieved and consensus is not achieved. Necessary conditions, some of them also appear here involve the strenght of the communication rates and not their appearance/disappearance throughout the evolution of the system dynamics.

Chapter 4 Delays play a very important role in the existence, uniqueness and stability of the solutions of functional differential equations. The objective of the chapter is to develop a rigorous framework for the study of the stability of linear systems with asymptotically constant solutions both in the scalar and the multi-dimensional case. The emphasis is, yet again, concentrated on explicit estimates on the rate of convergence. Each subject has an interest of its own. While consensus networks lie in the heart of distributed control dynamics, the scalar dynamics resurrect an old discussion, started nearly 30 years ago, and concerned problems on simple biological processes. Our approach extends and unifies the majority of the mentioned results under a framework that is partitioned between a fixed point theory approach for the systems with time-invariant parameters and a Lyapunov-Razumikhin approach. The latter measures the contraction rate of sets constructed out of segments of the solution for the systems with time-varying parameters.

The conclusion for the scalar case revolves around the sign of the weight parameters. Following the notation of the chapter, whenever the weights are positive, the systems exhibit a co-operative behavior that yields convergence to a constant value regardless of the magnitude of the delay. The magnitude has effect only on the rate of the asymptotic convergence. The larger the delay is, the slower the rate estimate. Whenever the sign is negative, the system exhibits competitive behavior and instability occurs for large delays. The scalar equation with constant parameters incorporates both types of coupling weights and fully illustrates this phenomenon. In the general linear case however, the analysis becomes more complex. Coupling rates of processing delays were systematically neglected and we merely commented on ways of solving the problem in the general case.

An important distinction is this between the time-varying and time-invariant case for a number of reasons. Regardless the dimension of the problem, time-invariant parameters are not only easier to handle, but they also provide additional information such as the closed form of the consensus point. The simplicity of these dynamics allowed us to consider multiple processing delays as well, a feature not present in the general model.

A central open question is whether fixed point and Lyapunov-Razumikhin methodologies could be combined to enhance the rate estimates in the presence of both propagation and processing delays. Indeed, several different approaches may be implemented. All of them follow the spirit of variational approximation of systems and each option is to be chosen on the base of the particular perspective the researcher wants to impose. For example, in the LTI network, Eq. (4.30) is an approximation of the system without the processing delay σ. Then the analysis can hold only for sufficiently small values of σ. The variation of parameters method could possibly be used more efficiently. Based on the stability

$$\dot{y} = -ay(t-\sigma)$$

for certain values of a and σ, we can express the solution of

$$\dot{x}_i = -\sum_j a_{ij}x_i(t-\sigma) + a_{ij}x_j(t-\tau)$$

as a variation of the y. The price to pay is that this approach may weaken even more the stability bounds for τ. Eventually, it is only a matter of choice and taste on behalf of the researcher. He is the one who will determine the corresponding necessary assumptions.

Chapter 5 The research in the non-linear alternatives reveals a number of interesting topics for discussion. While passivity or monotonicity properties clearly resemble the linear dynamics so that the same machinery can be used, the case of non-monotonicity signifies the need for developing more efficient mathematical tools. The contraction coefficient is clearly unsuitable to deal with this case. As it is remarked in [18] competitive dynamics exhibit much richer behavior. A very interesting problem would be to study the effect of non-linear (i.e. state-dependent) delays in competitive systems. We recall from Chapter 4 that for a competitive scalar equation a large delay induces instability. A state-dependent delay would then establish an interplay between stability and instability creating a much more interesting behavior such as this of periodic or chaotic solutions. The case can then be elevated to the multidimensional consensus based networks that will not converge to an agreement value but possibly to new chaotic invariant sets.

Finally, the dynamics of neutral consensus systems were introduced and studied on a primary and epidermal level. The motivation of this model occurred from the observation that the dynamics of a living organism or a modern computing machine should evolve in time as a function of the current or previous state only, may be an over-simplistic hypothesis. In real world problems the rate of change of a state also depends on the rate of change of the same state at a previous time. In Nature, for example, the acceleration of a flying bird at a particular moment cannot but be also a function of its acceleration at a previous time. These sorts of correlations are consistently ignored when designing mathematical models, exactly because their analysis is particularly difficult. In this end of this chapter we introduced and developed a framework on such distributed systems of autonomous agents that execute a simple consensus algorithm with a supplementary non-linear neutral term. A next step is to further clarify the form of the term under more realistic scenarios and exploit these assumptions to establish clearer results on the long term behavior of the systems in question and to investigate the potential physical meaning of its solutions. For example, if we really want to model the neutral term as a term of dynamic weakening of the stamina of the agent, then the function must clearly satisfy a monotonic condition, the effect of which is obviously crucial for the stability of the solutions and should be taken into account in the rigorous analysis.

Consensus couplings are a special case of synchronization networks [124]. A natural question is if we can extend the analysis for this type of systems. Do, for instance, networks of coupled chaotic oscillators accept the methodologies developed in the thesis? If yes, to what extent? Can stability in variation provide an, alternative to the standard Lyapunov methods, approach for the study for these systems, as well? Would the method of combining the solution lead to any efficient, or at least meaningful, convergence conditions?

Chapter 6 In the case of nonlinear flocking networks we have a much clearer view for future research. A first generalization is to consider processing delays as well. This is not an easy problem as the general approach developed in Chapter 4 does not apply any more. In addition, sharpening the noisy flocking results along the lines explained at the end of Chapter 6 is a feasible extension of the results derived so far. Furthermore, the development of flocking algorithms with collision avoidance along the lines of [91] is an interesting future research topic. Also, the effect of state-dependent delays in these systems considered not only on the microscopic scale (differential equations) but also on the mesoscopic / macroscopic scales (partial differential equations) (see for example [94, 89, 55]) pave the way for many interesting research topics. We conjecture that the effect of state-dependent delays on the meso/macro-scopic scale may allow for more important behavior especially on the transient level.

Chapter 7 The solution of the EDP problem in a distributed manner is very important for the modern smart grid architectures. In this paper, we introduced a consensus based optimization algorithm that greatly improves the existing ones [119, 121, 120, 122]. The theory develops a decentralized version of the lambda iteration algorithm with emphasizing on the communication network that controls the process. The optimal economic point is dynamically achieved via a communication network that suffers from multiple and complex delays. The present work is but a small step towards merging two very interesting fields of networked control systems: this of agreement dynamics and this of the modern electric power networks in the smart grid environment. Several things are yet to be addressed concerning this work.

The elementary EDP which we basically analyzed, neglects the fact that every power generator operates within limits. For the classical EDP one must substitute (7.6) with

$$i = 1, \ldots, N : \begin{cases} \frac{dF_i}{P_i} = \lambda \\ P_{i,\min} \leq P_i \leq P_{i,\max} \\ \sum_i P_i = P_{load} \end{cases}$$

Within the developed setup, this issue can be tackled in two steps. The first is to recall that theory predicts explicit bounds on the difference of $\min_i \lambda_i - \max_i \lambda_i$. Since $P_{gen}^{(i)}$ can be expressed as a linear function of λ_i we are half way far from explicit bounds on the generated power $P_{gen}^{(i)}$. Indeed every sensor needs information on all the cost parameters α_j, β_j. Unless one is willing to set these parameters within universal standard bounds, these are information a sensor needs to learn from the network. Note that an important point is that for the EDP to have a solution, λ_∞ to be within the operation region of the generators. This is not always the case. Therefore a first extension is to develop algorithms that will, or attempt, to solve the EDP with given operating constraints.

A second problem with our approach is that we require an all-to-all connectivity even with arbitrary delays. It is very important to decentralize the architecture even further. It is not clear, however, how this could be achieved without critically destabilizing the algorithm. Even in the toy example of Section 7.4, this complete communication regime could not suffice to stabilize the algorithm. This reveals the sensitivity of the consensus algorithm on perturbations and delays. Future research along this line would require a study on the connectivity regimes and how these affect the rate estimates and the stability bounds.

Another simplification assumption followed here, is that the power network is loss-less. If the energy network has energy losses, the EDP derives a slightly more complex Langrangian with an extra incremental cost value. Incorporating this factor on the dynamic algorithm is also a necessary step for designing more realistic theoretical algorithms. Given both $F_i(P_i)$ and the cost function of the transmission lines $F_{loss}(P_i)$ in quadratic form we conjecture that our theory could be adequately extended.

All the above observations unavoidably point to the final remark of our work: The form of the cost functions and the quadratic assumption. All the aforementioned questions can (and should) be repeated for more general cost functions. This is a fairly challenging issue even for the conventional methods [16].

We conclude by repeating the importance of applying distributed learning techniques as part of the MAS technology in networks where the main objective is fast managing operations handled automatically and effectively, with respect to strict quality standards, by as few central authorities as possible. A most typical network of this type is the electric Smart Grid. This engineering vision automatically highlights the importance of consensus based optimization algorithms and their contribution to even elementary problems such the one we studied here.

Bibliography

[1] C. Somarakis and J.S. Baras. The consensus problem under vanishing communication weights. *Proceedings of the 20th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2012)*, pages 1–8, 2012.

[2] C. Somarakis and J.S. Baras. Linear processes under vanishing communications - the consensus problem. Technical Report TR-2012-03, Univ. Of Md., College Park, USA, 2012.

[3] C. Somarakis and J.S. Baras. Asymptotic consensus solutions in non-linear delayed distributed algorithms under deterministic & stochastic perturbations. In *52nd IEEE Conference on Decision and Control*, December 2013.

[4] C. Somarakis and J.S. Baras. Fixed point theory approach to exponential convergence in ltv continuous time consensus dynamics with delays. In *SIAM Control Conference*, July 2013.

[5] C. Somarakis and J.S. Baras. A fixed point theory approach to multi-agent consensus dynamics with delays. Technical Report TR-2013-01, Univ. Of Md., College Park, USA, 2013.

[6] C. Somarakis and J.S. Baras. On the convergence rate of non-linear consensus networks with delays. In *Proceedings of 9th IFAC Symposium on Nonlinear Control Systems*, September 2013.

[7] C. Somarakis and J.S. Baras. LTI continuous time consensus dynamics with multiple and distributed delays. In *21st Mediterranean Conference on Control & Automation*, pages 1495–1501, June 2013.

[8] C. Somarakis and J.S. Baras. Towards a unified theory of consensus. Technical Report doi:10.13016/M2MW3B, Univ. Of Md., College Park, USA, 2014.

[9] C. Somarakis and J.S. Baras. Consensus networks of neutral type. In *23rd Mediterranean Conference on Control & Automation*, Torremolinos, Spain, June 2015. to appear.

[10] C. Somarakis and J.S. Baras. Delay-independent convergence for linear consensus networks with applications to non-linear flocking systems. In *The 12th IFAC Workshop on Time Delay Systems*, Ann Arbor, MI, USA, June 2015. to appear.

[11] C. Somarakis and J.S. Baras. A general framework for consensus systems. submitted to SIAM Journal of Mathematical Analysis doi:10.13016/M2WK77, Univ. Of Md., College Park, USA, 2015. http://hdl.handle.net/1903/16324.

[12] C. Somarakis and J.S. Baras. New resuls on stochastic consensus networks. In *SIAM Control Conference*, Paris,France, July 2015. to appear.

[13] C. Somarakis and J.S. Baras. A simple proof of the continuous time linear consensus problem with applications in non-linear flocking networks. In *the 14th European Control Conference*, Linz,Austria, July 2015. to appear.

[14] C. Somarakis, J.S. Baras, and E. Paraskevas. On the dynamics of linear functional differential equations with asymptotically constant solutions. Technical Report doi:10.13016/M2CC7H, Univ. Of Md., College Park, USA, 2014. http://hdl.handle.net/1903/15580.

[15] C. Somarakis, J.S. Baras, and E. Paraskevas. Stability by fixed point theory in time delayed distributed consensus networks. Technical Report (submitted to Automatica) doi:10.13016/M27P47, Univ. Of Md., College Park, USA, 2014. http://hdl.handle.net/1903/15581.

[16] A.J. Wood, B.F. Wollenberg, and G.B. Sheblé. *Power Generation, Operation and Control*. Wiley, 2013.

[17] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.

[18] H.L Smith. *Monotone Dynamical Systems, An Introduction to the Theory of Competitive and Cooperative Systems*, volume 41. AMS, 1995.

[19] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226–1229, 1995.

[20] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[21] M.H. DeGroot. Reaching a consensus. *Journal of American Statistical Association*, 69:118–121, 1974.

[22] A.A. Markov. Extensions of the law of large numbers to dependent quantities (in russian). *Izn. Fiz. - Matern. Obsch. Kazan Univ.*, 15, 1906.

[23] B. Bollobás. *Modern Graph Theory*, volume 184. Springer, 1998.

[24] C. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207. Springer, 2001.

[25] M. Mesbahi and M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.

[26] D.J. Hartfiel. *Markov Set-Chains*, volume 1695. Springer, 1998.

[27] D.J. Hartfiel. *Nonhomogeneous Matrix Products*. World Scientific, 2002.

[28] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer, 2006.

[29] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*, volume 54. Cambridge University Press, 1996.

[30] P. Walters. *An introduction to Ergodic Theory*, volume 79. Springer, 1982.

[31] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. Wiley-Interscience publication, 1974.

[32] G. Grimmett and D. Stirzaker. *Probability and Random Processes*. Oxford Univ. Press, 3 edition, 2003.

[33] L.B. Koralov and Y.G. Sinai. *Theory of Probability and Random Processes*. Springer, 2 edition, 2007.

[34] T.A. Burton. *Stability by Fixed Point Theory for Functional Differential Equations*. Dover, 2006.

[35] H.L. Royden. *Real Analysis*. McMillan Publishing Company, 3 edition, 1989.

[36] D.R. Smart. *Fixed Point Theorems*. Cambridge University Press, 1980.

[37] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1997.

[38] N. G. Markley. *Principles of Differential Equations*. Pure and Applied Mathematics. Wiley-Intersciense Series of Texts, Monographs, and Tracts, New Jersey, 2004.

[39] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(98):298–305, 1973.

[40] F.R.K. Chung. *Spectral Graph Theory*, volume 92. American Mathematical Society, 1997.

[41] I. C. F. Ipsen and T. M. Selee. Ergodicity coefficients defined by vector norms. *SIAM J. Matrix Anal. Appl.*, 32:153–200, March 2011.

[42] J. Schauder. Der Fixpunktsatz in Funktionalraümen. *Studia Math.*, 2:171–180, 1930.

[43] N.A. Krasnoselskii. Some problems of nonlinear analysis. *Amer. Math. Soc. Transl.*, 2(10):345–409, 1958.

[44] J. Schauder. Über den zusammenhang zwischen der eindeutigkeit und lösbarkeit partieller differentialgleichungen zweiter Ordnung von elliptischen Typus. *Math. Ann.*, 106:661–721, 1932.

[45] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520 – 1533, sept. 2004.

[46] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11):1867–1872, 2005.

[47] M. Porfiri and D. L. Stilwell. Consensus seeking over random weighted directed graphs. *IEEE Transactions on Automatic Control*, 52(9):1767–1773, 2007.

[48] A. Tahbaz-Salehi and A. Jadbabaie. A necessary and sufficient condition for consensus over random networks. *IEEE Transactions on Automatic Control*, 53(3):791–795, 2008.

[49] A. Tahbaz-Salehi and A. Jadbabaie. Consensus over ergodic stationary graph processes. *IEEE Transactions on Automatic Control*, 55(1):225–230, 2010.

[50] I. Matei, J. S. Baras, and C. Somarakis. Convergence results for the linear consensus problem under markovian random graphs. *SIAM J. Control Optim.*, pages 1574–1591, 2013.

[51] L. Moreau. Stability of continuous-time distributed consensus algorithms. *43rd IEEE Conference on Decision and Control (2004)*, page 21, 2004.

[52] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on Automatic Control*, 52(5):852–862, 2007.

[53] U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. *Communications in difference equations*, pages 227–236, 2000.

[54] F. Riesz and B. Szőkefalvi-Nagy. *Functional Analysis.* Frederick Ungar Publishing, New York, 1972.

[55] S. Motsch and E. Tadmor. Heterophilious dynamics enhances consensus. *SIAM: Review*, 56(4):577–621, 2014.

[56] T.A. Burton. *Stability and Periodic Solutions of Ordinary and Functional Differential Equations.* Dover, 2005.

[57] R.D. Driver. *Ordinary and Delay Differential Equations.* Number 20 in Applied Mathematical Sciences. Springer, 1977.

[58] K. Gu, J. Chen, and V.L. Kharitonov. *Stability of Time-Delay Systems*. Addison-Wesley series in electrical and computer engineering: Control engineering. Birkhäuser Boston, 2003.

[59] J.K. Hale and S.M.V. Lunel. *Introduction to Functional Differential Equations*. Number v. 99 in Applied mathematical sciences. Springer-Verlag, 1993.

[60] F.M. Atay. *Complex Time-Delay Systems: Theory and Applications*. Springer Publishing Company, Incorporated, 1 edition, 2010.

[61] B. Balachandran, T. Kalmár-Nagy, and D.E. Gilsinn. *Delay Differential Equations: Recent Advances and New Directions*. Springer U.S., 2009.

[62] M. Lakshmanan and D.V. Senthilkumar. *Dynamics of Nonlinear Time-Delay Systems*. Springer Series in Synergetics. Springer, 2011.

[63] W. Michiels and S.I. Niculescu. *Stability, Control and Computation for Time-Delay Systems: An Eigenvalue-Based Approach*. Advances in Design and Control. Society for Industrial and Applied Mathematics, 2 edition, 2014.

[64] R. Sipahi, T. Vyhlídal, S.I. Niculescu, and P. Pepe. *Time Delay Systems: Methods, Applications and New Trends*. Lecture Notes in Control and Information Sciences. Springer, 2012.

[65] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pages 2996 – 3000, Dec. 2005.

[66] A. Papachristodoulou, A. Jadbabaie, and U. Munz. Effects of delay in multiagent consensus and oscillator synchronization. *IEEE Transactions on Automatic Control*, 55(6):1471 –1477, june 2010.

[67] U. Munz, A. Papachristodoulou, and F. Allgower. Delay-dependent rendezvous and flocking of large scale multi-agent systems with communication delays. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 2038 –2043, dec. 2008.

[68] V. Lakshminkantham and Deo S.G. *Method of Variation of Parameters for Dynamic Systems*, volume 1. Gordon and Breach Science, 1998.

[69] J.R. Haddock and J. Terjeki. Liapunov-razumikhin functions and an invariance principle for functional differential equations. *Journal of Differential Equations*, 48:95–122, 1983.

[70] K.L. Cooke and J.A. Yorke. Some equations modeling growth processes and gonorrhea epidemics. *Math. Biosci.*, 16:75–101, 1973.

[71] F.V. Atkinson, J.R. Haddock, and O.J. Staffans. Integral inequalities and exponential convergence of solutions of differential equations with bounded delay. *Ordinary and partial differential equations-Lecture Notes in Math.*, pages 56–68, 1982.

[72] F.V. Atkinson and J.R. Haddock. Criteria for asymptotic constancy of solutions of functional differential equations. *Journal of Math. Analysis and Applications*, 91:410–423, 1983.

[73] H. Bereketoglu and M. Pituk. Asymptotic constancy for non-homogeneous linear differential equations with unbounded delays. In *Proc. of the* 4^{th} *International Conference of Dynamical Systems and Differential Equations*, pages 100–107, Wilmington, NC, USA,, 2002.

[74] J. Gallardo and M. Pinto. Asymptotic constancy of solutions of delay-differential equations of implicit type. *Rocky Mountain Journal of Mathematics*, 28(2):487–504, 1998.

[75] J.R. Haddock, T. Krisztin, and J. Terjeki. Comparison theorems and convergence properties for functional differential equations with infinite delay. *Acta Sci. Math.*, 52:399–414, 1988.

[76] J.R. Haddock and Y. Kuang. Asymptotic theory for a class of non-autonomous delay differential equations. *Journal of Math. Analysis and Applications*, 168:147–162, 1992.

[77] J.R. Haddock, M.N. Nkashama, and J. Wu. Asymptotic constancy for linear neutral volterra integrodifferential equations. *Tohoku Math. J.*, 41:689–710, 1989.

[78] T. Krisztin. On the rate of convergence of solutions of functional differential equations. *Funkcialaj Ekvacioj*, 29:1–10, 1986.

[79] J. Petzold and M. Sorg. Generalized caldirola equations for the classically radiating electron. *Z. Physik*, A(283):207–215, 1977.

[80] E. J. Putzer. Avoiding the jordan canonical form in the discussion of linear systems with constant coefficients. *The American Mathematical Montly*, 73(1):2–7, 1966.

[81] M. Arcak. Passivity as a design tool for group coordination. *IEEE Transactions on Automatic Control*, 52(8):1380–1390, 2007.

[82] J. Lorenz and D. A. Lorenz. On conditions for convergence to consensus. *IEEE Transactions on Automatic Control*, 55(7):1651–1656, 2010.

[83] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50(02):169–182, 2005.

[84] A. Papachristodoulou, A. Jadbabaie, and U. Mnz. Effects of delay in multi-agent consensus and oscillator synchronization. *IEEE Transactions on Automatic Control*, 55(6):1471–1477, 2010.

[85] H. Qing and W.M. Haddad. Distributed nonlinear control algorithms for network consensus. *Automatica*, 44(9):2375 – 2381, 2008.

[86] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, 1984.

[87] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

[88] F. Cucker and S. Smale. On the mathematics of emergence. *Japan Journal of Mathematics*, 2:197–227, 2007.

[89] S-Y. Ha and J. G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Commun. Math. Sci.*, 7(2):297–325, 2009.

[90] S.M. Ahn and S.-Y. Ha. Stochastic flocking dynamics of the Cucker-Smale type with multiplicative white noises. *Journal of Mathematical Physics*, 51:1–17, 2010.

[91] F. Cucker and J. G. Dong. A general collision-avoiding flocking framework. *IEEE Transactions on Automatic Control*, 56(5):1124–1129, 2011.

[92] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. Pures Appl.*, 89:278–296, 2008.

[93] K. Ha, S-Y. Lee and D. Levy. Emergence of time-asymptotic flocking in a stochastic cucker-smale system. *Commun. Math. Sci.*, 7(2):453–469, 2009.

[94] S-Y. Ha and Tadmor E. From particle to kinetic and hydrodynamic description of flocking. *Kinetic and Related Models*, 1(3):415–435, 2008.

[95] F. Cucker, S. Smale, and D. Zhou. Best choices for regularization parameters in learning theory. *Found. Comput. Math.*, 4:315–343, 2004.

[96] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic. Interaction ruling animal collective behavior depends on topological rather than metric distance. *Proc. Natl. Acad. Sci. USA*, 105:12321237, 2008.

[97] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics*, 144(5):923–947, 2011.

[98] H. Saadat. *Power System Analysis*. PSA Pub., 2010.

[99] P.W. Sauer and A. Pai. *Power System Dynamics and Stability*. Stipes Publishing L.L.C., 2006.

[100] T. Jayabarathi, G. Sadasivam, and V. Ramachandran. Evolutionary programming based economic dispatch of generators with prohibited operating zones. *Electric Power Systems Research*, 52(3):261 – 266, 1999.

[101] J.O. Kim, D.J. Shin, J.N Park, and C. Singh. Atavistic genetic algorithm for economic dispatch with valve point effect. *Electric Power Systems Research*, 62(3):201 – 207, 2002.

[102] F. Li, R. Morgan, and D. Williams. Hybrid genetic approaches to ramping rate constrained dynamic economic dispatch. *Electric Power Systems Research*, 43(2):97 – 103, 1997.

[103] A. Pereira-Neto, C. Unsihuay, and O.R. Saavedra. Efficient evolutionary strategy optimisation procedure to solve the nonconvex economic dispatch problem with generator constraints. *Generation, Transmission and Distribution, IEE Proceedings-*, 152(5):653–660, Sept 2005.

[104] N. Sinha, R. Chakrabarti, and P.K. Chattopadhyay. Evolutionary programming techniques for economic load dispatch. *IEEE Transactions on Evolutionary Computation*, 7(1):83–94, Feb 2003.

[105] D.C. Walters and G.B. Sheble. Genetic algorithm solution of economic dispatch with valve point loading. *IEEE Transactions on Power Systems*, 8(3):1325–1332, Aug 1993.

[106] C.A. Roa-Sepulveda and B.J. Pavez-Lazo. A solution to the optimal power flow using simulated annealing. *International Journal of Electrical Power & Energy Systems*, 25(1):47 – 57, 2003.

[107] K.P. Wong and Y.W. Wong. Genetic and genetic/simulated-annealing approaches to economic dispatch. *Generation, Transmission and Distribution, IEE Proceedings-*, 141(5):507–513, Sep 1994.

[108] I.N. Da Silva, L. Nepomuceno, and T.M. Bastos. An efficient Hopfield network to solve economic dispatch problems with transmission system representation. *International Journal of Electrical Power & Energy Systems*, 26(9):733 – 738, 2004.

[109] J. Kumar and G.B. Sheble. Clamped state solution of artificial neural network for real-time economic dispatch. *IEEE Transactions on Power Systems*, 10(2):925–931, May 1995.

[110] T. Yalcinoz and M.J. Short. Neural networks approach for solving economic dispatch problem with transmission capacity constraints. *IEEE Transactions on Power Systems*, 13(2):307–313, May 1998.

[111] S. Khamsawang, C. Boonseng, and S. Pothiya. Solving the economic dispatch problem with tabu search algorithm. In *Industrial Technology, 2002. IEEE*

ICIT '02. 2002 IEEE International Conference on, volume 1, pages 274–278 vol.1, 2002.

[112] W. M. Lin, F. S. Cheng, and M. T. Tsay. An improved tabu search for economic dispatch with multiple minima. *IEEE Power Engineering Review*, 22(1):70–70, Jan 2002.

[113] L.D.S. Coelho and C.S. Lee. Solving economic load dispatch problems in power systems using chaotic and gaussian particle swarm optimization approaches. *International Journal of Electrical Power & Energy Systems*, 30(5):297 – 307, 2008.

[114] Z.L. Gaing. Particle swarm optimization to solving the economic dispatch considering the generator constraints. *IEEE Transactions on Power Systems*, 18(3):1187–1195, Aug 2003.

[115] J.B. Park, K.S. Lee, J.R. Shin, and K.Y. Lee. A particle swarm optimization for economic dispatch with nonsmooth cost functions. *IEEE Transactions on Power Systems*, 20(1):34–42, Feb 2005.

[116] T.A.A. Victoire and A.E. Jeyakumar. Hybrid psosqp for economic dispatch with valve-point effect. *Electric Power Systems Research*, 71(1):51 – 59, 2004.

[117] P. Fox-Penner. *Smart Power: Climate Change, the Smart Grid, and the Future of Electric Utilities*. Island Press, 2010.

[118] J. Momoh. *Smart Grid: Fundamentals of Design and Analysis*. Wiley, 2012.

[119] R. Mudumbai, S. Dasgupta, and B.B. Cho. Distributed control for optimal economic dispatch of a network of heterogeneous power generators. *IEEE Transactions on Power Systems*, 27(4):1750–1760, Nov 2012.

[120] Z. Zhang and M.Y. Chow. Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment. In *North American power symposium*, pages 4–6, 2011.

[121] Z. Zhang and M.Y. Chow. Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid. *IEEE Transactions on Power Systems*, 27(4):1761–1768, Nov 2012.

[122] Z. Zhang and M.Y. Chow. The influence of time delays on decentralized economic dispatch by using incremental cost consensus algorithm. In Aranya Chakrabortty and Marija Ilic, editors, *Control and Optimization Methods for Electric Smart Grids*, Power Electronics and Power Systems. Springer, 2012.

[123] M. Krebs and A. Shaheen. *Expander Families and Cayley Graphs: A Beginner's Guide*. Oxford University Press, USA, 2011.

[124] C.W. Wu. *Synchronization in coupled chaotic circuits and systems*. World Scientific, 2002.