공학박사학위논문

Design of Distributed Algorithms via Discrete-time Blended Dynamics Theorem

이산시간 혼합 동역학 정리를 통한 분산 알고리즘의 설계

2023년 2월

서울대학교 대학원

전기정보공학부

김 정 우

Design of Distributed Algorithms via Discrete-time Blended Dynamics Theorem

이산시간 혼합 동역학 정리를 통한 분산 알고리즘의 설계

지도교수 심 형 보

이 논문을 공학박사 학위논문으로 제출함.

2023년 2월

서울대학교 대학원

전기정보공학부

김 정 우

김 정 우의 공학박사 학위논문을 인준함.

2023년 2월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

ABSTRACT

Design of Distributed Algorithms via Discrete-time Blended Dynamics Theorem

by

Jeong Woo Kim

Department of Electrical and Computer Engineering
College of Engineering
Seoul National University

February 2023

In this thesis, a discrete-time version of the blended dynamics theorem is developed for the use of designing distributed computation algorithms. The blended dynamics theorem enables to predict the behavior of heterogeneous multi-agent systems. Therefore, once we get a blended dynamics for a particular computational task, design idea of node dynamics for individual heterogeneous agents can easily occur. In the continuous-time case, prediction by blended dynamics was enabled by high coupling gain among neighboring agents. In the discrete-time case, we propose an equivalent action, which we call multi-step coupling in this thesis. The discrete-time approach maintains the advantages of the continuous-time case, such as the plug-and-play operation, and that the individual node dynamics need not be stable as long as the blended dynamics is stable.

Compared to the continuous-time case, the discrete-time blended dynamics can have more variety depending on the coupling matrix. This benefit is

demonstrated with applications including distributed PageRank estimation problem where each node estimates its relative importance in a network which is possibly agent-wise different. In particular, the proposed algorithm does not require an initialization process, while most of other distributed PageRank algorithms have assumed.

Furthermore, the previous result of the discrete-time version of the blended dynamics theorem is extended to a heterogeneous multi-agent system under the limited exchange of state information (we call this as rank-deficient coupling). To achieve this, a coordinate change which separates the system into vanishing and non-vanishing dynamics as each agent repeats the coupling dynamics for multiple times is presented. Based on the non-vanishing part, the blended dynamics for the rank-deficient coupling is derived, which could predict the behavior of the heterogeneous network.

Again to emphasize the practical utility of the blended dynamics, the extended result is applied to a distributed state estimation problem where numerous sensors monitor a target plant with partial measurements. The distributed state observer is designed such that the local observer of each agent estimates the plant state on its detectable part while compensating the lacking information on undetectable space through the network communication. Even though the previous blended dynamics results only guarantee the approximate convergence, it is shown that the proposed observer guarantees the asymptotic performance.

Keywords: discrete-time heterogeneous multi-agent system, multi-step coupling, blended dynamics, multi-time scale, open multi-agent system
Student Number: 2015–22779

사랑하는 나의 가족과 친구들,
그리고 끊임없는 격려와 따뜻한 응원으로
도움을 주신 모든 분들께 본 논문을 바칩니다.

Contents

List of Figures

Symbols and Acronyms

$\mathbb{N}$	field of natural numbers
$\mathbb{Z}$	field of integers
$\mathbb{R}$	field of real numbers
$\mathbb{C}$	field of complex numbers
$\mathbb{R}^n$	Euclidean space of dimension n
$\mathbb{R}^{m\times n}$	space of $m \times n$ matrices with real entries
$\mathbb{R}_{>0}$	set of positive real numbers
$\mathbb{R}_{\geq 0}$	set of non-negative real numbers
$\mathrm{Re}\,(s)$	real part of a complex number s
$\mathrm{Im}\,(s)$	imaginary part of a complex number s
$\mathbb{1}_n$	$n \times 1$ column vector of all ones
I_n	$n \times n$ identity matrix
$\mathbb{0}_n$	$n \times 1$ column vector having all elements equal to 0
$O_{m\times n}$	$m \times n$ matrix having all elements equal to 0
A^{-1}	inverse of the nonsingular matrix A
A^T	transpose of the matrix A
$\mathrm{diag}(a_1, \ldots, a_k)$	diagonal matrix whose i-th diagonal is a_i
$\mathrm{diag}(A_1, \ldots, A_k)$	block diagonal matrix whose i-th block diagonal is A_i

$\mathrm{col}(a_1, \ldots, a_k)$	column vector whose i-th entry is a_i
$\mathrm{col}(A_1, \ldots, A_k)$	block column matrix whose i-th column block is A_i
$A \otimes B$	Kronecker product of matrices A and B
$\sum_{i=m}^{n} x_i$	summation of the sequence x_i; i.e., $x_m + x_{m+1} + \cdots + x_n$ if $m < n$, x_m if $m = n$, and 0 if $m > n$
$[a, b]$	interval of real numbers a and b; i.e., $\{x \in \mathbb{R} : a \leq x \leq b\}$
$\lvert X \rvert$	cardinal number of the set X
$\max\{a_1, \ldots, a_n\}$	maximum value among a_1, a_2, $\ldots$, a_n
$\min\{a_1, \ldots, a_n\}$	minimum value among a_1, a_2, $\ldots$, a_n
$\lvert x \rvert$	Absolute value of the scalar $x \in \mathbb{R}$
$\lVert x \rVert$	Euclidean norm of the vector $x \in \mathbb{R}^n$
$\lVert x \rVert_\infty$	maximum norm of the vector $x \in \mathbb{R}^n$
$\lVert A \rVert$	induced 2-norm of the matrix $A \in \mathbb{R}^{m \times n}$
$\lVert A \rVert_\infty$	induced maximum norm of the matrix $A \in \mathbb{R}^{m \times n}$
$\mathrm{im}A$	image space of the matrix $A \in \mathbb{R}^{m \times n}$
$\ker A$	kernel of the matrix $A \in \mathbb{R}^{m \times n}$
$\lambda_{\min}(A)$	the eigenvalue of A with the minimum real value
$\lambda_{\max}(A)$	the eigenvalue of A with the maximum real value
$\rho(A)$	spectral radius of A
$\mathrm{spec}(A)$	spectrum of A
$A > 0$ ($A \geq 0$)	symmetric matrix A is positive definite (positive semi-defnite)
$f : A \to B$	f is a function on the set A into the set B
$\mathfrak{C}^i$	i-th times continuously differentiable
$:=$	defined as
$\Rightarrow$	implies

$\forall$	for all
$\lozenge$	end of theorems, lemmas, propositions, assumptions, remarks, and so on
$\square$	end of proof

- For notational convenience, we use convention $A_{\otimes n} := A \otimes I_n$ for any size of matrix A.
- A square matrix A is said to be Hurwitz (matrix) if every eigenvalue λ of A has strictly negative real parts, i.e., $\mathrm{Re}(\lambda) < 0$.
- For any state variable $x(t)$, its initial condition will be denoted by $x(0)$.
- In order to avoid messy notation, the time symbol $\mathtt{t}$ or k is omitted when there in no confusion.
- For simplicity, we often use I_n, $\mathbb{0}_n$, and $O_{m\times n}$ without subscripts if their dimensions are obvious.

Acronyms

MAS	multi-agent system

Chapter 1

Introduction

1.1 Research Background

For a multi-agent system or networked system where multiple agents are interconnected through the network, a centralized algorithm to control the overall network or solve a specific computation problem which requires a global information might cause many problems including heavy load in the computation and communication of the centralized unit and difficulties in maintenance. Moreover, since the network has been expanded for various subjects with a recent advances in communication technologies, these problems become critical issue as the complexity and dimension of the network increases.

To overcome this issue, a distributed algorithm has been proposed, which has the following benefits:

- Computational burden of one node is lessened as the burden is distributed over many nodes in the network.
- Reliability against faults is improved as a fault on one node can be compensated by redundancy of many other nodes.
- Privacy of each node is preserved as the private information need not be transferred to a central node for computation.

Based on these benefits, many distributed algorithms have been actively studied with a wide range of application over the last decades. One of the major applications of the distributed algorithms are about synchronization where the state

or output information of every agent in a network is synchronized through the exchange of the information among neighboring agents. For example, the application of synchronization includes formation control [OPA15, Sak17], distributed state estimation [KLS19, MS18, MS16], and so on (further details on the synchronization can be found on [Wie10, RBA05]). Other examples of the distributed algorithms are distributed optimization [NO09, NL18] and distributed computation of PageRank [IT10, ITB12].

Nevertheless, constructive design methods for general distributed algorithms are not well developed yet, except the distributed optimization. In the distributed optimization, each agent minimizes a global cost function which is given by a sum of local cost functions. Initially, [NO09] combines a synchronization algorithm with the classical gradient descent method and this work was diversely extended with such as the fixed step size or gradient tracking method [SLWY15, YLY16, QL17]. Recently, the above studies were unified by [Jak18, AS20] which proposed a common framework to analyze different variations of the distributed algorithm and improved by utilizing an accelerated methods in [QL19]. However, most of the distributed optimization algorithms commonly require the convexity of each local cost function or a specific initialization.

One potential approach towards the constructive design is the *blended dynamics* approach [LS20], which is motivated by [KYS+16, PL17]. In the approach, the following multi-agent system is considered

$$\dot{\mathsf{x}}_i = \mathsf{f}_i(\mathsf{t}, \mathsf{x}_i) + \kappa \sum_{j \in \mathcal{N}_i} (\mathsf{x}_j - \mathsf{x}_i), \quad i \in \mathcal{N}, \tag{1.1.1}$$

where $\mathcal{N} := \{1, \cdots, N\}$ is the set of node indices, $\mathcal{N}_i$ is the index set of nodes that send information to node i, and $\mathsf{f}_i(\mathsf{t}, \mathsf{x}_i)$ represents the heterogeneity of each agent. Under the assumption that the communication graph is undirected and connected, if the blended dynamics is incrementally stable, then every agent in (1.1.1) behaves like the blended dynamics defined as

$$\dot{\mathsf{s}}(\mathsf{t}) = \frac{1}{N} \sum_{i=1}^{N} \mathsf{f}_i(\mathsf{t}, \mathsf{s}(\mathsf{t})) \tag{1.1.2}$$

	Homogeneous /Heterogeneous	Linear/Nonlinear	Leaderless	Constraints on MAS
Y. Yan et al., TCNS, 2020	Heterogeneous	Linear	O	Simultaneous stabilizable
D. Zhang et al., TC, 2018	Heterogeneous	Linear	X	-
X. Xu et al., TAC, 2017	Heterogeneous	Linear	O	Common internal model
S. Li et al., TC, 2015	Heterogeneous	Linear	O	Common internal model

Figure 1.1: Literature survey on synchronization of heterogeneous multi-agent system in discrete-time domain

for a sufficiently large coupling gain κ. More precisely, for any $\epsilon > 0$, there exists $\kappa^{\min}$ such that, if $\kappa > \kappa^{\min}$,

$$\limsup_{t\to\infty} |x_i(t) - s(t)| \leq \epsilon, \qquad \forall i \in \mathcal{N}. \tag{1.1.3}$$

Further details on the blended dynamics approach will be reviewed in Chapter 3.

As seen in (1.1.3), the blended dynamics approach provides a way to analyze the synchronized behavior of the heterogeneous multi-agent system. In fact, initial studies on the synchronization consider a homogeneous multi-agent system due to its simplicity, but the attention has soon shifted to the heterogeneous multi-agent system which could handle uncertainty, disturbance, noise, and agent-wise different input. Compared to the homogeneous case, achieving an asymptotic synchronization in the heterogeneous network is challenging unless every agent embeds a common internal model [KSS11, WSA11]. This difficulty naturally leads to studies on a practical synchronization (or approximate synchronization) as an alternative and one of the solution is suggested by the above blended dynamics approach.

In parallel to the studies in the continuous-time domain, discrete-time synchronization in a heterogeneous multi-agent system has also attracted attention due to the practical utility. For instance, the leader-following synchronization of a heterogeneous network is studied in [ZLF18] and an output synchronization problem of the heterogeneous system is addressed with uncertainty and disturbance [LFLG15] or switching network and time delay [XLF17] under the common in-

ternal model assumption. Similarly, an autonomous synchronization is proposed for a heterogeneous network of simultaneous stabilizable agents in [YCM20], such that the heterogeneity among the agent dynamics vanishes. However, all of the discrete-time studies mentioned so far handles only a linear system. Furthermore, most of the previous results on the synchronization of leaderless multi-agent system commonly assume a common internal model or simultaneous stabilizability as shown in Figure 1.1.

It should be emphasized that the blended dynamics is a simple average of the heterogeneous function of each agent. This could be usefully utilized as a design tool for many distributed algorithms, i.e., the designer firstly designs a desired algorithm as the blended dynamics (1.1.2), and then, splits it into f_i of (1.1.1). Indeed, this philosophy has been successfully employed in many applications such as distributed economic power dispatch problem [YSA19], distributed state estimator [KLS19], secure estimation by distributed median computation [LKS20], distributed least square solver [LS19], distributed optimization without convexity of each node [LS22], and decentralized controller design [KLS20]. See [LS21] for more comprehensive summary of these applications.

On the contrary, since all the above results are in the continuous-time domain, it is necessary to implement the designed algorithm in the discrete-time domain so that it could operate on digital devices in practice. Recalling that the arbitrarily small error in x_i and s in (1.1.3) can be guaranteed by increasing the coupling gain κ (i.e., strong coupling), this however prohibits using common discretization methods such as forward difference. This is because increasing κ unboundedly yields instability of the overall system unless its sampling time is decreased with the same ratio. As a result, this inspire the development of the discrete-time version of the blended dynamics approach which is main result of this dissertation.

1.2 Outline and Contributions of Dissertation

The following overview provides the outline of this thesis and briefly summarizes its contributions.

Chapter 2. Preliminaries

In this chapter, we review basic definitions of graph theory and related results. We also classify a non-negative matrix into special classes and present useful properties of each class. Based on these, by associating a non-negative matrix to the concept of directed graph, we characterize the matrix properties in the graph theoretical viewpoint. Finally, a definition of Kronecker product operation and its properties are provided which will be used in the dissertation. This chapter summarizes the followings:

- We review preliminary results from algebraic graph theory and matrix analysis which will be used in the dissertation.
- We provide a technical tool to handle more various communication protocols including the static diffusive type coupling for general directed graph.

Chapter 3. Behavior of Continuous-time Heterogeneous Multi-agent System under Strong Coupling

In this chapter, we review an initial studies on practical synchronization of heterogeneous multi-agent system by strong coupling. We define a virtual dynamics called blended dynamics whose solution could approximate an overall behavior of the system when the coupling gain among agents is sufficiently large. The contents of this chapter are contained on [KYS+16, KLLS22] and we provide the followings in this chapter:

- We recall an analysis tool for the synchronized behavior of the heterogeneous network under the strong coupling.
- We review that the practical synchronization is achieved even unstable agents are included in the network as long as the blended dynamics is assumed to be stable.
- We demonstrate the utility of the blended dynamics approach in the sense that the blended dynamics has been widely employed in designing various distributed algorithms.

- We explain that the implementation of the designed algorithms in the discrete-time domain is required to operate on digital devices in practice, but not trivial because the high-gain concept in the strong coupling causes an instability of the overall system.

Chapter 4. Behavior of Discrete-time Heterogeneous Multi-agent System under Multi-step Coupling

In this chapter, we develop a discrete-time version of the blended dynamics approach as a counter part of the results in the previous chapter. In particular, we propose a multi-step coupling concept where each agent repeats a weighted averaging action among its neighbors for many times before progressing through the heterogeneous individual node dynamics. The contents of this work are based in [KLLS22] and the main contributions of this work are listed as follows:

- We propose the multi-step coupling concept which corresponds to the strong coupling in the continuous-time blended dynamics approach.
- We define the discrete-time blended dynamics in a way the behavior of the multi-agent system can be predicted by its solution.
- We maintain the advantages of the continuous-time case including the plug-and-play operation and stability assumption of the blended dynamics, not individual node dynamics, in the discrete-time case.
- We demonstrate that the discrete-time approach can handle more diverse communication protocols comparing to the continuous-time case by providing three applications.

Chapter 5. Application to Initialization-free Distributed PageRank Estimation for Network of Web-pages

In this chapter, we apply the main result in the previous chapter to a distributed PageRank estimation problem. PageRank score represents the relative importance of each node in a network so that it has widely utilized in diverse areas. The contribution of this chapter is as follows:

- We show the benefit of the discrete-time approach that can have more variety than synchronization by proposing a distributed algorithm for each node to estimate its relative importance so that overall network is not synchronized.

- We propose an initialization-free distributed algorithm, while most of other distributed PageRank algorithms require the initialization process which is hard to be achieved in a distributed manner.

- The initialization-free benefit allows the algorithm to adopt plug-and-play feature, i.e., some nodes and their associate link can join or leave the network during the operation of the algorithm.

- We verify the advantage of the proposed scheme by simulation for real web-data of large scale.

Chapter 6. Behavior of Discrete-time Heterogeneous Multi-agent System under Rank-deficient Coupling

In this chapter, the result of Chapter 4 has been extended for rank-deficient coupling where only a partial information of communication affects the update of every agent. The main contributions of this work are:

- We derive a coordinate transformation to separate the system into vanishing and non-vanishing dynamics with respect to the repeating number of coupling dynamics.

- We introduce the blended dynamics based on the derived coordinate transformation, which approximates the overall behavior of the network under rank-deficient coupling.

Chapter 7. Application to Distributed State Estimation

In order to emphasize the practical utility of the result in Chapter 6, we propose a distributed state estimation algorithm based on the rank-deficient multi-step coupling framework. The distributed state estimation problem considers a

discrete-time linear time-invariant plant whose partial measurements are monitored by numerous sensors. It is supposed that no sensor can recover whole state information so that each sensor should compensate the lacking information through the network communication. The contribution of this work is:

- We formulate the distributed state estimation problem into multi-step coupling framework under rank-deficient coupling.
- We show that the proposed observer guarantees the asymptotic performance even though the result in the previous chapter only guarantees the approximate convergence.

Chapter 2

Preliminaries

2.1 Graph Theory

Graph theory has been widely utilized for modeling the interaction among agents in the network. In general, each agent corresponds to a node and the information flow from one agent to another agent is represented by an edge in the graph theory. In this section, we introduce the basic definitions and useful results in the graph theory, which will be used throughout the dissertation. For more details, refer to [GR01, Bul19].

2.1.1 Basic Definitions in Graph Theory

In this part, we introduce the weighted directed graph and other related concepts.

Definition 2.1.1. A weighted directed graph $\mathcal{G} := (\mathcal{N}, \mathcal{E})$ consisting of a nonempty finite set of nodes $\mathcal{N} := \{1, \dots, N\}$, an edge set of ordered pairs of nodes $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$, and a weighted adjacency matrix $\mathcal{A} \in \mathbb{R}^{N \times N}$ whose element α_{ij} is positive if an edge (j, i) is contained in $\mathcal{E}$ and $\alpha_{ij} = 0$ otherwise. We exclude the multiple edges, but consider the self-loop in the graph. $\Diamond$

Using the previous definition, the interconnection between agents is represented as a graph as the following convention: if the j-th agent sends information to the i-th agent or equivalently the i-th agent receives information from the j-th

agent, then the information flow is modeled as an edge $(j, i) \in \mathcal{E}$. From this, the following special classes of graphs can be derived.

Definition 2.1.2. We define the following special graphs.

1. A graph $\mathcal{G}$ is *unweighted* if $\alpha_{ij} \in \{0, 1\}$ for all $i, j \in \mathcal{N}$. In this case, its adjacency matrix is denoted as *binary adjacency matrix* $A \in \mathbb{R}^{N \times N}$.

2. A graph $\mathcal{G}$ is *undirected* if $(j, i) \in \mathcal{E}$ implies $(i, j) \in \mathcal{E}$. $\Diamond$

2.1.2 Connectivity and Periodicity of the Graph

The concepts of connectivity and periodicity of a graph are presented in this part. Before introducing them, we first define two types of neighbors for each node as follows.

Definition 2.1.3. Consider a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$.

1. *In-neighbors* of the i-th node is a set defined as $\mathcal{N}_i := \{j \in \mathcal{N} \mid (j, i) \in \mathcal{E}\}$ and *out-neighbors* of the i-th node is a set defined as $\mathcal{N}_i^{\text{out}} := \{j \in \mathcal{N} \mid (i, j) \in \mathcal{E}\}$.

2. *In-degree* and *out-degree* of the i-th node is defined as $d_i := |\mathcal{N}_i|$ and $d_i^{\text{out}} := |\mathcal{N}_i^{\text{out}}|$, respectively. $\Diamond$

From the convention of the previous part, every agent which sends information to the i-th agent is included in $\mathcal{N}_i$ and every agent which receives information from the i-th agent is included in $\mathcal{N}_i^{\text{out}}$. Thus, d_i and d_i^{out} represent the number of agents which sends information to the i-th agent and the number of agents which receives information from the i-th agent, respectively. The sequence of neighboring pair of nodes yields the concept of the directed path.

Definition 2.1.4. A *directed path* from the i-th node to the j-th node is a sequence $(i_0, i_1, \ldots, i_L)$ such that $i_0 = i$, $i_L = j$, $(i_l, i_{l+1}) \in \mathcal{E}$ for any $l \in \{0, \ldots, L-1\}$, and every i_l is distinct. Here, the length of the directed path is L. $\Diamond$

Through the directed path from the i-th node to the j-th node, an information of the i-th agent can be propagated to the j-th agent.

Now, we can define the connectivity of the graph using the definition of the directed path as follows.

Definition 2.1.5. A directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{A})$ is said to be *strongly connected* if there exists a direct path from any node to any other node. When $\mathcal{G}$ is undirected, we simply say *connected*. ◊

Meanwhile, we introduce a special directed path which starts and ends at the same node as follows.

Definition 2.1.6. A *cycle* is a directed path that starts and ends at the same node and any node does not appear more than once in it. ◊

Based on the definition of the directed path, a cycle of length L can be represented by $(i_0, i_1, \ldots, i_{L-1}, i_0)$ for some node i_0.

From the concept of the cycle, the following special class of the strongly connected graph is of interest.

Definition 2.1.7. A strongly connected directed graph is said to be *periodic* if there exists $\mathfrak{P} > 1$, called the *period*, that divides the length of every cycle in the graph, otherwise the graph is said to be *aperiodic*. ◊

If a strongly connected directed graph includes at least one self-loop, i.e., there exist $i \in \mathcal{N}$ such that $(i, i) \in \mathcal{E}$, then the graph is aperiodic because any self-loop is the cycle whose length is one.

2.1.3 Laplacian Matrix and Its Properties

The Laplacian matrix of a directed graph is defined as follows.

Definition 2.1.8. Given a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E}, \mathcal{A})$, the Laplacian matrix $\mathcal{L} := [l_{ij}] \in \mathbb{R}^{N \times N}$ is defined as

$$l_{ij} = \begin{cases} -\alpha_{ij}, & i \neq j, \\ \sum_{j=1, j \neq i}^{N} \alpha_{ij}, & i = j. \end{cases}$$ ◊

By its definition, every row-sum of the Laplacian matrix $\mathcal{L}$ is 0, i.e., $\mathcal{L}\mathbb{1}_N = \mathbb{0}_N$, so that the Laplacian matrix has a zero eigenvalue with corresponding eigenvector of $\mathbb{1}_N$. Let $\lambda_i(\mathcal{L})$ for $i \in \mathcal{N}$ be the eigenvalues of the Laplacian matrix $\mathcal{L}$, then it is well-known that all the eigenvalues of $\mathcal{L}$ has negative real part as stated below.

Lemma 2.1.1. All the eigenvalues of the Laplacian matrix $\mathcal{L}$ lie on the closed right-half complex plane, i.e., $\mathrm{Re}\,(\lambda_i(\mathcal{L})) \geq 0$ for all $i \in \mathcal{N}$. ◊

Proof. Proof directly follows by applying Geršgorin disc theorem [HJ19, Theorem 6.1.1]. □

The strong connectivity of a directed graph can be characterized by the following lemma [MRC15, Lemma 2.1].

Lemma 2.1.2. Let $\mathcal{G}$ be a strongly connected directed graph with N nodes and $\mathcal{L}$ be the associated Laplacian matrix. Then, the following statements hold.

1. The Laplacian matrix $\mathcal{L}$ has a simple zero eigenvalue and all other eigenvalues have positive real parts.
2. There exists a positive vector $\Theta = [\Theta_1, \cdots, \Theta_N]^\top$, i.e., $\Theta_i > 0$ for all $i = 1, \ldots, N$ such that $\sum_{i=1}^{N} \Theta_i = 1$ and $\Theta^\top \mathcal{L} = 0$ hold. ◊

2.2 Matrix Analysis

This section presents special classes of matrix as well as their useful properties. Based on this, we associate a matrix to weighted directed graph and characterize the matrix properties with the graph's connectivity and related concepts through the sufficient and necessary conditions. More details can be found in [Bul19].

2.2.1 Stochastic Matrix

In this part, we are interested in the special classes of a non-negative square matrix, i.e., every component of the matrix is non-negative.

Definition 2.2.1. Consider a non-negative square matrix $M \in \mathbb{R}^{N \times N}$.

1. M is *row-stochastic* if $M\mathbb{1}_N = \mathbb{1}_N$.
2. M is *column-stochastic* if $M^\top \mathbb{1}_N = \mathbb{1}_N$.
3. M is *doubly-stochastic* if it is row-stochastic and column-stochastic. $\Diamond$

By its construction, a row-stochastic matrix $M \in \mathbb{R}^{N\times N}$ has an eigenvalue of 1 with a right eigenvector of $\mathbb{1}_N$. Similarly, if M is a column-stochastic matrix, it has an eigenvalue of 1 with a left eigenvector of $\mathbb{1}_N$ and, for the doubly-stochastic matrix, $\mathbb{1}_N$ is both right and left eigenvector for the eigenvalue 1. It is well known that the spectrum and the spectral radius of any stochastic matrix hold the following lemma by adopting the Lemma 2.9 of [Bul19].

Lemma 2.2.1. Consider any stochastic matrix $M \in \mathbb{R}^{N\times N}$.

1. The spectrum of M is contained in the unit disk centered at the origin in the complex plane, i.e., $\text{spec}(M) \subset \{z \in \mathbb{C} : \|z\| \leq 1\}$.
2. The spectral radius of M is 1, i.e., $\rho(M) = 1$. $\Diamond$

2.2.2 Irreducible and Primitive Matrix

In this part, we have interest in two groups of non-negative matrices as follows.

Definition 2.2.2. Consider a non-negative matrix $M \in \mathbb{R}^{N\times N}$.

1. M is *irreducible* if $\sum_{n=0}^{N-1} M^n$ is positive matrix, i.e., every component of $\sum_{n=0}^{N-1} M^n$ is positive.
2. M is *primitive* if there exist $n \in \mathbb{N}$ such that M^n is positive matrix. $\Diamond$

From the definitions, it is clear that any primitive matrix is irreducible.

The following result, which is well known as Perron-Frobenius theorem, characterizes the properties of the spectral radius and its corresponding eigenvectors of a non-negative matrix based on the aforementioned classification. The proof can be found in [Bul19, Theorem 2.12].

Lemma 2.2.2. For a non-negative matrix $M \in \mathbb{R}^{N\times N}$, the following statements hold.

1. There exist a real eigenvalue $\lambda \geq |\sigma| \geq 0$ for any other eigenvalue σ of M,
2. The right and left eigenvectors of λ can be selected non-negative.

If M is irreducible, the following statements hold.

1. The eigenvalue λ is strictly positive and simple,
2. The right and left eigenvectors of λ are positive.

If M is primitive, the following statement holds.

1. The eigenvalue λ satisfies $\lambda > |\sigma|$ for any other eigenvalue σ of M. ◊

Here, the eigenvalue λ of M is denoted as *Perron-Frobenius eigenvalue*, *dominant eigenvalue*, or simply *Perron root*.

2.2.3 Graph Theoretical Characterization

In this part, we introduce some results which relate the aforementioned classes of a non-negative matrix with the graph theoretical characterizations.

A directed graph can be associated from a non-negative square matrix as stated below.

Definition 2.2.3. Given a non-negative square matrix $M = [m_{ij}] \in \mathbb{R}^{N\times N}$, the associated graph $\mathcal{G}_M$ of M is defined as the weighted directed graph whose adjacency matrix is M, i.e., $\mathcal{G}_M = (\mathcal{N}_M, \mathcal{E}_M, \mathcal{A}_M)$ such that $\mathcal{N}_M = \{1, \cdots, N\}$, $\mathcal{A}_M = M$, and $(j, i) \in \mathcal{E}_\mathcal{M}$ if $m_{ij} > 0$. ◊

Using this concept, the irreducible or primitive property of a non-negative square matrix can be verified by the following results [Bul19, Theorem 4.3, Theorem 4.7].

Lemma 2.2.3. Given a non-negative square matrix M and its associated weighted graph $\mathcal{G}_M$, the following statements are equivalent:

1. M is irreducible.
2. $\mathcal{G}_M$ is strongly connected. ◊

Lemma 2.2.4. Given a non-negative square matrix M and its associated weighted graph $\mathcal{G}_M$, the following statements are equivalent:

1. M is primitive.
2. $\mathcal{G}_M$ is strongly connected and aperiodic. $\Diamond$

2.3 Kronecker Product

Kronecker product is a useful tool for representing an overall networked system. In this section, the basic definition of the Kronecker product and useful properties are discussed in this section.

Basic definition of the Kronecker product is introduced as follows.

Definition 2.3.1. For $A = [a_{ij}] \in \mathbb{R}^{m\times n}$ and $B = [b_{ij}] \in \mathbb{R}^{p\times q}$, the Kronecker product $A \otimes B$ is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{pm\times qn}. \qquad \Diamond$$

Through the dissertation, the following properties of the Kronecker product will be used. More general properties for the Kronecker product can be found in [Ber09].

Proposition 2.3.1. Consider real matrices A, B, C and D of appropriate dimensions for each item.

1. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
2. $\|A \otimes B\| = \|A\|\|B\|$
3. Let $\lambda_1(A), \ldots, \lambda_n(A)$ be the eigenvalues of $A \in \mathbb{R}^{n\times n}$ and $\lambda_1(B), \ldots, \lambda_m(B)$ be the eigenvalues of $B \in \mathbb{R}^{m\times m}$, then the eigenvalues of $A \otimes B$ are $\lambda_i(A)\lambda_j(B)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. $\Diamond$

Chapter 3

Behavior of Continuous-time Heterogeneous Multi-agent System under Strong Coupling

3.1 Problem Formulation

In this chapter, an initial study on synchronization of heterogeneous multi-agent systems is introduced. Unlike homogeneous multi-agent systems, achieving an asymptotic synchronization in the heterogeneous multi-agent systems is difficult unless a common internal model is embedded in each heterogeneous agent [WSA11, KSS11]. This challenge naturally lead to studies on a practical synchronization (or approximate synchronization) as an alternative.

One potential approach to achieve the practical synchronization is utilizing a strong coupling such as [KYS+16]. This paper considers a continuous-time heterogeneous multi-agent system whose individual dynamics is given as

$$\dot{x}_i = f_i(t, x_i) + u_i, \quad i \in \mathcal{N} := \{1, \dots, N\} \tag{3.1.1}$$

where $x_i \in \mathbb{R}$ is the state, $u_i \in \mathbb{R}$ is the input, t is the continuous-time index, N is the number of agents in the network, and $f_i : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ is continuously differentiable with respect to x_i, globally Lipschitz with respect to x_i uniformly in t, and uniformly bounded in t. Note that the heterogeneous vector field f_i could include parametric variations or uncertainties such as disturbances and noises as

well as an external inputs which are possibly agent-wise different. Suppose that all agents are interconnected by static diffusive-type coupling

$$\mathsf{u}_i = \kappa \sum_{j \in \mathcal{N}_i} \alpha_{ij} (\mathsf{x}_j - \mathsf{x}_i) \tag{3.1.2}$$

where κ represents the coupling gain, $\mathcal{N}_i$ is the in-neighbors of the i-th agent, and α_{ij} is the (i,j)-th entry of the adjacency matrix of the given network.

In the paper, the *blended dynamics* is proposed in order to handle the heterogeneity of the network, whose dynamics is an average of the vector fields of all agents as follows:

$$\dot{\mathsf{s}}(\mathsf{t}) = \frac{1}{N} \sum_{i=1}^{N} \mathsf{f}_i(\mathsf{t}, \mathsf{s}(\mathsf{t})) =: \bar{\mathsf{f}}(\mathsf{t}, \mathsf{s}(\mathsf{t})). \tag{3.1.3}$$

It will be reviewed in Section 3.2 that, as long as the blended dynamics (3.1.3) is stable, every solution $\mathsf{x}_i(t)$ satisfies

$$\limsup_{\mathsf{t} \to \infty} |\mathsf{x}_i(\mathsf{t}) - \mathsf{s}(\mathsf{t})| \leq \beta\left(\frac{1}{\kappa}\right), \quad \forall \kappa > \kappa^{\min}, \quad \forall i \in \mathcal{N} \tag{3.1.4}$$

where $\kappa^{\min}$ is a threshold and β is a class $\mathcal{K}$ function[1]. It should be emphasized that the stability of the blended dynamics does not necessarily implies the stability of each agents. Rather it allows unstable agents if their instabilities are compensated by other stable agents so that overall average is stable.

Recently, this blended dynamics approach is extended to the multi-agent systems whose coupling matrices are possibly all different and singular [LS20]. For the related discussion, we refer to Chapter 6.

3.2 Synchronization of Multi-agent System due to Strong Coupling

In this section, we assume the following for heterogeneous function f_i.

[1] A continuous function $\beta : [0, b) \to [0, \infty)$ for some positive constant b is class $\mathcal{K}$ function if it is strictly increasing and $\beta(0) = 0$ [Kha02].

Assumption 3.2.1. There exist a constant $\mathsf{L} > 0$ and a non-decreasing continuous function $\mathsf{M} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that, $\forall \mathsf{x}_i \in \mathbb{R}, \mathsf{t} \geq 0, i \in \mathcal{N}$,

$$\left| \frac{\partial \mathsf{f}_i}{\partial \mathsf{x}_i}(\mathsf{t}, \mathsf{x}_i) \right| \leq \mathsf{L}, \quad |\mathsf{f}_i(\mathsf{t}, \mathsf{x}_i)| \leq \mathsf{M}(|\mathsf{x}_i|). \qquad \Diamond$$

In addition, we assume the network connectivity as follows.

Assumption 3.2.2. The communication network induced by the adjacency element α_{ij} is undirected and connected. $\Diamond$

Now, the stability of the blended dynamics (3.1.3) is assumed in the following way.

Assumption 3.2.3. There exists a constant $\mathsf{p} > 0$ such that, for all $\mathsf{s} \in \mathbb{R}$ and $\mathsf{t} \geq 0$,

$$\frac{\partial \bar{\mathsf{f}}}{\partial \mathsf{s}}(\mathsf{t}, \mathsf{s}) = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial \mathsf{f}_i}{\partial \mathsf{x}_i}(\mathsf{t}, \mathsf{s}) \leq -\mathsf{p}. \qquad \Diamond$$

From the stability of the blended dynamics, the ultimate boundedness of its solution $\mathsf{s}(\mathsf{t})$ can be shown as the following lemma.

Lemma 3.2.1. [KYS+16, Lemma 1] For a scalar system $\dot{\mathsf{s}} = \mathsf{F}(\mathsf{t}, \mathsf{s})$ with $\mathfrak{C}^1$ function F satisfying $(\partial \mathsf{F})/(\partial \mathsf{s}) \leq -\mathsf{p} < 0$ for all s and $\mathsf{t} \geq 0$,

$$\limsup_{\mathsf{t} \to \infty} |\mathsf{s}(\mathsf{t})| \leq \frac{\limsup_{\mathsf{t} \to \infty} |F(\mathsf{t}, 0)|}{\mathsf{p}}. \qquad \Diamond$$

Let $\mathsf{x} := \mathrm{col}\,(\mathsf{x}_1, \cdots, \mathsf{x}_N)$, then the dynamics of the overall system (3.1.1) and (3.1.2) is written by

$$\dot{\mathsf{x}} = -\kappa \mathcal{L} \mathsf{x} + \mathsf{f}(\mathsf{t}, \mathsf{x})$$

where $\mathcal{L}$ is the Laplacian matrix and $\mathsf{f}(\mathsf{t}, \mathsf{x}) := \mathrm{col}\,(\mathsf{f}_1(\mathsf{t}, \mathsf{x}_1), \cdots, \mathsf{f}_N(\mathsf{t}, \mathsf{x}_N))$. Since the Laplacian $\mathcal{L}$ is symmetric, all the eigenvalues are real and the zero eigenvalue of $\mathcal{L}$ is simple by Lemma 2.1.2. Thus, without loss of generality, let $\lambda_i(\mathcal{L})$ for $i \in \mathcal{N}$ be the eigenvalues of $\mathcal{L}$ such that $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \leq \ldots \leq \lambda_N(\mathcal{L})$.

Now, the following proposition shows that the strong coupling among agents makes the trajectories of all the agents remain in an arbitrarily small neighborhood of the trajectory of the blended dynamics.

Proposition 3.2.2. [KYS+16, Theorem 1] Under Assumptions 3.2.1, 3.2.2 and 3.2.3, there exists a class $\mathcal{K}$ function Γ such that the solutions of the overall system, composed of (3.1.1) and (3.1.2), with arbitrary initial conditions, and the solution $\mathsf{s}(\mathsf{t})$ to the blended dynamics (3.1.3) with $\mathsf{s}(0) = \sum_{i=1}^{N} \mathsf{x}_i(0)/N$ satisfy

$$\limsup_{\mathsf{t}\to\infty} |\mathsf{x}_i(\mathsf{t}) - \mathsf{s}(\mathsf{t})| \leq \Gamma\left(\frac{1}{\kappa\lambda_2(\mathcal{L}) - \mathsf{L}}\right), \quad \forall \kappa > \kappa^{\min} \tag{3.2.1}$$

for all $i = 1, \ldots, N$, where

$$\kappa^{\min} = \frac{3\mathsf{L}^2}{\mathsf{p}\lambda_2(\mathcal{L})} + \frac{\mathsf{L}}{\lambda_2(\mathcal{L})}.$$

In particular, the function Γ is defined on $[0, \mathsf{p}/(3\mathsf{L}^2))$ and given by

$$\Gamma(\chi) = \mathsf{M}\left(\frac{\mathsf{M}(0)}{\mathsf{p}}\right)\sqrt{N}\sqrt{\mathsf{r}(\chi)} \tag{3.2.2}$$

in which

$$\mathsf{r}(\chi) = \begin{cases} 0, & \text{if } \chi = 0 \\ \frac{4\chi}{\mathsf{p}-3\mathsf{L}^2\chi}, & \text{if } 0 < \chi \leq \frac{4\mathsf{p}}{\mathsf{p}^2+20\mathsf{L}^2} \\ \frac{(\mathsf{p}^2+8\mathsf{L}^2)\chi^2}{(\mathsf{p}-3\mathsf{L}^2\chi)^2}, & \text{if } \frac{4\mathsf{p}}{\mathsf{p}^2+20\mathsf{L}^2} < \chi < \frac{\mathsf{p}}{3\mathsf{L}^2}. \end{cases} \quad \lozenge$$

Remark 3.2.1. [KYS+16, Remark 2] It should be noted that the function Γ and the value of $\kappa^{\min}$ are affected by the number N. The former is obvious due to the appearance of $\sqrt{N}$ in (3.2.2), but the latter is indirect through the value of $\lambda_2(\mathcal{L})$. The second smallest eigenvalue $\lambda_2(\mathcal{L})$ of the Laplacian matrix is called the algebraic connectivity (or, density) of a graph $\mathcal{G}$, which indicates how well connected the graph is. It depends both on the topology of the graph and the number N of the nodes. For the all- to-all network (with unit weights), $\lambda_2(\mathcal{L})$ is the same as the number N, but for the ring network, increasing N decreases

$\lambda_2(\mathcal{L})$ (because $\lambda_2(\mathcal{L}) = 2(1 - \cos(2\pi/N))$). We refer to [Fie73] for more details on their relation. ◊

Remark 3.2.2. [KYS+16, Remark 3] Remark 3.2.1 implies that, in order to maintain the same level of error while the number N increases, the coupling strength κ may need to be increased. In fact, the ultimate error bound is given in (3.2.1) and it can be seen that

$$\Gamma\left(\frac{1}{\kappa\lambda_2(\mathcal{L}) - \mathsf{L}}\right) = \mathcal{O}\left(\sqrt{N/\kappa\lambda_2(\mathcal{L})}\right)^{2}$$

when κ is large enough (so that χ is small enough). Therefore, for the ring network where $\lambda_2(\mathcal{L}) = 2(1 - \cos(2\pi/N))$, we have

$$\Gamma\left(\frac{1}{\kappa\lambda_2(\mathcal{L}) - \mathsf{L}}\right) = \mathcal{O}\left(\sqrt{N^3/\kappa}\right),$$

and for the same level of error, κ should be increased when N is increased. On the other hand, for the all-to-all network, the error bound is given by

$$\Gamma\left(\frac{1}{\kappa\lambda_2(\mathcal{L}) - \mathsf{L}}\right) = \mathcal{O}\left(\sqrt{1/\kappa}\right)$$

since $\lambda_2(\mathcal{L}) = N$. ◊

3.3 Utility of the Blended Dynamics Theory

Since the blended dynamics is a simple average of individual node dynamics, it has been utilized as a design tool for many distributed algorithms; that is, one designs a desired algorithm as the blended dynamics (3.1.3) first, and then, splits it into different node dynamics (3.1.1) with the diffusive coupling (3.1.2).

This philosophy has been successfully employed in many applications such as distributed economic power dispatch problem [YSA19], distributed state estimator [KLS19], secure estimation by distributed median computation [LKS20], distributed least square solver [LS19], distributed optimization without convexity

[2]For real valued function f and g, $f(x) = \mathcal{O}(g(x))$ if there exists a positive real number M and a real number x_0 such that $|f(x)| \leq Mg(x)$ for all $x \geq x_0$ [Bla08].

of each node [LS22], and decentralized controller design [KLS20]. See [LS21] for more comprehensive summary of these applications.

The distributed algorithms designed by the blended dynamics theorem does not require each node dynamics to be stable, as long as their average (i.e., the blended dynamics) is contractive[3], which yields flexibility of the design. Moreover, as long as the blended dynamics remains contractive, a new node can join the network or an existing node can leave the network during the operation, which is called as a plug-and-play feature. This is because the designed distributed algorithms are initialization-free [LS20].

3.4 Necessity of Discrete-time Blended Dynamics Theory

While all the above results are in the continuous-time domain, it is however required to implement the designed algorithm in the discrete-time domain so that it operates on digital devices in practice. A naive idea is to use simple discretization methods such as forward difference. For example, a discretized model of the overall dynamics composed of (3.1.1) and (3.1.2) becomes

$$\mathsf{x}_i(\mathsf{t}+\Delta_\mathsf{t}) = \mathsf{x}_i(\mathsf{t}) + \Delta_\mathsf{t}\mathsf{f}_i(\mathsf{t},\mathsf{x}_i(\mathsf{t})) + \kappa\Delta_\mathsf{t}\sum_{j\in\mathcal{N}_i}(\mathsf{x}_j(\mathsf{t})-\mathsf{x}_i(\mathsf{t})), \tag{3.4.1}$$

where Δ_t is the sampling time.

In the continuous-time case, we recall that arbitrarily small error between x_i and s in (3.1.4) can be obtained by increasing the coupling gain κ, so that an emergent behavior of the multi-agent system arises with strong coupling. However, in the discrete-time case of (3.4.1), increasing κ unboundedly yields instability of the network unless Δ_t is decreased with the same ratio. One can verify it with the following example.

[3] A system $\dot{x} = f(t,x)$ is contractive if there exist a positive definite matrix H_c and a positive constant λ_c such that $H_c(\partial f/\partial x)(t,x) + (\partial f/\partial x)^\top(t,x)H_c \leq -\lambda_c H_c$ for all $x \in \mathbb{R}^m$ and $t \geq 0$ [PWN05].

Let the heterogeneous function $\mathsf{f}_i(\mathsf{t}, \mathsf{x}_i) = -\mathsf{x}_i$, then (3.4.1) is rewritten as

$$\mathsf{x}(\mathsf{t} + \Delta_\mathsf{t}) = \{(1 - \Delta_\mathsf{t})I - \kappa\Delta_\mathsf{t}\mathcal{L}\}\,\mathsf{x}(t)$$

where $\mathsf{x} = \mathrm{col}\,(\mathsf{x}_1, \cdots, \mathsf{x}_N)$ and $\mathcal{L}$ is the Laplacian matrix of a connected graph. With κ sufficiently large, some eigenvalues of the system matrix lie outside of the unit circle unless $\kappa\Delta_\mathsf{t}$ remains small. Therefore, the discrete-time algorithm (3.4.1) cannot be a discrete-time version of the blended dynamics approach.

Chapter 4

Behavior of Discrete-time Heterogeneous Multi-agent System under Multi-step Coupling

4.1 Problem Formulation

This chapter proposes a new form of a multi-agent system, which is given by (4.2.1) in Section 4.2. We note that the meaning of using a large coupling gain κ in the continuous-time case of (3.1.2) is that synchronization is taken more care of than the progress through the node dynamics. Based on the observation, and motivated by [WLMA19], the proposed form repeats a weighted averaging action many times before progressing through the node dynamics, which we call 'multi-step coupling.'

This approach maintains the advantages of the continuous-time case, such as the plug-and-play operation, and that the individual node dynamics need not be stable as long as the blended dynamics is stable, which we do not repeat in this paper but refer the reader to [KYS$^+$16, LS20, KLS20].

Moreover, while the continuous-time approach predicted collective synchronization behavior of the multi-agent system in [KYS$^+$16, LS20], this discrete-time approach estimates not only emergent but also individually scaled behavior, i.e., each agent behaves similarly to the solution of the blended dynamics with an agent-wise scaling factor. For example, in Chapter 5, we will introduce an application example where each node estimates its relative importance which is

possibly agent-wise different so that overall nodes are not synchronized in the network.

This chapter is organized as follows. In Section 4.2, the behavior of discrete-time heterogeneous multi-agent system under the multi-step coupling is studied. Section 4.3 is devoted to demonstrate the utility of the proposed approach. In particular, we specifies some special couplings which are frequently utilized in other studies, analyzes a behavior of each multi-step coupling system using the results in Section 4.2, and introduces practical examples of application.

4.2 Prediction on Emergent Behavior under Multi-step Coupling

For a discrete-time version of the blended dynamics theorem, we propose the following discrete-time algorithm: for each agent $i \in \mathcal{N}$,

$$x_i[t_{k+1}] = \begin{cases} f_i(t_k, x_i[t_k]), & \text{if } k = 0, \quad (4.2.1a) \\ \displaystyle\sum_{j \in \mathcal{N}_i \cup \{i\}} w_{ij} x_j[t_k], & \text{if } k = 1, \dots, K-1, \quad (4.2.1b) \end{cases}$$

where $x_i \in \mathbb{R}^n$ is the state, the function $f_i : \mathbb{Z} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and represents the time-varying heterogeneous node dynamics (4.2.1a), and the coefficient w_{ij}, called *coupling weight*, determines the behavior of the coupling dynamics (4.2.1b). Here, t_k is the symbol defined by

$$t_k = t + \frac{k}{K}, \quad (4.2.2)$$

where $K \in \mathbb{N}$, and we call t_k by *fractional discrete-time index*. In particular, we call $t \in \mathbb{Z}$ as *integer count* and $k \in \mathbb{N}$ as *fraction count*. The fraction count k varies from 0 to $K-1$. Keeping in mind that $t_K = t + K/K = (t+1) + 0/K = (t+1)_0$, we see that the fractional discrete-time t_k advances as $0_0, 0_1, \cdots, 0_{K-1}, 1_0, 1_1, \cdots$. The time t_0 will often be written as t for convenience. The fractional discrete-time has nothing to do with real time, and can be implemented in practice just as a sequential order in an algorithm.

We will choose K sufficiently large, which determines how many times the

coupling dynamics (4.2.1b) is executed before the next node dynamics (4.2.1a) is executed. It will be shown that, in this way, the effect of *strong coupling* κ in the continuous-time blended dynamics theorem can be similarly reflected in discrete-time. To emphasize the difference, we call this type of coupling in (4.2.1) by *multi-step coupling*.

The coupling weights w_{ij} in (4.2.1b) have the property:

$$w_{ij} \begin{cases} > 0, & j \in \mathcal{N}_i \cup \{i\}, \\ = 0, & \text{otherwise.} \end{cases} \tag{4.2.3}$$

Now, we assume that the matrix $W := [w_{ij}] \in \mathbb{R}^{N \times N}$, which we call a *weight matrix*, satisfies the following.

Assumption 4.2.1. The spectral radius of W is 1. ◊

Note that the communication protocols in many discrete-time multi-agent systems in the literature have the form of linear combination like in (4.2.1b) and their weight matrices satisfy Assumption 4.2.1. Examples include [IT10, ITB12, LC14, OSFM07, RB05], in which the weight matrices are given by stochastic matrices whose spectral radius is 1.

Meanwhile, the communication network under consideration is represented by the directed graph $\mathcal{G}$, which does not have self-connection at any node by definition, and we assume the following.

Assumption 4.2.2. The network $\mathcal{G}$ is strongly connected. ◊

Then, under Assumptions 4.2.1 and 4.2.2, the following is well-known (but we put its proof for readers' convenience).

Lemma 4.2.1. Let $\lambda_i(W)$, $i \in \mathcal{N}$, be the eigenvalues of W such that $|\lambda_1(W)| \geq |\lambda_2(W)| \geq \ldots \geq |\lambda_N(W)|$. Under Assumptions 4.2.1 and 4.2.2, $\lambda_1 = 1$, $\lambda_1(W) > |\lambda_j(W)|$ for all $j = 2, \ldots, N$, and there exist positive vectors $p, q \in \mathbb{R}^N$ such that

$$Wp = p, \quad q^\top W = q^\top, \quad q^\top p = 1. \tag{4.2.4}$$

◊

Proof. From (4.2.3), the associated graph of W not only contains all edges of $\mathcal{G}$ but also has a self-connection for every node because all diagonal entries of W are positive. Thus, the associated graph is aperiodic as well as strongly connected. This implies that W is primitive by Lemma 2.2.4, and, by Assumption 4.2.1 and Lemma 2.2.2, W has the simple Perron-Frobenius eigenvalue 1, i.e., $\lambda_1(W) = 1$ and $\lambda_1(W) > |\lambda_j(W)|$ for all $j = 2, \dots, N$, with positive right and left eigenvectors p and q, respectively. Scaling p and q yields that $q^\top p = 1$. □

With p and q from Lemma 4.2.1 at hand, we now introduce *discrete-time blended dynamics*, which is defined as a *weighted* average of node dynamics:

$$s[t+1] = \sum_{i=1}^{N} q_i f_i(t, p_i s[t]) =: f_s(t, s[t]) \quad \in \mathbb{R}^n \tag{4.2.5}$$

where t is the integer count of the fractional time (i.e., $t = t_0$). In particular, we assume the blended dynamics is stable in the sense of [LS98, TRK18] as follows.

Assumption 4.2.3. The blended dynamics (4.2.5) is *contractive*; i.e., there exist a (symmetric) positive definite matrix $H \in \mathbb{R}^{n\times n}$ and a positive constant $\gamma < 1$ such that

$$\frac{\partial f_s}{\partial s}(t,s)^\top H^2 \frac{\partial f_s}{\partial s}(t,s) \le \gamma H^2, \quad \forall s \in \mathbb{R}^n, t \in \mathbb{Z}.$$

Remark 4.2.1. Assumption 4.2.3 does not ask each node dynamics $x_i[t+1] = f_i(t, x_i[t])$ to be stable. Rather it allows unstable node dynamics whose instability can be compensated by other node dynamics so that the blended dynamics becomes stable in the sense of Assumption 4.2.3. For example, when there are four agents with $f_1(t,x) = f_2(t,x) = 0.1x$ and $f_3(t,x) = f_4(t,x) = 1.5x$, the agents 1 and 2 have stable node dynamics while the agents 3 and 4 have unstable ones. If the weight matrix has the vectors $p = \mathbb{1}_4$ and $q = \mathbb{1}_4/4$, then Assumption 4.2.3 holds because $f_s(s) = 0.8s$. ♢

We will see that the blended dynamics (4.2.5) allows to predict the behavior of (4.2.1) when K is large. To make the prediction effective from any initial conditions globally in the state-space, we impose the following assumption.

Assumption 4.2.4. The function $f_i(t,x)$ is uniformly bounded in t and globally Lipschitz with respect to x uniformly in t: i.e., $\exists$ a non-decreasing continuous function $M_f : \mathbb{R} \to \mathbb{R}$ and a constant $L_f \geq 0$ such that, $\forall x, y \in \mathbb{R}^n$, $t \in \mathbb{Z}$, and $i \in \mathcal{N}$,

$$\|f_i(t,x)\| \leq M_f\left(\|x\|\right),$$
$$\|f_i(t,x) - f_i(t,y)\| \leq L_f\|x-y\|. \qquad \Diamond$$

Theorem 4.2.2. Under Assumptions 4.2.1, 4.2.2, 4.2.3, and 4.2.4, for any $\epsilon > 0$, there exists $K^{\min}$ such that, for all $K > K^{\min}$, the solution x_i of (4.2.1) and the solution s of (4.2.5) with arbitrary initial conditions satisfy

$$\limsup_{t\to\infty} \left\|x_i[t] - p_i s[t]\right\| \leq \epsilon, \quad \forall i \in \mathcal{N}. \tag{4.2.6}$$

In addition, for each $k \in \{1, 2, \ldots, K-1\}$ and $i \in \mathcal{N}$,

$$\limsup_{t\to\infty} \left\|x_i[t_k] - p_i s[t+1]\right\| \leq \frac{\epsilon}{2}\left(1 + \frac{1}{|\lambda_N(W)|^{K-k}}\right). \tag{4.2.7}$$
$\Diamond$

Theorem 4.2.2 states that, with sufficiently large number of steps for the coupling (4.2.1b), the behavior of node dynamics (4.2.1a), which is represented by the state x_i at the integer count t, can be approximately predicted by the solution s of the blended dynamics with the scaling factor p_i, and the approximation error can be made arbitrarily small by increasing K. Moreover, the behavior of x_i over the fraction counts is also bounded with respect to the scaled trajectory of s.

Remark 4.2.2. Selection of the eigenvectors p and q as (4.2.4) is not unique, but the result of Theorem 4.2.2 remains the same. To see this, we note that different selection of p' and q' from p and q, respectively, should satisfy $p' = cp$ and $q' = (1/c)q$ for some $c > 0$ because of the Perron-Frobenius theorem (Lemma 2.2.2). In addition, we note that the new blended dynamics becomes $s'[t+1] = (1/c)\sum_{i=1}^{N} q_i f_i(t, cp_i s'[t]) =: f_{s'}(t, s'[t])$. Comparing it with (4.2.5), it is seen that $s'[t] = (1/c)s[t]$, and thus, we have $\|x_i[t] - p_i' s'[t]\| = \|x_i[t] - cp_i(1/c)s[t]\| = \|x_i[t] - p_i s[t]\|$. Finally, it is also seen that the new blended dynamics satisfies

Assumption 4.2.3 because $(\partial f_{s'}/\partial s')(t,s') = (\partial f_s/\partial s)(t,s)$ with $s = cs'$. $\Diamond$

Remark 4.2.3. In (4.2.7), the state $x_i[t_k]$ is compared not with $s[t]$ but with $s[t+1]$. One may find this is natural considering the behavior of the overall system. At each integer time $t = t_0$, each x_i obeys the heterogeneous node dynamics (4.2.1a), which potentially updates $x_i[t_1]$ in different directions from the updated $p_i s[t+1]$ (even if $x_i[t_0]$ is close to $p_i s[t]$). Instead, repeated execution of (4.2.1b) drives $x_i[t_k]$ to $p_i s[t+1]$, which is well reflected in (4.2.7). $\Diamond$

We now present intuitive explanations for Theorem 4.2.2 before providing a rigorous proof. For simplicity, define $\bar{x} := \mathrm{col}\,(x_1, \cdots, x_N) \in \mathbb{R}^{nN}$. Then, (4.2.1) is simply written as

$$\bar{x}[t_k] = W_{\otimes n}^{k-1} \begin{bmatrix} f_1(t_0, x_1[t_0]) \\ \vdots \\ f_N(t_0, x_N[t_0]) \end{bmatrix} =: W_{\otimes n}^{k-1} F(t_0, \bar{x}[t_0]), \tag{4.2.8}$$

for $k = 1, \ldots, K-1$ and $t \in \mathbb{Z}$, where $W_{\otimes n} = W \otimes I_n$, and, since $t_K = (t+1)_0 = t+1$, we have

$$\bar{x}[t+1] = W_{\otimes n}^{K-1} F(t, \bar{x}[t]). \tag{4.2.9}$$

Similar to (4.2.9), the blended dynamics (4.2.5) is written as

$$s[t+1] = \sum_{i=1}^{N} q_i f_i(t, p_i s[t]) = q_{\otimes n}^{\top} F(t, p_{\otimes n} s[t]). \tag{4.2.10}$$

By Lemma 4.2.1, there exist $P, Q \in \mathbb{R}^{N \times (N-1)}$ such that

$$W = \begin{bmatrix} p & P \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} q^{\top} \\ Q^{\top} \end{bmatrix}, \tag{4.2.11}$$

$$Q^{\top} P = I_{N-1}, \quad \text{and} \quad Q^{\top} p = P^{\top} q = 0_{N-1}, \tag{4.2.12}$$

where $\Lambda \in \mathbb{R}^{(N-1)\times(N-1)}$ is a matrix whose eigenvalues are $\lambda_2(W), \ldots, \lambda_N(W)$.

With them, we consider the coordinate transformation

$$\xi = \begin{bmatrix} \xi_1 \\ \tilde{\xi} \end{bmatrix} = \begin{bmatrix} q_{\otimes n}^\top \\ Q_{\otimes n}^\top \end{bmatrix} \bar{x} \tag{4.2.13}$$

whose inverse is

$$\bar{x} = p_{\otimes n}\xi_1 + P_{\otimes n}\tilde{\xi}.$$

The overall dynamics (4.2.9) at each integer time t becomes

$$\begin{aligned} \xi_1[t+1] &= q_{\otimes n}^\top W_{\otimes n}^{K-1} F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) \\ &= q_{\otimes n}^\top F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]), \\ \tilde{\xi}[t+1] &= Q_{\otimes n}^\top W_{\otimes n}^{K-1} F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) \\ &= (\Lambda^{K-1}Q^\top)_{\otimes n} F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]). \end{aligned} \tag{4.2.14}$$

Define the error variable $e := \xi_1 - s$. Then the above dynamics is rewritten by

$$\begin{aligned} e[t+1] &= q_{\otimes n}^\top F(t, p_{\otimes n}(e[t] + s[t]) + P_{\otimes n}\tilde{\xi}[t]) \\ &\quad - q_{\otimes n}^\top F(t, p_{\otimes n}s[t]), \\ \tilde{\xi}[t+1] &= (\Lambda^{K-1}Q^\top)_{\otimes n} F(t, p_{\otimes n}(e[t] + s[t]) \\ &\quad + P_{\otimes n}\tilde{\xi}[t]). \end{aligned}$$

Since the spectral radius $\rho(\Lambda) = |\lambda_2(W)| < 1$, it may be inferred that $\|\tilde{\xi}\|$ gets small if K is sufficiently large. On the other hand, if $\tilde{\xi}$ happens to be identically zero, then $e[t]$ converges to zero as t tends to infinity, which follows from the following lemma.

Lemma 4.2.3. Under Assumption 4.2.3,

$$\|H\{f_s(t, s_2) - f_s(t, s_1)\}\| \le \sqrt{\gamma}\|H(s_2 - s_1)\|$$

for all $t \in \mathbb{Z}$ and $s_1, s_2 \in \mathbb{R}^n$. ◊

Proof. Before proving Lemma 4.2.3, we first claim that, for each $t \in \mathbb{Z}$ and $s_1, s_2 \in$

$\mathbb{R}^n$, there exists $\tilde{s} \in \mathbb{R}^n$ in the line connecting s_1 and s_2 such that

$$\begin{aligned}&\{f_s(t,s_2)-f_s(t,s_1)\}^\top H^2\{f_s(t,s_2)-f_s(t,s_1)\}\\&\quad\le (s_2-s_1)^\top\frac{\partial f_s}{\partial s}(t,\tilde{s})^\top H^2\frac{\partial f_s}{\partial s}(t,\tilde{s})(s_2-s_1).\end{aligned}\tag{4.2.15}$$

It can be proved by the mean-value theorem with a trick. Let the left-hand side of the equality (4.2.15) as $\Delta(t,s_1,s_2)$ for convenience. With a variable $c\in\mathbb{R}$, define

$$g_{t,s_1,s_2}(c) := \{f_s(t,s_2)-f_s(t,s_1)\}^\top H^2\{f_s(t,cs_2+(1-c)s_1)-f_s(t,s_1)\}.$$

Since the function $g_{t,s_1,s_2}:\mathbb{R}\to\mathbb{R}$ is continuously differentiable, by the mean-value theorem, there exists $\tilde{c}\in[0,1]$ such that $g_{t,s_1,s_2}(1)-g_{t,s_1,s_2}(0)=g'_{t,s_1,s_2}(\tilde{c})(1-0)$, which is equivalent to

$$\Delta(t,s_1,s_2)=\{f_s(t,s_2)-f_s(t,s_1)\}^\top H^2\frac{\partial f_s}{\partial s}(t,\tilde{s})(s_2-s_1)\tag{4.2.16}$$

where $\tilde{s}=\tilde{c}s_2+(1-\tilde{c})s_1$. Using this, we have

$$\begin{aligned}\Delta(t,s_1,s_2)&=\|H\{f_s(t,s_2)-f_s(t,s_1)\}\|^2\\&\le\|H\{f_s(t,s_2)-f_s(t,s_1)\}\|\left\|H\frac{\partial f_s}{\partial s}(t,\tilde{s})(s_2-s_1)\right\|,\end{aligned}$$

which in turn implies $\|H\{f_s(t,s_2)-f_s(t,s_1)\}\|\le\|H(\partial f_s/\partial s)(t,\tilde{s})(s_2-s_1)\|$. From this, the claim (4.2.15) is justified. Finally, it follows from Assumption 4.2.3 that

$$\begin{aligned}\|H\{f_s(t,s_2)-f_s(t,s_1)\}\|^2&=\Delta(t,s_1,s_2)\\&\le(s_2-s_1)^\top\frac{\partial f_s}{\partial s}(t,\tilde{s})^\top H^2\frac{\partial f_s}{\partial s}(t,\tilde{s})(s_2-s_1)\\&\le\gamma(s_2-s_1)^\top H^2(s_2-s_1)\\&=\gamma\|H(s_2-s_1)\|^2,\end{aligned}$$

which proves Lemma 4.2.3. □

In fact, if $\tilde{\xi} \equiv \mathbb{0}_{n(N-1)}$, then,

$$\begin{aligned}\left\|He[t+1]\right\| &= \|H\{q_{\otimes n}^{\top}F(t, p_{\otimes n}(e[t]+s[t])) - q_{\otimes n}^{\top}F(t, p_{\otimes n}s[t])\}\| \\ &= \|H\{f_s(t, e[t]+s[t]) - f_s(t, s[t])\}\| \\ &\leq \sqrt{\gamma}\|He[t]\|.\end{aligned}$$

This implies that ξ_1 and s tend to get closer as time goes on. When this happens, $\bar{x} = p_{\otimes n}\xi_1 + P_{\otimes n}\tilde{\xi}$ tends to $p_{\otimes n}s$, which motivates Theorem 4.2.2.

However, $\tilde{\xi}$ does not become identically zero in general, and the above arguments need rigorous analysis with a Lyapunov function as follows. Before providing the proof of Theorem 4.2.2, we first claim that the solution $s[t]$ of the blended dynamics is bounded by Assumptions 4.2.3 and 4.2.4 as follows.

Lemma 4.2.4. Under Assumption 4.2.3, the solutions of the blended dynamics (4.2.5), initiated at time $t = t^0$, are bounded as

$$\|Hs[t]\| \leq \sqrt{\gamma}^{t-t^0}\|Hs[t^0]\| + \frac{\sup_{\tau \geq t^0}\|Hf_s(\tau, \mathbb{0}_n)\|}{1-\sqrt{\gamma}}. \tag{4.2.17}$$

Additionally, with Assumption 4.2.4, we have

$$\limsup_{t\to\infty}\|s[t]\| \leq \frac{\sqrt{N}\|q\|\|H^{-1}\|\|H\|M_f(0)}{1-\sqrt{\gamma}} =: M_s \tag{4.2.18}$$

$\lozenge$

Proof. We prove (4.2.17) by showing that

$$\|Hs[t]\| \leq w[t], \quad \forall t \geq t^0 \tag{4.2.19}$$

where $w \in \mathbb{R}$ is the solution of

$$w[t+1] = \sqrt{\gamma}w[t] + \|Hf_s(t, \mathbb{0}_n)\|, \quad w[t^0] = \|Hs[t^0]\|. \tag{4.2.20}$$

Indeed, since $\|Hs[t^0]\| \leq w[t^0]$, let us suppose $\|Hs[\tau]\| \leq w[\tau]$ for some integer $\tau \geq t^0$. Then,

$$\|Hs[\tau+1]\| = \|H\{f_s(\tau, s[\tau]) - f_s(\tau, \mathbb{0}_n) + f_s(\tau, \mathbb{0}_n)\}\|$$

$$
\begin{aligned}
&\leq \|H\{f_s(\tau, s[\tau]) - f_s(\tau, \mathbb{0}_n)\}\| \\
&\quad + \|Hf_s(\tau, \mathbb{0}_n)\| \\
&\leq \sqrt{\gamma}\|Hs[\tau]\| + \|Hf_s(\tau, \mathbb{0}_n)\| \\
&\leq \sqrt{\gamma}w[\tau] + \|Hf_s(\tau, \mathbb{0}_n)\| \\
&= w[\tau + 1],
\end{aligned}
$$

where the second inequality comes from Lemma 4.2.3. This justifies (4.2.19). Meanwhile, the solution w of (4.2.20) is given by

$$
w[t] = \sqrt{\gamma}^{t-t^0} w[t^0] + \sum_{\tau=t^0}^{t-1} \sqrt{\gamma}^{t-\tau-1} \|Hf_s(\tau, \mathbb{0}_n)\|,
$$

which yields (4.2.17).

Now, with Assumption 4.2.4, it follows from (4.2.17) that

$$
\begin{aligned}
\limsup_{t\to\infty} \|s[t]\| &\leq \|H^{-1}\| \limsup_{t\to\infty} \|Hs[t]\| \\
&\leq \|H^{-1}\|\|H\| \frac{\sup_{\tau\geq t^0} \|q_{\otimes n}^\top F(t, p_{\otimes n}\mathbb{0}_n)\|}{1 - \sqrt{\gamma}} \\
&\leq \|H^{-1}\|\|H\| \frac{\|q\|\sqrt{N} M_f(0)}{1 - \sqrt{\gamma}}
\end{aligned}
$$

which completes the proof. □

Based on the aforementioned observations, now we can prove Theorem 4.2.2 in the following way.

Proof. We analyze the behavior of (4.2.14) with (4.2.10), which describes the evolution of the overall system at every integer time t. For this, we introduce a Lyapunov function

$$
V = \|H(\xi_1 - s)\| + \eta\|\tilde{\xi}\|
$$

where $\eta > L_f\|q\|\|P\|\|H\|/\sqrt{\gamma}$. Then,

$$
\begin{aligned}
V[t+1] &= \|H(\xi_1[t+1] - s[t+1])\| + \eta\|\tilde{\xi}[t+1]\| \\
&= \big\|H\{q_{\otimes n}^\top F(t, p_{\otimes n}\xi_1[t]) - q_{\otimes n}^\top F(t, p_{\otimes n}s[t])
\end{aligned}
$$

$$
\begin{aligned}
&+ q_{\otimes n}^{\top} F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) - q_{\otimes n}^{\top} F(t, p_{\otimes n}\xi_1[t]))\}\big\| \\
&+ \eta\big\|(\Lambda^{K-1}Q^{\top})_{\otimes n}\{F(t, p_{\otimes n}s[t]) \\
&+ F(t, p_{\otimes n}\xi_1[t]) - F(t, p_{\otimes n}s[t]) \\
&+ F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) - F(t, p_{\otimes n}\xi_1[t]))\}\big\| \\
\leq\ & \big\|H\{q_{\otimes n}^{\top} F(t, p_{\otimes n}\xi_1[t]) - q_{\otimes n}^{\top} F(t, p_{\otimes n}s[t]))\}\big\| \\
&+ \big\|Hq_{\otimes n}^{\top}\{F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) - F(t, p_{\otimes n}\xi_1[t]))\}\big\| \\
&+ |\lambda_2(W)|^{K-1}\eta\Big\{\|Q_{\otimes n}^{\top}\{F(t, p_{\otimes n}\xi_1[t]) - F(t, p_{\otimes n}s[t]))\}\| \\
&\quad + \|Q_{\otimes n}^{\top}\{F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) - F(t, p_{\otimes n}\xi_1[t]))\}\| \\
&\quad + \|Q_{\otimes n}^{\top} F(t, p_{\otimes n}s[t]))\|\Big\}.
\end{aligned}
$$

The above inequality can be simplified by the following properties:

- by Lemma 4.2.3,

$$
\begin{aligned}
&\|H\{q_{\otimes n}^{\top} F(t, p_{\otimes n}\xi_1[t]) - q_{\otimes n}^{\top} F(t, p_{\otimes n}s[t]))\}\| \\
&\quad = \|H\{f_s(t, \xi_1[t]) - f_s(t, s[t])\}\| \\
&\quad \leq \sqrt{\gamma}\|H(\xi_1[t] - s[t])\|
\end{aligned}
$$

- by Assumption 4.2.4,

$$
\begin{aligned}
&\|F(t, p_{\otimes n}\xi_1[t] + P_{\otimes n}\tilde{\xi}[t]) - F(t, p_{\otimes n}\xi_1[t])\| \\
&\leq \left(\sum_{i=1}^{N} \|f_i(t, p_i\xi_1[t] + (P_i \otimes I_n)\tilde{\xi}[t]) - f_i(t, p_i\xi_1[t])\|^2\right)^{1/2} \\
&\leq \left(\sum_{i=1}^{N} L_f^2\|(P_i \otimes I_n)\tilde{\xi}[t]\|^2\right)^{1/2} \\
&= L_f\|P_{\otimes n}\tilde{\xi}[t]\| \leq L_f\|P\|\|\tilde{\xi}[t]\|,
\end{aligned}
$$

  where P_i is the i-th row of P, and similarly

$$
\begin{aligned}
&\|F(t, p_{\otimes n}\xi_1[t]) - F(t, p_{\otimes n}s[t])\| \\
&\leq L_f\|p\|\|\xi_1[t] - s[t]\|
\end{aligned}
$$

$$
\begin{aligned}
&= L_f\|p\|\|H^{-1}H(\xi_1[t]-s[t])\| \\
&\le L_f\|p\|\|H^{-1}\|\|H(\xi_1[t]-s[t])\|.
\end{aligned}
$$

Using the above three inequalities, we have that

$$
\begin{aligned}
V[t+1] \le{}& \sqrt{\gamma}\|H(\xi_1[t]-s[t])\| + \|H\|\|q\|L_f\|P\|\|\tilde{\xi}[t]\| \\
&+ |\lambda_2(W)|^{K-1}\eta\|Q\|L_f\|p\|\|H^{-1}\|\|H(\xi_1[t]-s[t])\| \\
&+ |\lambda_2(W)|^{K-1}\eta\|Q\|L_f\|P\|\|\tilde{\xi}[t]\| \\
&+ |\lambda_2(W)|^{K-1}\eta\|Q\|\|F(t,p_{\otimes n}s[t])\| \\
\le{}& \sqrt{\gamma}V[t] + |\lambda_2(W)|^{K-1}\eta L_f\|Q\|M_1V[t] \\
&+ |\lambda_2(W)|^{K-1}\eta\|Q\|\|F(t,p_{\otimes n}s[t])\|,
\end{aligned}
$$

where $M_1 := \max\{\|p\|\|H^{-1}\|, \|P\|/\eta\}$.

For the given ϵ, let $K^{\min}$ be a positive integer such that

$$
|\lambda_2(W)|^{K^{\min}}\eta L_f M_1\|Q\| \le \frac{1-\sqrt{\gamma}}{2} \tag{4.2.21}
$$

$$
|\lambda_2(W)|^{K^{\min}}\frac{2\eta M_1 M_f(\|p\|M_s)\sqrt{N}\|Q\|}{1-\sqrt{\gamma}} \le \frac{\epsilon}{2}. \tag{4.2.22}
$$

Then, for all $K > K^{\min}$,

$$
V[t+1]-V[t] \le -\frac{(1-\sqrt{\gamma})}{2}V[t] + |\lambda_2(W)|^{K-1}\eta\|Q\|\|F(t,p_{\otimes n}s[t])\|. \tag{4.2.23}
$$

By Assumption 4.2.4 and by Lemma 4.2.4,

$$
\begin{aligned}
\limsup_{t\to\infty}\|F(t,p_{\otimes n}s[t])\| &\le \sqrt{N}M_f(\|p\|\limsup_{t\to\infty}\|s[t]\|) \\
&\le \sqrt{N}M_f\left(\|p\|M_s\right)
\end{aligned}
$$

Using this and (4.2.23), the ultimate bound of V is obtained as

$$
\limsup_{t\to\infty} V[t] \le |\lambda_2(W)|^{K-1}\frac{2\eta\|Q\|}{1-\sqrt{\gamma}}\sqrt{N}M_f(\|p\|M_s). \tag{4.2.24}
$$

Therefore, for each agent $i \in \mathcal{N}$ and $K > K^{\min}$,

$$\begin{aligned}\limsup_{t\to\infty} \|x_i[t] - p_i s[t]\| &= \limsup_{t\to\infty} \left\| p_i(\xi_1[t] - s[t]) + (P_i \otimes I_n)\tilde{\xi}[t] \right\| \\ &\le \max\left\{ \|p\|\|H^{-1}\|, \frac{\|P\|}{\eta} \right\} \limsup_{t\to\infty} V[t] \\ &\le |\lambda_2(W)|^{K-1} \frac{2\eta\|Q\|}{1-\sqrt{\gamma}} \sqrt{N} M_f(\|p\|M_s) M_1 \\ &\le \epsilon\end{aligned}$$

where we used $\bar{x} = p_{\otimes n}\xi_1 + P_{\otimes n}\tilde{\xi}$. This completes the proof for (4.2.6) of Theorem 4.2.2.

Now, in order to inspect the behavior of the system over the fractional time, let us apply the transformation of (4.2.13) to (4.2.8), which yields, for $k = 1, \cdots, K-1$,

$$\begin{aligned}\xi_1[t_k] &= q_{\otimes n}^\top W_{\otimes n}^{k-1} F(t_0, \bar{x}[t_0]) \\ &= q_{\otimes n}^\top F(t_0, \bar{x}[t_0]) = q_{\otimes n}^\top \bar{x}[(t+1)_0] \\ &= \xi_1[(t+1)_0]\end{aligned} \tag{4.2.25}$$

in which, the third equality can also be seen from (4.2.9). Similarly, for $k = 1, \cdots, K-1$,

$$\begin{aligned}\tilde{\xi}[t_k] &= Q_{\otimes n}^\top W_{\otimes n}^{k-1} F(t_0, \bar{x}[t_0]) \\ &= (\Lambda^{k-1} Q^\top)_{\otimes n} F(t_0, \bar{x}[t_0]) \\ &= \Lambda_{\otimes n}^{k-K} (\Lambda^{K-1} Q^\top)_{\otimes n} F(t_0, \bar{x}[t_0]) \\ &= \Lambda_{\otimes n}^{k-K} Q_{\otimes n}^\top W_{\otimes n}^{K-1} F(t_0, \bar{x}[t_0]) \\ &= (\Lambda^{-1})_{\otimes n}^{K-k} \tilde{\xi}[(t+1)_0].\end{aligned}$$

Hence,

$$\|\tilde{\xi}[t_k]\| \le \frac{\|\tilde{\xi}[(t+1)_0]\|}{|\lambda_N(W)|^{K-k}} \tag{4.2.26}$$

for each $k = 1, \cdots, K-1$.

On the other hand, by (4.2.24) and (4.2.22), we have

$$\begin{aligned}\limsup_{t\to\infty}\max\{\|H(\xi_1[t]-s[t])\|,\eta\|\tilde{\xi}[t]\|\} &\le \limsup_{t\to\infty} V[t]\\ &\le \frac{\epsilon}{2M_1}.\end{aligned}$$

Therefore, for each $k=1,\cdots,K-1$,

$$\begin{aligned}&\limsup_{t\to\infty}\|x_i[t_k]-p_i s[t+1]\|\\ &\quad=\limsup_{t\to\infty}\|p_i(\xi_1[t_k]-s[t+1])+(P_i\otimes I_n)\tilde{\xi}[t_k]\|\\ &\quad\le\limsup_{t\to\infty}\|p\|\|H^{-1}\|\|H(\xi_1[t_k]-s[t+1])\|\\ &\qquad+\frac{\|P\|}{\eta}\eta\|\tilde{\xi}[t_k]\|\\ &\quad\le\limsup_{t\to\infty}\|p\|\|H^{-1}\|\|H(\xi_1[t+1]-s[t+1])\|\\ &\qquad+\frac{\|P\|}{\eta}\frac{\eta\|\tilde{\xi}[t+1]\|}{|\lambda_N(W)|^{K-k}}\\ &\quad\le\frac{\|p\|\|H^{-1}\|}{M_1}\frac{\epsilon}{2}+\frac{\|P\|/\eta}{M_1}\frac{\epsilon}{2|\lambda_N(W)|^{K-k}}\\ &\quad\le\frac{\epsilon}{2}\left(1+\frac{1}{|\lambda_N(W)|^{K-k}}\right),\end{aligned}$$

which completes the proof. □

Remark 4.2.4. From (4.2.21) and (4.2.22), $K^{\min}$ can be explicitly defined as the smallest integer such that

$$\begin{aligned}K^{\min}&\ge\log_{|\lambda_2(W)|}\left(\frac{1-\sqrt{\gamma}}{2\eta\|Q\|M_1}\max\left\{\frac{1}{L_f},\frac{\epsilon}{2M_f(\|p\|M_s)\sqrt{N}}\right\}\right)\\ &=\log_{|\lambda_2(W)|}\left(\frac{1-\sqrt{\gamma}}{2\eta\|Q\|M_1}\right)+\log_{|\lambda_2(W)|}\left(\max\left\{\frac{1}{L_f},\frac{\epsilon}{2M_f(\|p\|M_s)\sqrt{N}}\right\}\right).\end{aligned}$$

This yields a reasonable interpretation that $K^{\min}$ increases as the second largest eigenvalue $\lambda_2(W)$ of the weight matrix W approaches to 1, the performance index ϵ decreases, or the Lipschitz constant L_f and the network size N get larger, while it decreases as the degree of stability $(1-\sqrt{\gamma})$ of the blended dynamics gets

larger. ◊

In Theorem 4.2.2, the approximation of $x_i[t]$ by $p_i s[t]$ is stated in an asymptotic format, i.e., using lim sup. If one questions when the approximation becomes effective, the following corollary assures that it can be very early if K is sufficiently large. In particular, if the initial conditions of x_i are in a known compact set, then $K^{\min}$ can be computed.

Corollary 4.2.5. Under Assumptions 4.2.1, 4.2.2, 4.2.3, and 4.2.4, for any $\epsilon > 0$ and compact set $C \subset \mathbb{R}^n$, there exists $K^{\min} > 0$ such that, for all $K > K^{\min}$, the solution x_i of (4.2.1) with $x_i[0] \in C$, and the solution s of (4.2.5) with $s[1] = \sum_{i=1}^{N} q_i f_i(0, x_i[0])$ satisfy

$$\left\| x_i[t] - p_i s[t] \right\| \leq \epsilon, \quad \forall t \geq 1,\ i \in \mathcal{N}. \tag{4.2.27}$$

In addition, for all $k \in \{1, 2, \ldots, K-1\}$, $i \in \mathcal{N}$, and $t \geq 1$,

$$\left\| x_i[t_k] - p_i s[t+1] \right\| \leq \frac{\epsilon}{2} \left(1 + \frac{1}{|\lambda_N(W)|^{K-k}} \right). \tag{4.2.28}$$

◊

Proof. In this proof, we show that, after $K-1$ times execution of (4.2.1b), the solution of the overall system from the initial condition enters a positively invariant set, in which (4.2.27) holds. For this, let us first construct a few sets as

$$\begin{aligned} C_1^s &= \{ q_{\otimes n}^\top F(0, \bar{x}) : x_i \in C, i \in \mathcal{N} \} \subset \mathbb{R}^n, \\ C_{t+1}^s &= \{ q_{\otimes n}^\top F(t, p_{\otimes n} s) : s \in C_t^s \} \subset \mathbb{R}^n,\ \forall t \geq 1, \\ C_\infty^s &= \bigcup_{t=1}^{\infty} C_t^s \subset \mathbb{R}^n, \end{aligned}$$

in which, C_t^s is the set of all possible $s[t]$ for each $t \geq 1$, which is bounded. The set C_∞^s is also bounded by Lemma 4.2.4. Now, for the overall state $\bar{x}$, consider two more sets:

$$\begin{aligned} C' &= \{ \bar{x} : \|H(\xi_1 - s)\| \leq \epsilon_0, s \in C_\infty^s, \|\tilde{\xi}\| \leq \delta \} \cup C^N \subset \mathbb{R}^{nN}, \\ \bar{C} &= \{ (\bar{x}, s) \in C' \times C_\infty^s : \|H(\xi_1 - s)\| \leq \epsilon_0, \|\tilde{\xi}\| \leq \delta \} \subset \mathbb{R}^{nN} \times \mathbb{R}^n, \end{aligned}$$

where C^N is N-ary Cartesian power of C, i.e., $C^N = \underbrace{C \times C \times \cdots \times C}_{N\text{-times}}$ and

$$\epsilon_0 := \frac{\epsilon}{2\max\{\|p\|\|H^{-1}\|, \|P\|\}},$$
$$\delta := \min\left\{1, \frac{1-\sqrt{\gamma}}{L_f\|q\|\|P\|\|H\|}\right\}\epsilon_0.$$

Now, pick $K^{\min}$ such that

$$|\lambda_2(W)|^{K^{\min}}\|Q\| \sup_{\bar{x}\in\bar{C}'} \|F(t,\bar{x})\| \le \delta, \quad \forall t \ge 0.$$

We claim that the set $\bar{C}$ is positively invariant for any $K > K^{\min}$. To see this, suppose that, for any $\tau \ge 1$, $s[\tau] \in C^s_\tau \subset C^s_\infty$, $\|\tilde{\xi}[\tau]\| \le \delta$, and $\|H(\xi_1[\tau] - s[\tau])\| \le \epsilon_0$, so that $(\bar{x}[\tau], s[\tau]) \in \bar{C}$. Then, it follows that

$$\begin{aligned} s[\tau+1] &\in C^s_{\tau+1} \subset C^s_\infty \\ \|\tilde{\xi}[\tau+1]\| &\le |\lambda_2(W)|^{K-1}\|Q\|\|F(\tau,\bar{x}[\tau])\|\| \\ &\le \delta, \end{aligned}$$

and

$$\begin{aligned} \|H(\xi_1[\tau+1] - s[\tau+1])\| &\le \|H\{\xi_1[\tau+1] - q^\top_{\otimes n}F(\tau, p_{\otimes n}\xi_1[\tau])\}\| \\ &\quad + \|H\{q^\top_{\otimes n}F(\tau, p_{\otimes n}\xi_1[\tau]) - s[\tau+1])\}\| \\ &\le L_f\|q\|\|P\|\|H\|\delta + \|H\{f_s(\tau,\xi_1[\tau]) - f_s(\tau,s[\tau])\}\| \\ &\le L_f\|q\|\|P\|\|H\|\delta + \sqrt{\gamma}\|H(\xi_1[\tau] - s[\tau])\| \\ &\le \epsilon_0 \end{aligned}$$

in which, we used

$$\begin{aligned} &\|\xi_1[\tau+1] - q^\top_{\otimes n}F(\tau, p_{\otimes n}\xi_1[\tau])\| \\ &= \|q^\top_{\otimes n}\{F(\tau, p_{\otimes n}\xi_1[\tau] + P_{\otimes n}\tilde{\xi}[\tau]) - F(\tau, p_{\otimes n}\xi_1[\tau])\}\| \\ &\le L_f\|q\|\|P\|\|\tilde{\xi}[\tau]\| \le L_f\|q\|\|P\|\delta. \end{aligned}$$

On the other hand, the set $\bar{C}$ is reached within $K-1$ executions of (4.2.1b) from the initial time 0_0. Indeed,

$$\begin{aligned} s[1] &= q_{\otimes n}^{\top} F(0, \bar{x}[0]) \in C_1^s \subset C_\infty^s \\ \|\tilde{\xi}[1]\| &= \|(\Lambda^{K-1} Q^{\top})_{\otimes n} F(0, \bar{x}[0])\| \\ &\leq |\lambda_2(W)|^{K-1} \|Q\| \|F(0, \bar{x}[0])\| \\ &\leq \delta \\ \|H(\xi_1[1] - s[1])\| &= \|H\{q_{\otimes n}^{\top} W_{\otimes n}^{K-1} F(0, \bar{x}[0]) - q_{\otimes n}^{\top} F(0, \bar{x}[0])\}\| \\ &= 0. \end{aligned}$$

Therefore, $(\bar{x}[t], s[t])$ remains in $\bar{C}$ for all $t \geq 1$.

Finally, it is seen that, in the set $\bar{C}$,

$$\begin{aligned} \|x_i[t] - p_i s[t]\| &= \|p_i(\xi_1[t] - s[t]) + (P_i \otimes I_n)\tilde{\xi}[t]\| \\ &\leq \max\left\{\|p\|\|H^{-1}\|, \|P\|\right\} (\|H(\xi_1[t] - s[t])\| + \|\tilde{\xi}[t]\|) \\ &\leq \epsilon, \end{aligned}$$

which completes the proof of (4.2.27).

To show (4.2.28), we note that (4.2.25) and (4.2.26) still hold. Therefore, for all $k = 1, \cdots, K-1$, $i \in \mathcal{N}$, and $t \geq 1$,

$$\begin{aligned} \|x_i[t_k] - p_i s[t+1]\| &= \|p_i(\xi_1[t_k] - s[t+1]) + (P_i \otimes I_n)\tilde{\xi}[t_k]\| \\ &\leq \max\left\{\|p\|\|H^{-1}\|, \|P\|\right\} (\|H(\xi_1[t_k] - s[t+1])\| + \|\tilde{\xi}[t_k]\|) \\ &\leq \max\left\{\|p\|\|H^{-1}\|, \|P\|\right\} \\ &\quad \times (\|H(\xi_1[t+1] - s[t+1])\| + \|\tilde{\xi}[t+1]\|/|\lambda_N(W)|^{K-k}) \\ &\leq \max\left\{\|p\|\|H^{-1}\|, \|P\|\right\} \left(\epsilon_0 + \frac{\delta}{|\lambda_N(W)|^{K-k}}\right) \\ &\leq \frac{\epsilon}{2}\left(1 + \frac{1}{|\lambda_N(W)|^{K-k}}\right), \end{aligned}$$

which completes the proof of (4.2.28). □

Remark 4.2.5. In Corollary 4.2.5, the solution $s[t]$ of the blended dynamics is

initiated not at $t = 0$ but at $t = 1$. One may find this is natural because the state $s'[1] = f_s(0, s'[0])$ with $s'[0] = \sum_{i=1}^{N} q_i x_i[0]$, i.e., $s'[1] = \sum_{i=1}^{N} q_i f_i(0, p_i \sum_{i=1}^{N} q_i x_i[0])$, can be very different from the considered state $s[1] = \sum_{i=1}^{N} q_i f_i(0, x_i[0])$. In fact, the states $x_j[0_1]$, $j \in \mathcal{N}$, may be very different from each other, but they converge with sufficiently large K towards $\sum_{i=1}^{N} q_i x_i[0_1]$ with the scaling factor p_j, which is $s[1]$ (not $s'[1]$). ◊

4.3 Network Synthesis with Examples

As mentioned before, the proposed approach is useful as a design method for distributed algorithms by first designing a suitable blended dynamics such that it behaves as desired and then synthesizing each heterogeneity of the multi-agent system such that it has the pre-designed blended dynamics. Moreover, comparing with the continuous-time approach, the discrete-time version handles more general protocols as long as the spectral radius of its weight matrix is 1. In fact, many studies on discrete-time multi-agent system including PageRank [IT10, ITB12, LC14] or synchronization [OSFM07, RB05] have used the communication protocol which can be represented as a linear combination of agents' state information like (4.2.1b) with the weight matrix of unit spectral radius. Thus, in this section, some of the protocols are chosen to be considered as the coupling dynamics (4.2.1b) and the behaviors of corresponding multi-step coupling systems are illustrated using the results in previous section. Based on this, the design process for each application example is also provided.

4.3.1 Distributed Estimation of the Number of Agents in Network

In this subsection, we assume that the network is undirected and connected, and consider the following Metropolis-Hastings coupling weight w_{ij}^{MH} in [SHM14]:

$$w_{ij}^{\mathrm{MH}} := \begin{cases} \dfrac{1-\mu^{\mathrm{MH}}}{\max\{d_i, d_j\}}, & (j,i) \in \mathcal{E} \text{ and } i \neq j, \\ 0, & (j,i) \notin \mathcal{E} \text{ and } i \neq j, \\ 1 - \displaystyle\sum_{l \neq i} w_{il}^{\mathrm{MH}}, & i = j, \end{cases} \tag{4.3.1}$$

where $\mu^{\mathrm{MH}} \in (0,1)$. Note that w_{ij}^{MH} depends on both d_i and d_j, so that each agent should additionally exchange its degree information with neighbors to update its coupling weights online, and this will enable the plug-and-play operation to the multi-step coupling framework.

It can be easily seen that w_{ij}^{MH} satisfies (4.2.3) and Assumption 4.2.1 holds for $W^{\mathrm{MH}} := [w_{ij}^{\mathrm{MH}}]$ because it is doubly-stochastic. Hence, we can choose $p = \mathbb{1}_N$ and $q = (1/N)\mathbb{1}_N$ from Lemma 4.2.1. Then, the blended dynamics is obtained as a simple average of f_is as follows:

$$s[t+1] = \frac{1}{N}\sum_{i=1}^{N} f_i(t, s[t]). \tag{4.3.2}$$

Since all p_is are chosen evenly as 1, by Theorem 4.2.2 or Corollary 4.2.5, the behavior of every trajectory $x_i[t]$ is approximately synchronized to the solution s of the blended dynamics (4.3.2). It should be emphasized that this collective synchronized behavior comes from $p = \mathbb{1}_N$ and this choice of p is always possible for any row-stochastic (not necessarily to be doubly-stochastic) weight matrices. In Section 4.3.3, we will consider the row-stochastic weight matrix whose column-sums are not 1.

As an application example for the Metropolis-Hastings coupling, we can design a distributed algorithm for network size estimation as follows. In fact, many distributed algorithms such as [ITB12, NO09] are often assumed to know the network size N. The insight of the proposed algorithm is to make its blended dynamics converge to N. For example, if the blended dynamics is designed as the following scalar dynamics

$$s[t+1] = \left(1 - \frac{1}{N}\right) s[t] + 1, \tag{4.3.3}$$

such that it has the stable equilibrium point at N, then each state $x_i[t]$ will also approach to N under the Metropolis-Hastings coupling. Thus, by increasing K until the synchronization error ϵ in (4.2.6) or (4.2.27) gets smaller than 0.5, each agent can find the exact network size through the round-off to the nearest integer.

One idea to design the heterogeneous dynamics f_is whose average becomes

(4.3.3) comes from

$$\frac{1}{N}\left\{\left(1\right)+\left(\sum_{i=2}^{N}(s[t]+1)\right)\right\}=\left(1-\frac{1}{N}\right)s[t]+1,$$

this is, one agent has $f_i = 1$ and all others have $f_i(s) = s + 1$. For this, we intentionally add one specific node which does not leave the network during the operation of the algorithm. Without loss of generality, let an index of this node be 1 and it runs the following dynamics:

$$x_1[t_{k+1}] = \begin{cases} 1, & \text{if } k = 0, \\ \sum_{j\in\mathcal{N}_1\cup\{1\}} w_{1j}^{\mathrm{MH}} x_j[t_k], & \text{otherwise.} \end{cases} \tag{4.3.4}$$

On the other hand, all the other nodes of $i = 2, \ldots, N$ run the following dynamics:

$$x_i[t_{k+1}] = \begin{cases} x_i[t_k] + 1, & \text{if } k = 0, \\ \sum_{j\in\mathcal{N}_i\cup\{i\}} w_{ij}^{\mathrm{MH}} x_j[t_k], & \text{otherwise.} \end{cases} \tag{4.3.5}$$

Note that, even though the individual dynamics of $(N-1)$ nodes for $i = 2, \ldots, N$ are (marginally) unstable, the overall networked system becomes stable and the trajectories of individual agent approach close to N (less than the distance of 0.5 with sufficiently large K). In addition, this distributed algorithm can be applied even when some agents might join or leave the network during the process of the algorithm, because it does not rely on the initial condition of agents. This idea is motivated by [LLKS18] which proposed continuous-time distributed network size estimation algorithm.

Example 4.3.1. To verify the performance of the proposed algorithm (4.3.4) and (4.3.5), we consider a network where the number of the agents varies during the process of the algorithm as shown in Figure 4.1. Initially, there are 5 agents in the network, but, after an advance of 50 integer count t, the agent 2 leaves the network so that the number of agents, N, changes to 4.

The simulation result of the proposed algorithm (4.3.4) and (4.3.5) is given in Figure 4.2. Here, we use a parameter $\mu^{\mathrm{MH}} = 0.1$ in the Metropolis-Hastings

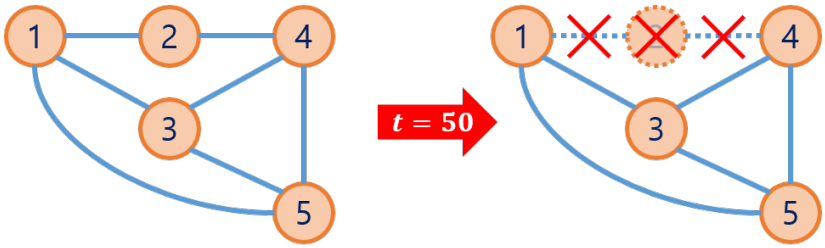

Figure 4.1: Time-varying network considered in Example 4.3.1

coupling weight (4.3.2). Moreover, the coupling dynamics is repeated for 5 times for each integer count (i.e., $K = 6$) in the algorithm so that the synchronization error (which is denoted as ϵ in Theorem 4.2.2) is less than 0.5. From this, each agent can estimate the exact network size N through the round-off to the closest integer. Moreover, even though all agents are arbitrarily initialized with 0 in the simulation, the algorithm could eventually estimate the network size N regardless of the initial conditions and the network change. This result confirms the plug-and-play feature of the proposed algorithm. ◊

4.3.2 Initialization-free Distributed PageRank Estimation for Strongly Connected Network

Now, we turn our attention to a different type of the weight matrix whose column-sums are all one. In this subsection, we consider the multi-step coupling framework whose coupling dynamics is the following iterative power method of PageRank [BP98]:

$$x_i[t_{k+1}] = \theta^{\mathrm{PR}} x_i[t_k] + (1 - \theta^{\mathrm{PR}}) \sum_{j \in \mathcal{N}_i} \frac{x_j[t_k]}{d_j^{\mathrm{out}}}, \tag{4.3.6}$$

where $x_i \in \mathbb{R}$ is the state, $\theta^{\mathrm{PR}} \in (0, 1)$ is the constant parameter, $\mathcal{N}_i$ is the in-neighbors of node i, and d_j^{out} is the out-degree of node j. In fact, PageRank score provides an information on relative importance of each node in the network, so it has been widely utilized in diverse areas such as informatics [CXMR07], bibliometrics [LBNVdS05], and biology [ZBE12].

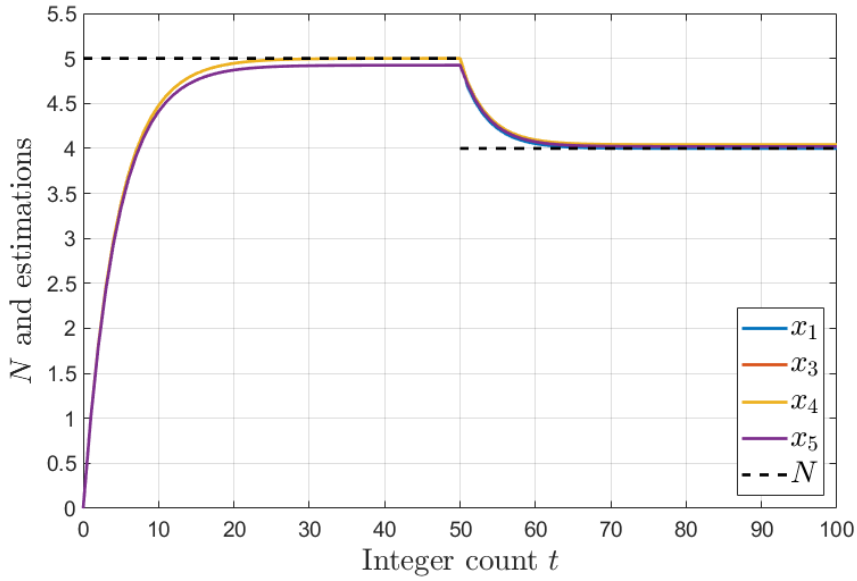

Figure 4.2: Simulation result for distributed network size estimation: Dashed line and solid lines represent the network size N and the estimate of each agent, respectively.

Then, the coupling weight w_{ij}^{PR} is given by

$$w_{ij}^{\mathrm{PR}} = \begin{cases} \theta^{\mathrm{PR}}, & i = j, \\ (1-\theta^{\mathrm{PR}})\dfrac{a_{ij}}{d_j^{\mathrm{out}}}, & i \neq j, \end{cases}$$

where a_{ij} is the ij-th element of the binary adjacency matrix A. It can be easily seen that w_{ij}^{PR} satisfies (4.2.3) and its weight matrix $W^{\mathrm{PR}} := [w_{ij}^{\mathrm{PR}}]$ is obtained as

$$W^{\mathrm{PR}} = \theta^{\mathrm{PR}} I + (1-\theta^{\mathrm{PR}}) A D_{\mathrm{out}}^{-1},$$

where D_{out} is a diagonal matrix whose diagonal components are $d_1^{\mathrm{out}}, \cdots, d_N^{\mathrm{out}}$ in sequence. Since W^{PR} is column-stochastic, it has the spectral radius of 1 with the left eigenvector $q = \mathbb{1}_N$. By Lemma 4.2.1, there exists a positive right eigenvector

$p \in \mathbb{R}^N$ for the eigenvalue 1 such that

$$W^{\mathrm{PR}} p = p, \quad \mathbb{1}_N^\top p = 1.$$

Here, each element p_i of p is called as PageRank score of node i in the strongly connected network and it represents the relative importance of each node in the network [BP98]. From this, the blended dynamics under PageRank coupling (4.3.6) is written as, with $s \in \mathbb{R}$,

$$s[t+1] = \mathbb{1}_N^\top F\left(t, ps[t]\right) = \sum_{i=1}^{N} f_i\left(t, p_i s[t]\right). \tag{4.3.7}$$

By Theorem 4.2.2, the i-th agent's trajectory over the integer count t is approximated by $p_i s[t]$, i.e., PageRank-scaled solution s of the blended dynamics (4.3.7). Therefore, if one is interested in solving the PageRank score of each node, then the blended dynamics can be designed to have a stable equilibrium of 1.

In fact, when the network has a large number of agents, these PageRank scores are not easy to be computed in a centralized manner. Thus, the *distributed* PageRank algorithms have been proposed in [IT10, ITB12, LC14]. Unfortunately, most of them commonly assume an initialization. However, when nodes are added to or removed from the network during the process of the algorithm, the whole algorithm must be re-initialized whenever a change occurs in the network, and this is not easy to be achieved in a distributed manner.

On the contrary, we can design an initialization-free distributed PageRank estimation algorithm by employing the proposed multi-step coupling framework. It can be easily inferred that, if the solution of the blended dynamics simply converges to 1, then every sampled state $x_i[t]$ will approach to its PageRank score p_i under the multi-step coupling of (4.3.6). Thus, we first design the blended dynamics which has a stable equilibrium point at 1 as

$$s[t+1] = \nu s[t] + (1-\nu), \tag{4.3.8}$$

where $\nu \in (0,1)$ is a design parameter. Since

$$\nu s[t] + (1-\nu) = \sum_{i=1}^{N} \{\nu p_i s[t] + (1-\nu)/N\},$$

we can divide (4.3.8) to each node by proposing the following algorithm

$$x_i[t_{k+1}] = \begin{cases} \nu x_i[t_k] + \dfrac{1-\nu}{N}, & \text{if } k = 0, \\ \theta^{\mathrm{PR}} x_i[t_k] + (1-\theta^{\mathrm{PR}}) \displaystyle\sum_{j \in \mathcal{N}_i} \frac{x_j[t_k]}{d_j^{\mathrm{out}}}, & \text{otherwise.} \end{cases} \tag{4.3.9}$$

Indeed, the proposed algorithm has the blended dynamics of (4.3.8). As stated in Theorem 4.2.2 or Corollary 4.2.5, the proposed distributed algorithm does not rely on a particular initialization. The algorithm (4.3.9) uses a global information of N, but it can be distributively estimated by the result in Section 4.3.1. Note that the proposed algorithm (4.3.9) could estimate the PageRank scores in the strongly connected network as Assumption 4.2.2. If the connectivity of the network is not guaranteed such as in a hyperlinked network of web-pages, then we need a new definition of the PageRank scores and modified algorithm, which will be introduced in Chapter 5.

4.3.3 Distributed Estimation of Degree Sequence of Network

Average consensus protocol has been widely utilized in many discrete-time synchronization problems including [OSFM07, RB05]. Thus, in this subsection, we consider a multi-step coupling framework whose coupling dynamics (4.2.1b) is the following average consensus protocol

$$x_i[t_{k+1}] = \theta^{\mathrm{avg}} x_i[t_k] + \frac{1-\theta^{\mathrm{avg}}}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} x_j[t_k],$$

where $\theta^{\mathrm{avg}} \in (0,1)$ is a parameter which represents weight between its own state and the average of the neighbors.

Then, the coupling weight w_{ij}^{avg} is obtained as

$$w_{ij}^{\mathrm{avg}} = \begin{cases} \theta^{\mathrm{avg}}, & i = j, \\ (1-\theta^{\mathrm{avg}})\dfrac{a_{ij}}{d_i}, & \text{otherwise}, \end{cases}$$

and the weight matrix $W^{\mathrm{avg}} := [w_{ij}^{\mathrm{avg}}]$ is given by

$$W^{\mathrm{avg}} = \theta^{\mathrm{avg}} I + (1-\theta^{\mathrm{avg}}) D^{-1} A,$$

where D is the diagonal matrix whose diagonal components are $d_1, \dots, d_N$ sequentially. Note that w_{ij}^{avg} satisfies (4.2.3) and Assumption 4.2.1 holds because W^{avg} is row-stochastic matrix. Thus, we can choose $p = \mathbb{1}_N$ and find a positive vector q by Lemma 4.2.1 such that

$$q^\top W^{\mathrm{avg}} = q^\top, \quad q^\top \mathbb{1}_N = 1.$$

Now, the blended dynamics is given by

$$s[t+1] = \sum_{i=1}^{N} q_i f_i(t, s[t]). \tag{4.3.10}$$

Meanwhile, if the network under consideration is undirected, q is easily obtained as $q = (1/d_{\mathrm{sum}})d$ for $p = \mathbb{1}_N$ where $d_{\mathrm{sum}} := \sum_{j=1}^{N} d_j$ and $d = \mathrm{col}\,(d_1, \cdots, d_N) \in \mathbb{R}^N$ because $d^\top D^{-1} A = \mathbb{1}_N^\top A = [d_1^{\mathrm{out}}, \dots, d_N^{\mathrm{out}}] = [d_1, \dots, d_N] = d^\top$. From this, the blended dynamics (4.3.10) is rewritten as

$$s[t+1] = \frac{1}{d_{\mathrm{sum}}} \sum_{i=1}^{N} d_i f_i(t, s[t]). \tag{4.3.11}$$

Similarly with Section 4.3.1, the overall trajectories of $x_i[t]$ are approximately synchronized to the solution of blended dynamics (4.3.10) (or (4.3.11)) for sufficiently large K. The difference between the doubly-stochastic and row-stochastic weight matrix is that the former has the blended dynamics as the simple average of f_is like (4.3.2), while the latter has the blended dynamics as the weighted av-

erage like (4.3.10) whose weights q_is are uneven in general.

For undirected graphs, a non-increasing sequence of all degrees is called as *degree sequence* [Die17]. Since the degree sequence does not uniquely identify a graph, there has been much attention to obtain information of the graph structure from the given degree sequence. For example, [VL05] realized the given degree sequence by a simple graph (realization problem) and [HP14] estimated the number of graphs with the given degree sequence (graph enumeration). If each agent can predict possible structures of the network with the degree sequence, it could obtain global information such as the algebraic connectivity. To achieve this, a distributed algorithm to estimate the degree sequence is required, and we proposed one implementation by employing the proposed framework as follows.

In the proposed algorithm, we additionally assume that each agent knows the network size N and has its unique id. Again, N can be distributively estimated using the application example in Section 4.3.1. Moreover, the assumption on the unique id for each agent is quite natural in the sense that, in practice, every communication device has its own identifier such as mac address. Let $\mathcal{X}_i \in \mathbb{Z}$ be the id of the agent i and suppose $1 < \mathcal{X}_i$ for all $i \in \mathcal{N}$ without loss of generality.

Under the above assumptions, an arbitrary i-th agent runs the following dynamics

$$x_i[t_{k+1}] = \begin{cases} (1 - 1/d_i)\, x_i[t_k] + N^{\mathcal{X}_i}, & \text{if } k = 0, \\ \sum_{j \in \mathcal{N}_i \cup \{i\}} w_{ij}^{\mathrm{avg}} x_j[t_k], & \text{otherwise,} \end{cases}$$

where $x_i \in \mathbb{R}$ is the state. From (4.3.11), the blended dynamics is obtained as

$$\begin{aligned} s[t+1] &= \frac{1}{d_{\mathrm{sum}}} \left\{ \sum_{i=1}^{N} d_i \left(1 - \frac{1}{d_i}\right) s[t] + \sum_{i=1}^{N} d_i N^{\mathcal{X}_i} \right\} \\ &= \left(1 - \frac{N}{d_{\mathrm{sum}}}\right) s[t] + \frac{1}{d_{\mathrm{sum}}} \sum_{i=1}^{N} d_i N^{\mathcal{X}_i}. \end{aligned}$$

For any connected network, $d_{\mathrm{sum}} \geq N$, and this guarantees the contraction stability of the above blended dynamics. In addition, its equilibrium point s^* is

obtained as

$$\begin{aligned} s^* &= \frac{1}{N}\sum_{i=1}^{N} d_i N^{\mathcal{X}_i} \\ &= \sum_{i=1}^{N} d_i N^{\mathcal{X}_i - 1}. \end{aligned}$$

Meanwhile, it is easily seen that $d_i < N$ for all $i \in \mathcal{N}$ because the network has no self-connection. Thus, the number s^* can be regarded as a representation of the numeral system with the base N:

$$\left[s^*\right]_N = \left[\delta_{(B-1)} \quad \delta_{(B-2)} \quad \cdots \quad \delta_1 \quad \delta_0\right]_N, \tag{4.3.12}$$

where $B = \max_{i\in\mathcal{N}} \mathcal{X}_i$ and δ_b represents the $(b+1)$-th right-most digit such that $\delta_b = d_i$ if $b = \mathcal{X}_i - 1$ and $\delta_b = 0$ otherwise.

As a result, since every state at each integer count, $x_i[t]$, is approximately synchronized to s^* of (4.3.12), every agent can estimate the degree sequence by removing zero digits in the N-base numeral representation of its state and ordering the rest digits in a non-increasing order. Here, we suppose that K is properly chosen such that the synchronization error $\epsilon < 1$ because this guarantees that the error does not affect even in the right-most digit δ_0.

Chapter 5

Application to Initialization-free Distributed PageRank Estimation for Network of Web-pages

5.1 Problem Formulation

PageRank is an algorithm used in the Google's search engine to rank numerous web-pages [BP98]. It assigns numerous values, called *PageRank scores*, to all web-pages in a hyperlinked network, which represents the relative importance or popularity of each web-pages. This concept of the PageRank has attracted a lot of attention in diverse areas where there are multiple objects and their interconnections. For example, the PageRank algorithm has widely utilized in informatics [CXMR07], bibliometrics [LBNVdS05], biology [ZBE12], sports [Rad11] and other fields [Gle15].

The PageRank score is designed such that it has larger value as it receives more citations from other web-pages and as the citing pages have larger PageRank scores. To compute the scores, a challenge is that it requires significant computational loads due to the large dimension of the network. To overcome this problem, there have been extensive studies regarding *distributed* PageRank algorithms which operate only by local information for each page. Distributed randomized algorithms are proposed, for example, in [IT10, ITB12, LC14] where each page builds its distributed hyperlink matrix using only its outgoing information and computes the PageRank score through a time average or a stochastic approxi-

mation. Recently, [SI19] extended this randomization approach to a deterministic update algorithm for multiple pages. In [LZD17], an optimization based distributed algorithm where the sum of the state values of all pages remains constant was proposed.

However, most of prior studies are commonly assume an initialization. A necessity of this initialization limits the algorithm's utility only for the time-invariant network. In other words, as pages or their hyperlinks are added or removed in the network during the operation of the algorithm (we call this *plug-and-play operation*), the algorithm must be re-initialized for each change in the network and this is hard to be achieved in a distributed manner.

In addition, many algorithms are designed assuming that every page knows the total number of pages N in the network. For example, the number of N is utilized as a parameter or the dimension of state variables. Not only the network size is a global information, but also using the number of N in the algorithm leads to challenge for plug-and-play operation. Although there has been various studies on distributed estimation of the network size in a multi-agent system [LLKS18, SCHJ12, BAMJ11], most of them cannot be applied for the PageRank problem. This is because the network in PageRank problem is different from other networks in the sense that the network of web-pages is not connected digraph in general.

To overcome this issue, [YTQ16] reformulated the PageRank problem as Least Square problem to distributively compute PageRank scores without specific initialization and the information of N. Nonetheless, the algorithm assumes that the time-varying network is strongly connected with the constant network size and there exists a global index for each page. Furthermore, to design of a random process for determining which pages are selected, a bidirectional communication is required in the directed graph structure.

In this chapter, we propose a distribute PageRank algorithm which does not require both the initialization and the information of the network size N. By forgetting an effect of initial conditions, the proposed algorithm enables the plug-and-play operation without any connectivity condition in the network or regardless of the network size. The trade-off of not employing N in the algorithm is

that each page estimates scaled PageRank scores which are rescaled by the network size N. However, since the relative ranking is not affected by the positive rescaling, an important information about PageRank scores is still retained. In addition, it is possible for each page to find the exact PageRank score if N is known.

This chapter is organized as follows. Section 5.2 introduces the basic definitions of PageRank and the iterative method to compute PageRank scores. In Section 5.3, the initialization-free distributed PageRank algorithm is proposed with the rigorous proof. In order to confirm aforementioned advantages of the proposed algorithm, Section 5.4 provides simulation results for a large dimensional real data on web-pages composed of more than 30,000 links and 3,700 pages.

5.2 Basic Definitions of PageRank for Teleportation Model

For a directed network $\mathcal{G}$ with N pages, the PageRank score $\mathsf{p}_i \geq 0$ for a page i is defined by

$$\mathsf{p}_i = \sum_{j \in \mathcal{N}_i} \frac{\mathsf{p}_j}{d_j^{\text{out}}} \quad \text{and} \quad \sum_{i=1}^{N} \mathsf{p}_i = 1 \tag{5.2.1}$$

where $\mathcal{N}_i$ is the index set of the pages that have outgoing link to page i (i.e., the in-neighbors of the node i) and d_j^{out} is the number of outgoing links from page j (i.e., the out-degree of the node j).

By stacking all the scores in a column vector $\mathsf{p} := \text{col}\,(\mathsf{p}_1, \cdots, \mathsf{p}_N) \in \mathbb{R}^N$ (we call this as *PageRank vector*), the definition of the PageRank (5.2.1) can be rewritten as

$$\mathsf{p} = \begin{bmatrix} a_{11}/d_1^{\text{out}} & a_{12}/d_2^{\text{out}} & \cdots & a_{1N}/d_N^{\text{out}} \\ a_{21}/d_1^{\text{out}} & a_{22}/d_2^{\text{out}} & \cdots & a_{2N}/d_N^{\text{out}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}/d_1^{\text{out}} & a_{N2}/d_2^{\text{out}} & \cdots & a_{NN}/d_N^{\text{out}} \end{bmatrix} \mathsf{p} = AD_{\text{out}}^{-1}\mathsf{p} \quad \text{and} \quad \mathbb{1}_N^\top \mathsf{p} = 1,$$

where A is the binary adjacency matrix of the network $\mathcal{G}$ and $D_{\text{out}} \in \mathbb{R}^{N \times N}$ is a diagonal matrix whose diagonal components are $d_1^{\text{out}}, \ldots, d_N^{\text{out}}$. Since $d_j^{\text{out}} =$

$\sum_{i=1}^{N} a_{ij}$, it should be emphasized that the matrix AD_{out}^{-1} is column-stochastic so that $\mathbb{1}_N^\top$ is a left eigenvector of AD_{out}^{-1} corresponding to the eigenvalue 1. This implies that the PageRank vector p can be considered as a scaled right eigenvector of AD_{out}^{-1} corresponding to the eigenvalue 1.

If the network $\mathcal{G}$ is strongly connected and aperiodic, then we can guarantee that AD_{out}^{-1} has a unique p by Lemma 2.2.2 and Lemma 2.2.4. Based on this observation, we already proposed an initialization-free distributed estimation algorithm for PageRank scores for the strongly connected network in Chapter 4.3.2.

However, asking the network of web-pages to be strongly connected is not realistic because there can be the pages that are isolated or have only outgoing/incoming hyperlinks. Therefore, we adopt a modified definition of the PageRank vector p as the solution to the following *teleportation model* [IT14]':

$$\mathsf{p} = \left[(1-m)AD_{\text{out}}^{-1} + \frac{m}{N}\mathbb{1}_N\mathbb{1}_N^\top\right]\mathsf{p} =: M\mathsf{p}, \quad \mathbb{1}_N^\top\mathsf{p} = 1. \tag{5.2.2}$$

where $m \in (0,1)$ is a parameter whose typical value is 0.15 [BP98]. More details related to this teleportation model can be found in [IT14]. It is clear that the associated graph of M is now strongly connected because each node is connected to every other node in the graph. Since M is a convex combination of two column-stochastic matrices, it is also column-stochastic matrix and $\rho(M) = 1$. Moreover, it can be easily seen that M is primitive for any graph topology of the network $\mathcal{G}$ by Lemma 2.2.4 because its associated graph is a complete digraph with a self-loop for each node. Using Lemma 2.2.2, this guarantees the uniqueness of the PageRank vector p as the scaled right eigenvector of M corresponding to the eigenvalue 1.

On the other hand, solving the linear equation (5.2.2) consumes significant computational power when the dimension of the network is large. One idea to overcome this problem is using a power method to iterate the following PageRank dynamics:

$$z[k+1] = Mz[k], \quad \mathbb{1}_N^\top z[0] = 1. \tag{5.2.3}$$

where $z = \text{col}\,(z_1, \cdots, z_N) \in \mathbb{R}^N$ is the state which estimates the PageRank vector

p. The state z asymptotically converges to the PageRank vector p because M is primitive.

Indeed, by Lemma 2.2.2, the matrix M has a simple eigenvalue 1 which is strictly larger than the absolute value of any other eigenvalues. Without loss of generality, we can denote the eigenvalues of M as $\lambda_i(M)$ for $i \in \mathcal{N}$ such that $1 = \lambda_1(M) > |\lambda_2(M)| \geq \cdots \geq |\lambda_N(M)|$. Moreover, there exist $R, Z \in \mathbb{R}^{N\times(N-1)}$ such that

$$M = \begin{bmatrix} \mathsf{p} & R \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \Lambda \end{bmatrix} \begin{bmatrix} \mathbb{1}_N^\top \\ Z^\top \end{bmatrix}, \tag{5.2.4}$$

$Z^\top R = I_{N-1}$, and $R^\top \mathbb{1}_N = Z^\top \mathsf{p} = \mathbb{0}_{N-1}$, where $\Lambda \in \mathbb{R}^{(N-1)\times(N-1)}$ is a matrix whose eigenvalues are $\lambda_2(M), \ldots, \lambda_N(M)$. Therefore, we have

$$\begin{aligned} \lim_{k\to\infty} z[k] &= M^k z[0] \\ &= \lim_{k\to\infty} \begin{bmatrix} \mathsf{p} & R \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \Lambda^k \end{bmatrix} \begin{bmatrix} \mathbb{1}_N^\top \\ Z^\top \end{bmatrix} z[0] \\ &= \mathbb{1}_N^\top z[0] \mathsf{p} = \mathsf{p}, \end{aligned} \tag{5.2.5}$$

where the last equality follows from $\sum_{i=1}^N z_i[0] = 1$. Thus, this initialization is essential for each state z_i to converge to its PageRank score p_i.

When the network is time-invariant, the sum of all the states is kept as 1 as long as the iterative method is initialized because $\mathbb{1}_N^\top z[k] = \mathbb{1}_N^\top M^k z[0] = \mathbb{1}_N^\top x[0] = 1$ for all $k > 0$. However, whenever some pages join or leave the network during the iteration of (5.2.3), the sum might vary from 1 by the amount of the joining/leaving state so that the convergence to the PageRank vector p in (5.2.5) is not guaranteed in general. This implies the power method (5.2.3) requires the re-initialization process for each plug-and-play operation, which is difficult to be achieved in a distributed manner.

5.3 Distributed PageRank Estimation without Initialization

In this section, we propose an initialization-free distributed PageRank algorithm, where each page estimates a scaled PageRank score $N\mathsf{p}_i$ (rather than p_i). There are several reasons why we aim to estimate N-times scaled PageRank scores instead of exact PageRank scores. Firstly, this positive rescaling does not affect the relative ranking among all pages. Next, this enables the algorithm not to contain a global information such as N so that its implementation becomes fully distributed. Finally, even if the network varies over time, each pages can estimate the exact PageRank score in the proposed algorithm only if N is known, while the power method (5.2.3) cannot.

In the proposed algorithm, an individual dynamics of an arbitrary page i is given by adopting the multi-step coupling framework (4.2.1) as follows:

$$x_i[t_{k+1}] = \begin{cases} \nu x_i[t_k] + (1-\nu) =: f_i^{\mathrm{PR}}(t_k, x_i[t_k]), & \text{if } k = 0, \quad (5.3.1a) \\ (1-m)\displaystyle\sum_{j\in\mathcal{N}_i} \frac{x_j[t_k]}{d_j^{\mathrm{out}}} + m, & \text{if } k = 1,\dots,K-1, \quad (5.3.1b) \end{cases}$$

where $x_i \in \mathbb{R}$ is the state, $\nu \in (0,1)$ is constant as a design parameter, and t_k is the fractional time-index in (4.2.2).

The underlying insight behind the proposed algorithm is motivated by the (discrete-time) blended dynamics approach in Chapter 4. However, unlike the coupling dynamics (4.2.1b) where each state is updated by a weighted average of neighbors' state, i.e., a linear combination of agents' state information, (5.3.1b) in the proposed algorithm cannot be represented as the linear combination form due to a constant term m.

So, it is required to modify the previous discrete-time blended dynamics (4.2.5) to predict the behavior of (5.3.1) and we insist that the discrete-time blended

dynamics for (5.3.1) can be defined as follows:

$$\begin{aligned} s[t+1] &= (1-m)^{K-1} \sum_{i=1}^{N} f_i^{\mathrm{PR}}(t, \mathsf{p}_i s[t]) + N(1-(1-m)^{K-1}) \\ &= (1-m)^{K-1} \left\{ \nu \sum_{i=1}^{N} \mathsf{p}_i s[t] + N(1-\nu) \right\} + N(1-(1-m)^{K-1}) \\ &= \nu(1-m)^{K-1}(s[t]-N) + N, \end{aligned} \tag{5.3.2}$$

where t is the integer count and $\mathsf{p} = [\mathsf{p}_1, \cdots, \mathsf{p}_N]^\top$ is the PageRank vector defined in (5.2.2). It should be emphasized that (5.3.2) has a unique stable equilibrium point at N and its convergence rate $\nu(1-m)^{K-1}$ can be arbitrarily small by increasing K. Then, we can infer from Theorem 4.2.2 that, if the new blended dynamics (5.3.2) is stable, then, for any initial conditions $x_i[0]$ and $s[0]$, every solution x_i at each integer count t can be approximately predicted by the p_i-times scaled solution s, i.e., $N\mathsf{p}_i$ for all $i \in \mathcal{N}$.

Indeed, the following main result of this chapter states that the proposed algorithm approximately estimates each scaled PageRank score for any initial condition. In particular, the proposed algorithm does not rely on initial conditions so that it enables the plug-and-play operation and the precision of the algorithm gets higher as K increases.

Theorem 5.3.1. For any $\epsilon > 0$, there exists $K^{\min}$ such that, for each $K > K^{\min}$, the solution x_i of (5.3.1) and the solution s of (5.3.2) with arbitrary initial conditions satisfy

$$\limsup_{t\to\infty} \left| x_i[t] - N\mathsf{p}_i \right| \le \epsilon, \quad \forall i \in \mathcal{N}. \qquad \Diamond$$

Proof. For convenience, let $\bar{x} := \mathrm{col}\,(\bar{x}_1, \cdots, \bar{x}_N)$ and define

$$\begin{aligned} H(\bar{x}) &= (1-m)AD_{\mathrm{out}}^{-1}\bar{x} + m\mathbb{1}_N, \\ H^k(\bar{x}) &= \underbrace{H \circ \cdots \circ H}_{k \text{ times}}(\bar{x}), \\ F(\bar{x}) &= \nu\bar{x} + (1-\nu)\mathbb{1}_N. \end{aligned}$$

Then, the evolution of the proposed algorithm (5.3.1) at each integer count t is rewritten as

$$\bar{x}[t+1] = H^{K-1}(F(\bar{x}[t])). \tag{5.3.3}$$

Based on the decomposition of M in (5.2.4), we consider the following coordinate transformation

$$\zeta = \begin{bmatrix} \zeta_1 \\ \tilde{\zeta} \end{bmatrix} = \begin{bmatrix} \mathbb{1}_N^\top \\ Z^\top \end{bmatrix} \bar{x}$$

whose inverse is

$$\bar{x} = \mathsf{p}\zeta_1 + R\tilde{\zeta}.$$

Along the overall dynamics (5.3.3), the dynamics of ζ_1 at each integer count t is given by

$$\begin{aligned}
\zeta_1[t+1] &= \mathbb{1}_N^\top H^{K-1}(F(\mathsf{p}\zeta_1[t] + R\tilde{\zeta}[t])) \\
&= (1-m)^{K-1}\mathbb{1}_N^\top F(\mathsf{p}\zeta_1[t] + R\tilde{\zeta}[t]) + Nm\sum_{l=0}^{K-2}(1-m)^l \\
&= (1-m)^{K-1}\{\nu\zeta_1[t] + N(1-\nu)\} + N\left\{1-(1-m)^{K-1}\right\} \\
&= (1-m)^{K-1}\nu(\zeta_1[t]-N) + N,
\end{aligned}$$

where we used $\mathbb{1}_N^\top \mathsf{p} = 1$ and $\mathbb{1}_N^\top R = 0$. This implies that $\zeta_1[t]$ converges to N as t goes to ∞ from any initial condition.

On the other hand, the $\tilde{\zeta}$ dynamics at each integer count t is obtained as

$$\begin{aligned}
\tilde{\zeta}[t+1] &= Z^\top H^{K-1}(F(\mathsf{p}\zeta_1[t] + R\tilde{\zeta}[t])) \\
&= Z^\top \left((1-m)AD_{\text{out}}^{-1}\right)^{K-1}(F(\mathsf{p}\zeta_1[t] + R\tilde{\zeta}[t]) - N\mathsf{p}), \\
&= Z^\top \left((1-m)AD_{\text{out}}^{-1}\right)^{K-1}\left\{\nu(\mathsf{p}\zeta_1[t] + R\tilde{\zeta}[t]) + (1-\nu)\mathbb{1}_N - N\mathsf{p}\right\},
\end{aligned}$$

because, for any vector $v \in \mathbb{C}^N$,

$$Z^\top H(v) = Z^\top\left\{\left(M - \frac{m}{N}\mathbb{1}_N\mathbb{1}_N^\top\right)v + m\mathbb{1}_N\right\}$$

$$
\begin{aligned}
&= \Lambda Z^\top (v - N\mathsf{p}) - \frac{m}{N} Z^\top \mathbb{1}_N \mathbb{1}_N^\top (v - N\mathsf{p}) \\
&= Z^\top \left(M - \frac{m}{N} \mathbb{1}_N \mathbb{1}_N^\top \right) (v - N\mathsf{p}),
\end{aligned}
$$

where we use $Z^\top M = \Lambda Z^\top$, $Z^\top \mathsf{p} = \mathbb{0}_{N-1}$ and $\mathbb{1}_N^\top \mathsf{p} = 1$ for the second equality, and

$$
\begin{aligned}
H(v) - N\mathsf{p} &= \left(M - \frac{m}{N} \mathbb{1}_N \mathbb{1}_N^\top \right) v + m\mathbb{1}_N - N\mathsf{p} \\
&= M(v - N\mathsf{p}) - \frac{m}{N} \mathbb{1}_N \mathbb{1}_N^\top (v - N\mathsf{p}) \\
&= \left(M - \frac{m}{N} \mathbb{1}_N \mathbb{1}_N^\top \right) (v - N\mathsf{p}),
\end{aligned}
$$

where the second equality comes from $M\mathsf{p} = \mathsf{p}$ and $\mathbb{1}_N^\top \mathsf{p} = 1$, so that, for example,

$$
\begin{aligned}
Z^\top H^2(v) &= Z^\top (M - (m/N)\mathbb{1}_N \mathbb{1}_N^\top)(H(v) - N\mathsf{p}) \\
&= Z^\top (M - (m/N)\mathbb{1}_N \mathbb{1}_N^\top)^2 (v - N\mathsf{p}).
\end{aligned}
$$

As a result, by letting $e := \zeta_1 - N$, the system (5.3.3) is transformed into

$$
\begin{aligned}
e[t+1] &= (1-m)^{K-1} \nu e[t], \\
\tilde{\zeta}[t+1] &= Z^\top \left((1-m) A D_{\text{out}}^{-1} \right)^{K-1} \\
&\qquad \times \left\{ \nu \mathsf{p} e[t] + \nu R \tilde{\zeta}[t] + (1-\nu)(\mathbb{1}_N - N\mathsf{p}) \right\}.
\end{aligned} \tag{5.3.4}
$$

Now, we define a Lyapunov function $V = |e| + \|\tilde{\zeta}\|$. Then, its time difference along (5.3.4) is

$$
\begin{aligned}
&V[t+1] - V[t] \\
&= |e[t+1]| + \|\tilde{\zeta}[t+1]\| - |e[t]| - \|\tilde{\zeta}[t]\| \\
&\leq -(1 - (1-m)^{K-1}\nu)|e[t]| - \|\tilde{\zeta}[t]\| \\
&\quad + \|Z^\top \left((1-m) A D_{\text{out}}^{-1} \right)^{K-1} \\
&\qquad \times \{ \nu \mathsf{p} e[t] + \nu R \tilde{\zeta}[t] + (1-\nu)(\mathbb{1}_N - N\mathsf{p}) \} \|
\end{aligned} \tag{5.3.5}
$$

Since $\rho((1-m)AD_{\text{out}}^{-1}) = 1-m$,

$$\begin{aligned}
&\|Z^\top \left((1-m)AD_{\text{out}}^{-1}\right)^{K-1} \\
&\qquad \times \{\nu \mathsf{p} e[t] + \nu R\tilde{\zeta}[t] + (1-\nu)(\mathbb{1}_N - N\mathsf{p})\}\| \\
&\le (1-m)^{K-1}\nu\|Z\|(\|\mathsf{p}e[t]\| + \|R\tilde{\zeta}[t]\|) \\
&\quad + (1-m)^{K-1}(1-\nu)\|Z\|\|\mathbb{1}_N - N\mathsf{p}\| \\
&\le (1-m)^{K-1}\nu\|Z\|(|e[t]| + \|R\|\|\tilde{\zeta}[t]\|) \\
&\quad + 2(1-m)^{K-1}(1-\nu)N\|Z\|,
\end{aligned} \tag{5.3.6}$$

where the last inequality comes from $\|p\| \le 1$ and $\|1_N - Np\| \le \|1_N\| + N\|p\| \le 2N$.

Substituting (5.3.6) into (5.3.5), it is obtained

$$\begin{aligned}
V[t+1] - V[t] &\le -(1-(1-m)^{K-1}\nu)|e[t]| - \|\tilde{\zeta}[t]\| \\
&\quad + (1-m)^{K-1}\nu\|Z\|(|e[t]| + \|R\|\|\tilde{\zeta}[t]\|) \\
&\quad + 2(1-m)^{K-1}(1-\nu)N\|Z\| \\
&\le -(1-(1-m)^{K-1}\nu)V[t] \\
&\quad + (1-m)^{K-1}\nu\|Z\|\max\{1,\|R\|\}V[t] \\
&\quad + 2(1-m)^{K-1}(1-\nu)N\|Z\|.
\end{aligned} \tag{5.3.7}$$

For given ϵ, let $K^{\min}$ be a positive integer such that

$$(1-m)^{K^{\min}}\nu\{1 + \|Z\|\max\{1,\|R\|\}\} \le \frac{1}{2} \tag{5.3.8}$$

$$4(1-m)^{K^{\min}}(1-\nu)N\max\{1,\|R\|\}\|Z\| \le \epsilon \tag{5.3.9}$$

Then, it follows from (5.3.7) by (5.3.8) that, for every $K > K^{\min}$,

$$V[t+1] - V[t] \le -\frac{1}{2}V[t] + 2(1-m)^{K-1}(1-\nu)N\|Z\|.$$

From this, the ultimate bound of V is obtained as

$$\limsup_{t\to\infty} V[t] \le 4(1-m)^{K-1}(1-\nu)N\|Z\|.$$

Therefore, for each agent $i \in \mathcal{N}$ and $K > K^{\min}$,

$$\begin{aligned}
\limsup_{t\to\infty} |x_i[t] - N\mathsf{p}_i| &\le \limsup_{t\to\infty} |\mathsf{p}_i\zeta_1[t] + R_i\tilde{\zeta}[t] - N\mathsf{p}_i| \\
&\le \limsup_{t\to\infty} \{\mathsf{p}_i|e[t]| + \|R\|\|\tilde{\zeta}[t]\|\} \\
&\le \max\{1, \|R\|\} \limsup_{t\to\infty} V[t] \\
&\le \max\{1, \|R\|\}\, 4(1-m)^{K-1}(1-\nu)N\|Z\| \\
&\le \epsilon,
\end{aligned} \tag{5.3.10}$$

where R_i is the ith row of R and the last inequality comes from (5.3.9). This completes the proof. □

5.4 Simulation Results

For simulation, we use real data on web-pages of Lincoln university in New Zealand, as in [IT14]. This web consists of 31,718 links and 3,756 pages including two pages without any incoming link. Since the two pages do not play any role in PageRank, they were removed in the data. Most of the rest pages do not have an outgoing link, which are called as dangling node. In order to make the matrix AD_{out}^{-1} stochastic, for all dangling node i, an outgoing link (i, j) was added whenever (j, i) is an incoming link. As a result, the resulting network has 3,764 pages with 40,646 links. More details on the data can be found in [IT14].

In order to confirm the performance of the proposed algorithm for the plug-and-play feature, the network is supposed to be varying twice during the operation of the proposed algorithm. In particular, two pages whose indices are 2308 and 2310, respectively, and all their associated incoming and outgoing links are removed at $t = 3$. Again, other two pages whose indices are 323 and 2306, respectively, and all their links are removed at $t = 6$. The varying distributions of the PageRank scores in the network are depicted in Figure 5.1. After two changes in the network, the two largest PageRank scores around 500 and 2500 are notably increased, while others are slightly altered.

The proposed algorithm is simulated for two different parameters $K = 20$

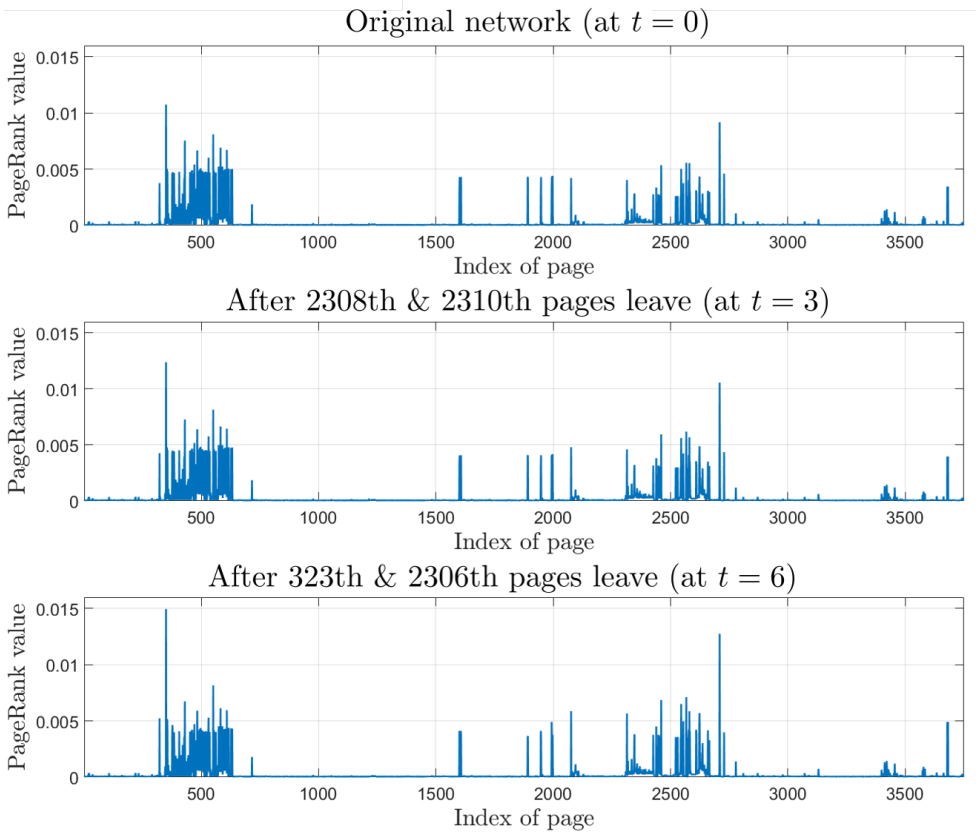

Figure 5.1: Distributions of PageRank scores for the original network (top) and changed networks (middle, bottom).

and $K = 40$ with the same initial condition. The initial conditions are set to be arbitrary for every agent. The parameters ν and m in (6) are set to be 0.1 and 0.15, respectively. Simulation results are represented in Figure 5.2. Trajectories of estimated PageRank scores in the proposed algorithm are depicted as colored solid curves, and true PageRank scores are given as black dashed lines.

The simulation result shows that the proposed algorithm estimates the true scaled PageRank scores for arbitrary initial condition, even for the plug-and-play operation. In addition, by comparing two results of $K = 20$ and $K = 40$, it can be seen that the level of error gets smaller by increasing K.

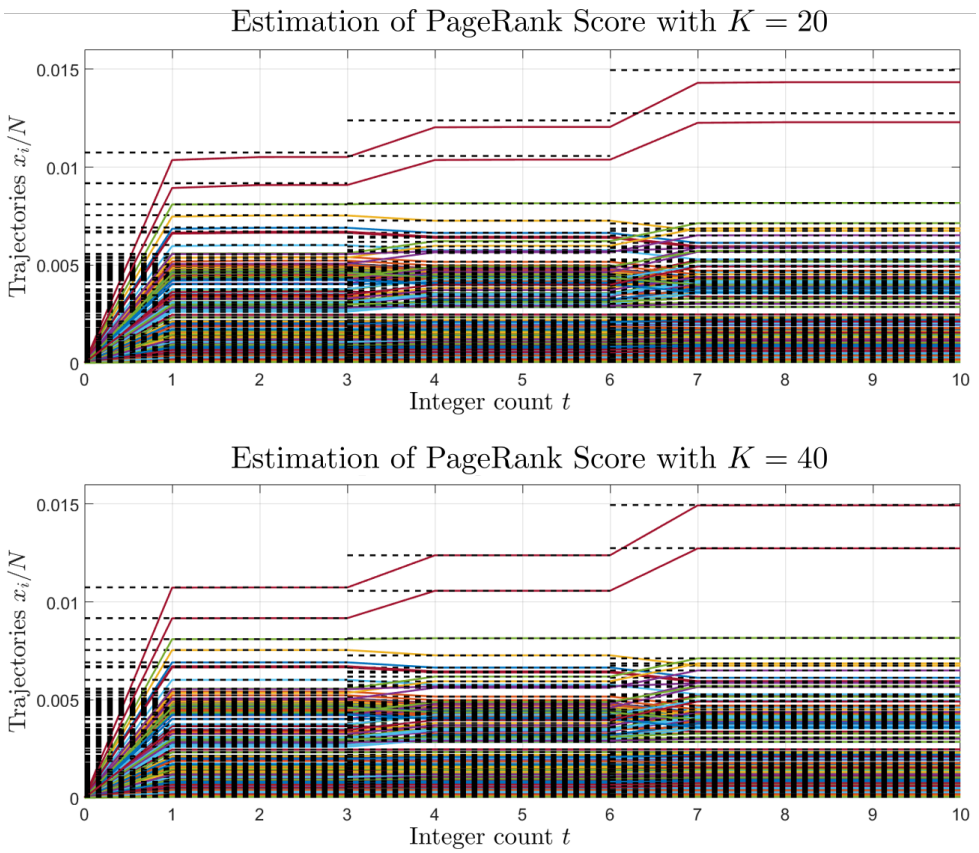

Figure 5.2: Estimated PageRank scores of the proposed algorithm with different parameters $K = 20$ (top) and $K = 40$ (bottom): Colored solid curves represent the estimation of each algorithm while black dashed lines represent true PageRank scores.

Chapter 6

Behavior of Discrete-time Heterogeneous Multi-agent System under Rank-deficient Coupling

6.1 Problem Formulation

This chapter studies the behavior of a discrete-time heterogeneous multi-agent system whose individual agent dynamics is given by

$$x_i[t_{k+1}] = \begin{cases} f_i(t_k, x_i[t_k]), & \text{if } k = 0, \quad (6.1.1a) \\ (I_n - B_i)x_i[t_k] + B_i \sum_{j \in \mathcal{N}_i \cup \{i\}} w_{ij} x_j[t_k], & \text{if } k = 1, \dots, K-1, \quad (6.1.1b) \end{cases}$$

where $x_i \in \mathbb{R}^n$ is the state, t_k is the fractional disrete-time index defined by (4.2.2), and $f_i : \mathbb{Z} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable function representing the time-varying heterogeneous node dynamics (6.1.1a). The coefficient w_{ij} is the coupling weight which satisfies the property (4.2.3). We assume the followings for the weight matrix $W := [w_{ij}]$.

Assumption 6.1.1. The weight matrix W is a stochastic matrix. ◊

Note that many studies on a discrete-time multi-agent system used a communication protocol whose weight matrix is a stochastic matrix such as in [IT10, ITB12, LC14, OSFM07, RB05]. In particular, if the weight matrix W is a row-stochastic

matrix, then the system (6.1.1) can be rewritten as

$$x_i[t_{k+1}] = \begin{cases} f_i(t_k, x_i[t_k]), & \text{if } k = 0, \quad (6.1.2a) \\ x_i[t_k] + B_i \sum_{j \in \mathcal{N}_i} w_{ij}(x_j[t_k] - x_i[t_k]), & \text{if } k = 1, \dots, K-1. \quad (6.1.2b) \end{cases}$$

so that the agents in the network are interconnected by the diffusive-type coupling [Hal97]. This type of the communication protocol has been frequently utilized in many synchronization problems [OSFM07, RB05, YSA19, LS22].

Here, B_i is supposed to be a symmetric positive semi-definite matrix so that only a partial information of the communication affects the update of x_i. Thus, we call this by *rank-deficient coupling*. Moreover, it is supposed that the spectral radius of every B_i is not larger than 1, i.e., $\rho(B_i) \leq 1$ for all $i \in \mathcal{N}$.

In fact, this work is closely related to [LS20] which extends the continuous-time blended dynamics approach, introduced in Chapter 3, to the systems whose coupling matrices are possibly all different and singular. In particular, this paper considers the heterogeneous multi-agent system whose individual dynamics is given by

$$\dot{\mathsf{x}}_i = \mathsf{f}_i(\mathsf{t}, \mathsf{x}_i) + \kappa \mathsf{B}_i \sum_{j \in \mathcal{N}_i} \alpha_{ij}(\mathsf{x}_j - \mathsf{x}_i), \quad (6.1.3)$$

where $\mathsf{x}_i \in \mathbb{R}^n$ is the state, κ is the coupling gain, α_{ij} is the ij-th element of the adjacency matrix $\mathcal{A}$, and $\mathsf{f}_i : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ is twice continuously differentiable with respect to their arguments, locally Lipschitz with respect to x_i uniformly in t, and $\mathsf{f}_i(\mathsf{t}, 0)$ is bounded. In this problem, it is also supposed that every B_i is positive semi-definite so that only a part of the integration of x_i is affected by other agents for all $i \in \mathcal{N}$.

Compared with B_i of (6.1.3), B_i is supposed to be not only a positive-semi definite matrix but also a stable matrix in the discrete-time sense (i.e., $\rho(B_i) \leq 1, \forall i \in \mathcal{N}$). To see the difference between the continuous-time rank-deficient coupling and the discrete-time rank-deficient coupling, let $\bar{x} := \mathrm{col}(x_1, \dots, x_N)$.

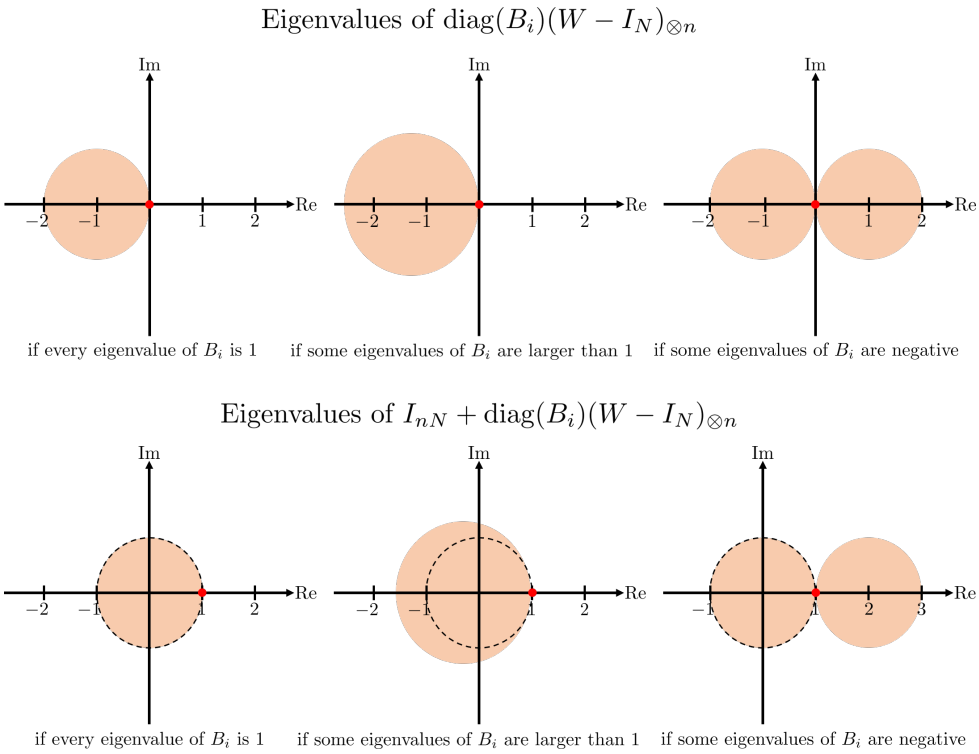

Figure 6.1: Possible locations of an eigenvalue of $\mathrm{diag}(B_i)(W - I_N)_{\otimes n}$ (top) and $I_{nN} + \mathrm{diag}(B_i)(W - I_N)_{\otimes n}$ (bottom): The colored area represents the possible location of all eigenvalues of the corresponding matrices and the dotted area represents the unit circle in the complex plane.

Then the overall system (6.1.1) at each integer count t is given by

$$\begin{aligned}
\bar{x}[t+1] &= \Big(\{I_{nN} - \mathrm{diag}(B_i)\} + \mathrm{diag}(B_i)(W \otimes I_n)\Big)^{K-1} \begin{bmatrix} f_1(t, x_1[t]) \\ \vdots \\ f_N(t, x_N[t]) \end{bmatrix} \\
&= \Big(I_{nN} + \mathrm{diag}(B_i)\{(W - I_N) \otimes I_n\}\Big)^{K-1} \begin{bmatrix} f_1(t, x_1[t]) \\ \vdots \\ f_N(t, x_N[t]) \end{bmatrix} \qquad (6.1.4) \\
&=: \Big(I_{nN} + \mathring{B}(W - I_N)_{\otimes n}\Big)^{K-1} F(t, \bar{x}[t]).
\end{aligned}$$

In order to guarantee the stability of the overall system (6.1.4), it is necessarily

required that the spectral radius of $I_{nN} + \mathrm{diag}(B_i)(W - I_N)_{\otimes n}$ is not larger than 1, i.e., $\rho(I_{nN} + \mathrm{diag}(B_i)(W - I_N)_{\otimes n}) \leq 1$. Equivalently, all eigenvalues of $\mathrm{diag}(B_i)(W - I_N)_{\otimes n}$ should be contained in the unit circle whose center is located at $(-1, 0)$ in the complex plane. It is easily known that every eigenvalue of B_i is real because B_i is symmetric for all $i \in \mathcal{N}$. In addition, every eigenvalue of $(W - I_N)_{\otimes n}$ is located in the unit circle whose center is located at $(-1, 0)$ on the complex plane because W is a stochastic matrix. If some eigenvalues of B_i are larger than 1 or negative, then some eigenvalues of $I_{nN}+\mathrm{diag}(B_i)(W-I_N)_{\otimes n}$ could be located in the outside of the unit circle in the complex plane (see Figure 6.1). This observation justifies $\rho(B_i) \leq 1$ for all $i \in \mathcal{N}$.

In this chapter, our interest is to analyze the behavior of the multi-agent system (6.1.1) under rank-deficient coupling when K is sufficiently large. To achieve this, similarly with Chapter 4, we introduce the discrete-time blended dynamics for rank-deficient coupling, which enables to predict the behavior of (6.1.1).

This chapter is organized as follows. Section 6.2 proposes the coordinate change for the multi-agent system under rank-deficient coupling, (6.1.1). In particular, by the proposed coordinate change, the system is separated into non-vanishing and vanishing dynamics with respect to the parameter K. Based on the observations, the blended dynamics for the rank-deficient coupling is introduced and this allows to predict an emergent behavior of the heterogeneous multi-agent system in Section 6.3.

6.2 Coordinate Change

Recalling (4.2.13) and (4.2.14) in Chapter 4, we propose the coordinate change which enables to separate the system into non-vanishing and vanishing parts as K gets larger. This allows to construct the blended dynamics which predicts the behavior of overall system for large K. Likewise, in this section, we introduce a linear coordinate change for (6.1.1) to separate the non-vanishing and vanishing dynamics with respect to the parameter K. In fact, a similar work is demonstrated in [LS20] for an undirected connected network in continuous-time domain. This

section introduces the coordinate change for a strongly connected directed network in discrete-time domain and this work is achieved with the aid of Jin Gyu Lee, lead author of [LS20].

The main result of this chapter is stated under the following assumptions.

Assumption 6.2.1. The function $f_i(t, x)$ is uniformly bounded in t and globally Lipschitz with respect to x uniformly in t.

Assumption 6.2.2. The communication network $\mathcal{G}$ is strongly connected. ◊

Under Assumption 6.2.2, by using Lemma 4.2.1, there exist positive eigenvectors $p, q \in \mathbb{R}^N$ such that

$$Wp = p, \quad q^\top W = q^\top, \quad q^\top p = 1. \tag{6.2.1}$$

Since W is a stochastic matrix, either p or q can be chosen as $\mathbb{1}_N$.

Before changing the coordinates of the system, a few matrices which consist of the transformation are introduced as follows:

1. For every positive semi-definite matrix B_i in (6.1.1), there exist $R_i \in \mathbb{R}^{n\times \mathrm{r}_i}$, $Z_i \in \mathbb{R}^{n\times(n-\mathrm{r}_i)}$ and a positive definite matrix $\Lambda_i \in \mathbb{R}^{\mathrm{r}_i\times \mathrm{r}_i}$, where r_i is the rank of B_i, such that $\begin{bmatrix} R_i & Z_i \end{bmatrix}$ is orthonormal matrix and

$$B_i = \begin{bmatrix} R_i & Z_i \end{bmatrix} \begin{bmatrix} \Lambda_i^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R_i^\top \\ Z_i^\top \end{bmatrix}.$$

It should be noted that the spectral radius of Λ_i is not greater than 1, i.e., $\rho(\Lambda_i) \leq 1$ because $\rho(B_i) \leq 1$ for all $i \in \mathcal{N}$. Let us define each block diagonal matrix of R_i, Z_i, and Λ_i for future use as follows:

$$\mathring{R} := \mathrm{diag}(R_1, \cdots, R_N)$$
$$\mathring{Z} := \mathrm{diag}(Z_1, \cdots, Z_N)$$
$$\mathring{\Lambda} := \mathrm{diag}(\Lambda_1, \cdots, \Lambda_N).$$

2. With $\bar{\mathrm{r}} := \sum_{i=1}^{N} \mathrm{r}_i$, there exists $V \in \mathbb{R}^{\bar{\mathrm{r}} \times \mathrm{r}_0}$ such that the columns of V are orthonormal vectors which satisfies

$$(W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi V = O_{nN \times \mathrm{r}_0} \tag{6.2.2}$$

where r_0 be the nullity of $(W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi$ and

$$\Xi = \mathrm{diag}(\Xi_1, \cdots, \Xi_N) = \mathrm{diag}\left(\sqrt{\frac{p_i}{q_i}} I_{\mathrm{r}_i} \right) \in \mathbb{R}^{\bar{\mathrm{r}} \times \bar{\mathrm{r}}}.$$

Let $V_i \in \mathbb{R}^{\mathrm{r}_i \times \mathrm{r}_0}$ for $i \in \mathcal{N}$ such that $V = \mathrm{col}(V_1, \cdots, V_N)$.

3. There exists $\bar{V} \in \mathbb{R}^{\bar{\mathrm{r}} \times (\bar{\mathrm{r}} - \mathrm{r}_0)}$ such that $\begin{bmatrix} V & \bar{V} \end{bmatrix}$ is an orthonormal matrix.

Based on aforementioned matrices, the following proposition can be shown.

Proposition 6.2.1.

$$(W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} V = O_{nN \times \mathrm{r}_0} \tag{6.2.3}$$

$\lozenge$

Proof. Since $\ker(W - I_N) = \mathrm{im}(p)$, it follows

$$\begin{aligned}
&\ker((W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi) = \{ v \in \mathbb{R}^{\bar{\mathrm{r}}} : \mathring{R} \mathring{\Lambda} \Xi v \in \mathrm{im}(p \otimes I_n) \} \\
&= \left\{ \mathrm{col}(v_1, \cdots, v_N) \in \mathbb{R}^{\bar{\mathrm{r}}} : \frac{R_1 \Lambda_1 \Xi_1 v_1}{p_1} = \cdots = \frac{R_N \Lambda_N \Xi_N v_N}{p_N} \right\} \\
&= \left\{ \mathrm{col}(v_1, \cdots, v_N) \in \mathbb{R}^{\bar{\mathrm{r}}} : \frac{R_1 \Lambda_1 v_1}{\sqrt{p_1 q_1}} = \cdots = \frac{R_N \Lambda_N v_N}{\sqrt{p_N q_N}} \right\}
\end{aligned}$$

where the last equality comes from $\Xi_i = \sqrt{p_i / q_i} I_{\mathrm{r}_i}$ for all $i \in \mathcal{N}$. This implies that, for any $\mathrm{col}(v_1, \cdots, v_N) \in \ker((W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi)$, there exists $w \in \mathbb{R}^n$ such that

$$R_i \Lambda_i v_i = \sqrt{p_i q_i} w, \quad i \in \mathcal{N}. \tag{6.2.4}$$

Meanwhile, by letting $v = \mathrm{col}(v_1, \cdots, v_N) \in \ker((W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi)$, it follows

$$\begin{aligned}\mathring{R}\mathring{\Lambda}\Xi^{-1}v &= \begin{bmatrix} \sqrt{\frac{q_1}{p_1}}R_1\Lambda_1 v_1 \\ \vdots \\ \sqrt{\frac{q_N}{p_N}}R_N\Lambda_N v_N \end{bmatrix} \\ &= \begin{bmatrix} q_1 w \\ \vdots \\ q_N w \end{bmatrix} \in \mathrm{im}(q \otimes I_n),\end{aligned}$$

where the last equality comes from (6.2.4).

Since $\ker(W^\top - I_N) = \mathrm{im}(q)$, we finally have

$$(W^\top - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi^{-1}V = O_{nN\times \mathrm{r}_0},$$

which completes the proof. □

Now, we consider the following coordinate transformation

$$\begin{bmatrix} z \\ r \\ \tilde{r} \end{bmatrix} = \begin{bmatrix} \mathring{Z}^\top \\ V^\top \Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top \\ S^{-1}\bar{V}^\top \Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top (W - I_N)_{\otimes n} \end{bmatrix} \bar{x} =: \mathcal{T}\bar{x} \tag{6.2.5}$$

where $z \in \mathbb{R}^{nN-\bar{\mathrm{r}}}$, $r \in \mathbb{R}^{\mathrm{r}_0}$, $\tilde{r} \in \mathbb{R}^{\bar{\mathrm{r}}-\mathrm{r}_0}$ and

$$S := \bar{V}^\top \Xi^{-1}\mathring{\Lambda}\mathring{R}^\top (W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V} \in \mathbb{R}^{(\bar{\mathrm{r}}-\mathrm{r}_0)\times(\bar{\mathrm{r}}-\mathrm{r}_0)}. \tag{6.2.6}$$

With (6.2.5), we claim that

$$\mathcal{T}^{-1} = \begin{bmatrix} \mathring{Z} - \mathring{R}\mathring{\Lambda}\Xi L & \mathring{R}\mathring{\Lambda}\Xi V & \mathring{R}\mathring{\Lambda}\Xi\bar{V} \end{bmatrix} \tag{6.2.7}$$

where

$$L := \bar{V}S^{-1}\bar{V}^\top \Xi\mathring{\Lambda}\mathring{R}^\top (W - I_N)_{\otimes n}\mathring{Z}.$$

Indeed, this claim can be easily proved by $\mathcal{T}\mathcal{T}^{-1} = I_{nN}$ using $\mathring{Z}^\top\mathring{R} = O$, $\mathring{Z}^\top\mathring{Z} = I$, $\mathring{R}^\top\mathring{R} = I$, $V^\top V = I$, $\bar{V}^\top\bar{V} = I$, and $V^\top\bar{V} = O$ (here, we omit the dimension of

identity matrices or the zero matrices for convenience). In particular, the $(3,1)$-th block component of $\mathcal{T}\mathcal{T}^{-1}$ is obtained by

$$\begin{aligned}
&S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\times(\mathring{Z}-\mathring{R}\mathring{\Lambda}\Xi L)\\
&=S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{Z}\\
&\quad-S^{-1}\underbrace{\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}}_{=S}S^{-1}\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{Z}\\
&=O,
\end{aligned}$$

and, the $(3,2)$-th block component is obtained from (6.2.2) as

$$S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi V=O.$$

Before applying coordinate transformation (6.2.5) for the system (6.1.4), we first confirm that

$$\begin{aligned}
\mathring{Z}^\top\Big(I_{nN}+\mathring{B}(W-I_N)_{\otimes n}\Big)&=\mathring{Z}\\
V^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top\Big(I_{nN}+\mathring{B}(W-I_N)_{\otimes n}\Big)&=V^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top\\
&\quad+V^\top\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\\
&=V^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top
\end{aligned}$$

where we use, for all $i\in\mathcal{N}$, $Z_i^\top B_i=0$, $R_i^\top B_i=\Lambda_i^2R_i$, and $(W^\top-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi^{-1}V=O$ from Proposition 6.2.1, and

$$\begin{aligned}
&S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\Big(I_{nN}+\mathring{B}(W-I_N)_{\otimes n}\Big)\\
&=S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\\
&\quad+S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi(VV^\top+\bar{V}\bar{V}^\top)\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\\
&=S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}\\
&\quad+S^{-1}\underbrace{\bar{V}^\top\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}}_{=S}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top(W-I_N)_{\otimes n}\\
&=(I+S)S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W-I_N)_{\otimes n}
\end{aligned}$$

where the first equality comes from

$$\begin{aligned}\mathring{B} &= \mathring{R}\mathring{\Lambda}^2\mathring{R}^\top \\ &= \mathring{R}\mathring{\Lambda}\Xi\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top \\ &= \mathring{R}\mathring{\Lambda}\Xi(VV^\top + \bar{V}\bar{V}^\top)\Xi^{-1}\mathring{\Lambda}\mathring{R}^\top.\end{aligned}$$

Using above equalities, by the coordinate change (6.2.5), the system (6.1.4) is transformed into

$$\begin{aligned}z[t+1] &= \mathring{Z}^\top F(t, \mathring{Z}z[t] - \mathring{R}\mathring{\Lambda}\Xi Lz[t] + \mathring{R}\mathring{\Lambda}\Xi Vr[t] + \mathring{R}\mathring{\Lambda}\Xi\bar{V}\tilde{r}[t]), \\ r[t+1] &= V^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top F(t, \mathring{Z}z[t] - \mathring{R}\mathring{\Lambda}\Xi Lz[t] + \mathring{R}\mathring{\Lambda}\Xi Vr[t] + \mathring{R}\mathring{\Lambda}\Xi\bar{V}\tilde{r}[t]) \\ &= \sum_{i=1}^{N} V_i^\top\Xi_i^{-1}\Lambda_i^{-1}R_i^\top \\ &\qquad \times f_i(Z_i z_i[t] + R_i\Lambda_i\Xi_i Lz[t] + R_i\Lambda_i\Xi_i V_i r[t] + R_i\Lambda_i\Xi_i\bar{V}_i\tilde{r}[t]), \\ \tilde{r}[t+1] &= (I+S)^{K-1}S^{-1}\bar{V}^\top\Xi^{-1}\mathring{\Lambda}^{-1}\mathring{R}^\top(W - I_N)_{\otimes n} \\ &\qquad \times F(t, \mathring{Z}z[t] - \mathring{R}\mathring{\Lambda}\Xi Lz[t] + \mathring{R}\mathring{\Lambda}\Xi Vr[t] + \mathring{R}\mathring{\Lambda}\Xi\bar{V}\tilde{r}[t]).\end{aligned} \tag{6.2.8}$$

If the spectral radius of $(I+S)$ is less than 1 (i.e., $\rho(I+S) < 1$), it is clear that $\tilde{r}$ becomes almost zero for sufficiently large K. Indeed, the following results show that the spectral radius of $(I+S)$ is less than 1 for a row-stochastic matrix (Proposition 6.2.2) and a column-stochastic matrix (Corollary 6.2.4).

Proposition 6.2.2. For a row-stochastic matrix W and positive vector $q \in \mathbb{R}^N$ satisfying $W^\top q = q$ and $\mathbb{1}_N^\top q = 1$, let $\Pi^q = \mathrm{diag}(q_i) \in \mathbb{R}^N$ and

$$G_1 := \bar{V}^\top\Xi\mathring{R}^\top\Pi^q_{\otimes n}\mathring{R}\Xi\bar{V}.$$

Then,

$$\Delta G_1 := (I+S)^\top G_1(I+S) - G_1 < 0. \tag{6.2.9}$$

$\Diamond$

It should be noted that (6.2.9) implies that the asymptotic stability of a linear

system $w[t+1] = (I+S)w[t]$ can be shown by the discrete-time Lyapunov function $w^\top Gw$, which is equivalent to $\rho(I+S) < 1$.

Before proving Proposition 6.2.2, we first introduce the following lemma which will be used for the proof.

Lemma 6.2.3. [WLMA19, Lemma 1] Consider a row-stochastic matrix W.

1. $L^W := \Pi^q - W^\top \Pi^q W$ is positive semi-definite and $L^W \mathbb{1}_N = \mathbb{0}_N$.
2. If every $W_{ii} > 0$ for all $i \in \mathcal{N}$, then the kernel of L^W is one-dimensional. $\Diamond$

Now, the proof of Proposition 6.2.2 can be shown as follows.

Proof. From (6.2.6), $I + S$ can be rewritten as

$$\begin{aligned} I + S &= I_{\bar{\mathrm{r}}-\mathrm{r}_0} + \bar{V}^\top \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \\ &= \bar{V}^\top \Xi^{-1} \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi \bar{V}. \end{aligned} \tag{6.2.10}$$

Substituting (6.2.10) into (6.2.9), it is obtained

$$\begin{aligned} \Delta G_1 = &\big(\bar{V}^\top \Xi \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi^{-1} \bar{V} \big) \big(\bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \big) \\ &\times \big(\bar{V}^\top \Xi^{-1} \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi \bar{V} \big) \\ &- \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \\ = &\, \bar{V}^\top \Xi \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi^{-1} (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \\ &\times (I_{\bar{\mathrm{r}}} - VV^\top) \Xi^{-1} \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi \bar{V} - \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \end{aligned}$$

where we use $\bar{V}\bar{V}^\top = I_{\bar{\mathrm{r}}} - VV^\top$. Then, we have

$$\begin{aligned} \Delta G_1 = &\, \bar{V}^\top (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi (I_{\bar{\mathrm{r}}} - VV^\top) \bar{V} \\ &+ \bar{V}^\top (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi (I_{\bar{\mathrm{r}}} - VV^\top) \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \\ &+ \bar{V}^\top \Xi \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi (I_{\bar{\mathrm{r}}} - VV^\top) \bar{V} \\ &+ \bar{V}^\top \Xi \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \\ &\times (I_{\bar{\mathrm{r}}} - VV^\top) \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \\ &- \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \end{aligned}$$

$$
\begin{aligned}
&= \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \\
&\quad + \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi (I_{\bar{\mathrm{r}}} - VV^\top) \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \\
&\quad + \bar{V}^\top \Xi \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V} \\
&\quad + \bar{V}^\top \Xi \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} (I_{\bar{\mathrm{r}}} - VV^\top) \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \\
&\quad \times (I_{\bar{\mathrm{r}}} - VV^\top) \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \\
&\quad - \bar{V}^\top \Xi \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \Xi \bar{V}
\end{aligned}
$$

where we use $\bar{V}^\top V = O$ in the last equality.

The above equality can be simplified by the following properties:

- From Proposition 6.2.1,

$$
\begin{aligned}
(W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} V V^\top &= O, \\
\text{or} \quad V V^\top \Xi^{-1} \mathring{\Lambda} \mathring{R} (W - I_N)_{\otimes n} &= O.
\end{aligned}
$$

- Since both $\Pi^q_{\otimes n}$ and $\mathring{R}\mathring{\Lambda}\mathring{R}^\top$ are block diagonal matrices with corresponding diagonal blocks of the same dimension, they commute. This yields

$$
\begin{aligned}
\mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \mathring{\Lambda} \mathring{R}^\top &= \mathring{R}^\top \mathring{R} \mathring{\Lambda} \mathring{R}^\top \Pi^q_{\otimes n} \\
&= \mathring{\Lambda} \mathring{R}^\top \Pi^q_{\otimes n}
\end{aligned}
$$

  where the second inequality comes from $\mathring{R}^\top \mathring{R} = I$. Similarly,

$$
\begin{aligned}
\mathring{R} \mathring{\Lambda} \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} &= \Pi^q_{\otimes n} \mathring{R} \mathring{\Lambda} \mathring{R}^\top \mathring{R} \\
&= \Pi^q_{\otimes n} \mathring{R} \mathring{\Lambda}.
\end{aligned}
$$

- Since Π^q is positive semi-definite, there exists $\sqrt{\Pi^q} \in \mathbb{R}^{N \times N}$ such that $\Pi^q = \sqrt{\Pi^q}\sqrt{\Pi^q}$. Moreover, $\sqrt{\Pi^q}$ also commutes with $\mathring{R}\mathring{\Lambda}\mathring{R}^\top$. From this, it follows that

$$
\begin{aligned}
\mathring{R} \mathring{\Lambda} \mathring{R}^\top \Pi^q_{\otimes n} \mathring{R} \mathring{\Lambda} \mathring{R}^\top &= (\mathring{R} \mathring{\Lambda} \mathring{R}^\top) \sqrt{\Pi^q}_{\otimes n} \sqrt{\Pi^q}_{\otimes n} (\mathring{R} \mathring{\Lambda} \mathring{R}^\top) \\
&= \sqrt{\Pi^q}_{\otimes n} (\mathring{R} \mathring{\Lambda} \mathring{R}^\top) (\mathring{R} \mathring{\Lambda} \mathring{R}^\top) \sqrt{\Pi^q}_{\otimes n}
\end{aligned}
$$

$$
\begin{aligned}
&= \sqrt{\Pi^q}_{\otimes n}\mathring{R}\mathring{\Lambda}\mathring{\Lambda}\mathring{R}^\top\sqrt{\Pi^q}_{\otimes n} \\
&\leq \sqrt{\Pi^q}_{\otimes n}\mathring{R}\mathring{R}^\top\sqrt{\Pi^q}_{\otimes n} \\
&\leq \sqrt{\Pi^q}_{\otimes n}\sqrt{\Pi^q}_{\otimes n} = \Pi^q_{\otimes n}
\end{aligned}
$$

where the first and the second inequalities come from $\rho(\Lambda_i) \leq 1$ and $R_i R_i^\top$ is an orthogonal projection matrix so that $R_i R_i^\top \leq I_n$ for all $i \in \mathcal{N}$, respectively.

Using above properties, we have

$$
\begin{aligned}
\Delta G_1 &\leq \bar{V}^\top \Xi\{\mathring{\Lambda}\mathring{R}^\top\Pi^q_{\otimes n}(W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda} + \mathring{\Lambda}\mathring{R}^\top(W^\top - I_N)_{\otimes n}\Pi^q_{\otimes n}\mathring{R}\mathring{\Lambda} \\
&\qquad + \mathring{\Lambda}\mathring{R}^\top(W^\top - I_N)_{\otimes n}\Pi^q_{\otimes n}(W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\}\Xi\bar{V} \\
&= -\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top\{\Pi^q_{\otimes n} - W^\top_{\otimes n}\Pi^q_{\otimes n}W_{\otimes n}\}\mathring{R}\mathring{\Lambda}\Xi\bar{V} \\
&= -\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}
\end{aligned}
$$

From this, it is enough to show that $\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}$ is positive definite to complete the proof.

Since $L^W_{\otimes n}$ is positive semi-definite from Lemma 6.2.3, so is $\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}$. To show $\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}$ is positive definite, consider $\zeta = \mathrm{col}(\zeta_1, \ldots, \zeta_N) \in \mathbb{R}^{\bar{\mathrm{r}}-\mathrm{r}_0}$ such that

$$
\zeta^\top\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}\zeta = 0.
$$

It can be easily shown that, if $\zeta^\top\bar{V}^\top\Xi\mathring{\Lambda}\mathring{R}^\top L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}\zeta = 0$, then $L^W_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}\zeta = \mathbb{0}$. Since the kernel of L^W is same with the kernel of $(W - I_N)$ as $\mathrm{im}(\mathbb{1}_N)$ by Lemma 6.2.3, it follows that

$$
(W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\bar{V}\zeta = \mathbb{0}_{nN}.
$$

Note that the columns of V spans the kernel of $(W - I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi$. Thus, we have $\bar{V}\zeta \in \mathrm{im}(V)$ which implies $\zeta = \mathbb{0}_{\bar{\mathrm{r}}-\mathrm{r}_0}$. □

If W is a column-stochastic matrix, then it can be easily shown that $\rho(I+S) < 1$ by the following corollary.

Corollary 6.2.4. For a column-stochastic matrix W and positive vector $p \in \mathbb{R}^N$ satisfying $Wp = p$ and $\mathbb{1}_N^\top p = 1$, let $\Pi^p = \mathrm{diag}(p_i) \in \mathbb{R}^N$ and

$$G_2 := \bar{V}^\top \Xi^{-1} \mathring{R}^\top \Pi^p_{\otimes n} \mathring{R} \Xi^{-1} \bar{V}.$$

Then,

$$\Delta G_2 := (I + S^\top)^\top G_2 (I + S^\top) - G_2 < 0. \tag{6.2.11}$$

$\lozenge$

Similarly with Proposition 6.2.2, Corollary 6.2.4 guarantees the asymptotic stability of $w[t+1] = (I + S^\top) w[t]$ using the Lyapunov function $w^\top G_2 w$, i.e., $\rho(I + S^\top) < 1$. Since $\rho(I + S^\top) = \rho((I + S)^\top) = \rho(I + S)$, Corollary 6.2.4 shows $\rho(I + S) < 1$ for any column-stochastic matrix W.

The key idea of the proof for Corollary 6.2.4 is to replace the column-stochastic matrix W to the row-stochastic matrix $W^\top$ in the proof of Proposition 6.2.2. The detail of the proof is as follows.

Proof. Since $I + S^\top$ can be written as

$$\begin{aligned} I + S^\top &= I_{\bar{\mathrm{r}} - \mathrm{r}_0} + \bar{V}^\top \Xi \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} \bar{V} \\ &= \bar{V}^\top \Xi \{ I_{\bar{\mathrm{r}}} + \mathring{\Lambda} \mathring{R}^\top (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi^{-1} \bar{V}, \end{aligned}$$

with the similar procedure in the proof of Proposition 6.2.2, we have

$$\begin{aligned} \Delta G_2 \leq \bar{V}^\top \Xi^{-1} \{ & \mathring{\Lambda} \mathring{R}^\top \Pi^p_{\otimes n} (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \\ & + \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \Pi^q_{\otimes n} \mathring{R} \mathring{\Lambda} \\ & + \mathring{\Lambda} \mathring{R}^\top (W - I_N)_{\otimes n} \Pi^q_{\otimes n} (W^\top - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \} \Xi^{-1} \bar{V} \\ = -\bar{V}^\top \Xi^{-1} & \mathring{\Lambda} \mathring{R}^\top \{ \Pi^p_{\otimes n} - W_{\otimes n} \Pi^p_{\otimes n} W^\top_{\otimes n} \} \mathring{R} \mathring{\Lambda} \Xi^{-1} \bar{V} \\ = -\bar{V}^\top \Xi^{-1} & \mathring{\Lambda} \mathring{R}^\top L^{W^\top}_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} \bar{V} \end{aligned}$$

where $L^{W^\top}_{\otimes n} = \Pi^p - W \Pi^p W^\top$.

By applying Lemma 6.2.3 to $W^\top$ which is row-stochastic, $L^{W^\top}$ is a positive semi-definite matrix. So $\bar{V}^\top \Xi^{-1} \mathring{\Lambda} \mathring{R}^\top L^{W^\top}_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi^{-1} \bar{V}$ is also positive semi-definite.

To show $\bar{V}^\top \Xi^1 \mathring{\Lambda} \mathring{R}^\top L_{\otimes n}^{W^\top} \mathring{R} \mathring{\Lambda} \Xi^1 \bar{V}$ is positive definite, let $\zeta = \mathrm{col}(\zeta_1, \ldots, \zeta_N) \in \mathbb{R}^{\bar{\mathrm{r}}-\mathrm{r}_0}$ such that

$$\zeta^\top \bar{V}^\top \Xi^1 \mathring{\Lambda} \mathring{R}^\top L_{\otimes n}^{W^\top} \mathring{R} \mathring{\Lambda} \Xi^1 \bar{V} \zeta = 0.$$

It follows that $L_{\otimes n}^{W^\top} \mathring{R} \mathring{\Lambda} \Xi^{-1} \bar{V} \zeta = 0$. By Lemma 6.2.3, the kernel of $L^{W^\top}$ is $\mathrm{im}(\mathbb{1}_N)$ because every diagonal entry of $W^\top$ is positive. This yields

$$(W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi \bar{V} \zeta = 0.$$

Since the columns of V are basis of the kernel of $(W - I_N)_{\otimes n} \mathring{R} \mathring{\Lambda} \Xi$, it follows that $\bar{V}\zeta \in \mathrm{im}(V)$. From the orthogonality of V and $\bar{V}$, we have $\zeta = 0$ and this implies that $\bar{V}^\top \Xi^1 \mathring{\Lambda} \mathring{R}^\top L_{\otimes n}^{W^\top} \mathring{R} \mathring{\Lambda} \Xi^1 \bar{V}$ is positive definite. □

6.3 Prediction on Emergent Behavior under Rank-deficient Coupling

In the previous section, it is proved that the spectral radius of $I + S$ for any stochastic matrix W is less than 1. By recalling the transformed system (6.2.8), the $\tilde{r}$ dynamics vanishes as K gets larger. Based on this observation, we introduce the blended dynamics for (6.1.1) as

$$\begin{aligned}
\hat{z}_i[t+1] &= Z_i^\top f_i(t, Z_i \hat{z}_i[t] - R_i \Lambda_i \Xi_i L_i \hat{z}[t] + R_i \Lambda_i \Xi_i V_i \hat{r}[t]) \in \mathbb{R}^{nN-\bar{\mathrm{r}}}, \\
\hat{r}[t+1] &= \sum_{i=1}^{N} V_i^\top \Xi_i^{-1} \Lambda_i^{-1} R_i^\top \\
&\qquad \times f_i(t, Z_i \hat{z}_i[t] - R_i \Lambda_i \Xi_i L_i \hat{z}[t] + R_i \Lambda_i \Xi_i V_i \hat{r}[t]) \in \mathbb{R}^{\mathrm{r}_0},
\end{aligned} \tag{6.3.1}$$

where $\hat{z} \in \mathbb{R}^{nN-\bar{\mathrm{r}}}$ and $\hat{r} \in \mathbb{R}^{\mathrm{r}_0}$, for all $i \in \mathcal{N}$.

Now, we can state the main result of this chapter where the system state $x_i[t]$ at integer count t can be approximated by the solution of the blended dynamics (6.3.1) for sufficiently large number of steps for the rank-deficient coupling in (6.1.1).

Theorem 6.3.1. Under assumptions 6.1.1, 6.2.1, and 6.2.2, assume the blended

dynamics (6.3.1) is contractive[1], then, for any $\epsilon > 0$, there exists $K^{\min}$ such that, for all $K > K^{\min}$, the solution x_i of (6.1.1) and the solution of (6.3.1) with arbitrary initial conditions satisfy

$$\limsup_{t\to\infty} \left\| x_i[t] - (Z_i\hat{z}_i[t] - R_i\Lambda_i\Xi_i L_i\hat{z}[t] + R_i\Lambda_i\Xi_i V_i\hat{r}[t]) \right\| < \epsilon, \quad \forall i \in \mathcal{N}. \tag{6.3.2}$$

♢

Proof. For convenience, let $\xi := \mathrm{col}(z_1, \ldots, z_N, r)$ and $\tilde{\xi} := \tilde{r}$. Then, the system (6.2.8) can be written as

$$\begin{aligned} \xi[t+1] &= h(t, \xi[t], \tilde{\xi}[t]) \\ \tilde{\xi}[t+1] &= (I+S)^{K-1} g(t, \xi[t], \tilde{\xi}[t]) \end{aligned} \tag{6.3.3}$$

where $\xi \in \mathbb{R}^m$, $\tilde{\xi} \in \mathbb{R}^{nN-m}$, and h and g are defined by (6.2.8), which are continuously differentiable with ξ and $\tilde{\xi}$, globally Lipschitz with respect to ξ and $\tilde{\xi}$, and uniformly bounded in t. In other words, there exist a non-decreasing continuous function $M_h : \mathbb{R} \to \mathbb{R}$, $M_g : \mathbb{R} \to \mathbb{R}$ and a constant $L_{h,\xi}, L_{h,\tilde{\xi}}, L_{g,\xi}, L_{g,\tilde{\xi}} \geq 0$ such that, $\forall \xi_1, \xi_2 \in \mathbb{R}^m$, $\tilde{\xi}_1, \tilde{\xi}_2 \in \mathbb{R}^{nN-m}$, $t \in \mathbb{Z}$, and $i \in \mathcal{N}$,

$$\begin{aligned} \|h(t,\xi,\tilde{\xi})\| &\leq M_h(\|\mathrm{col}(\xi,\tilde{\xi})\|), \\ \|g(t,\xi,\tilde{\xi})\| &\leq M_g(\|\mathrm{col}(\xi,\tilde{\xi})\|), \\ \|h(t,\xi_1,\tilde{\xi}) - h(t,\xi_2,\tilde{\xi})\| &\leq L_{h,\xi}\|\xi_1 - \xi_2\|, \\ \|h(t,\xi,\tilde{\xi}_1) - h(t,\xi,\tilde{\xi}_2)\| &\leq L_{h,\tilde{\xi}}\|\tilde{\xi}_1 - \tilde{\xi}_2\|, \\ \|g(t,\xi_1,\tilde{\xi}) - g(t,\xi_2,\tilde{\xi})\| &\leq L_{g,\xi}\|\xi_1 - \xi_2\|, \\ \|g(t,\xi,\tilde{\xi}_1) - g(t,\xi,\tilde{\xi}_2)\| &\leq L_{g,\tilde{\xi}}\|\tilde{\xi}_1 - \tilde{\xi}_2\|. \end{aligned} \tag{6.3.4}$$

[1]The system $x[t+1] = f(t, x[t]) \in \mathbb{R}^n$, where f is continuously differentiable, is said to be contractive [LS98, TRK18] if there exist a positive definite matrix $H \in \mathbb{R}^{n\times n}$ and a positive constant $\gamma < 1$ such that

$$\frac{\partial f}{\partial x}(t,x)^\top H^2 \frac{\partial f}{\partial x}(t,x) \leq \gamma H^2, \quad \forall x \in \mathbb{R}^n, t \in \mathbb{Z}.$$

Meanwhile, the blended dynamics 6.3.1 can be written as

$$\hat{\xi}[t+1] = h(t, \hat{\xi}[t], \mathbb{0}_{nN-m}) \tag{6.3.5}$$

where $\hat{\xi} \in \mathbb{R}^m$. Since the blended dynamics (6.3.1) is contractive, there exist a (symmetric) positive definite matrix $H \in \mathbb{R}^{n\times n}$ and a positive constant $\gamma < 1$ such that

$$\frac{\partial h}{\partial s}(t, \hat{\xi}, \mathbb{0}_{nN-m})^\top H^2 \frac{\partial h}{\partial s}(t, \hat{\xi}, \mathbb{0}_{nN-m}) \leq \gamma H^2, \quad \forall \hat{\xi} \in \mathbb{R}^m, t \in \mathbb{Z}.$$

Note that, since the blended dynamics (6.3.1) is contractive, by Lemma 4.2.3, the following inequality holds

$$\|H\{h(t, \hat{\xi}_2, \mathbb{0}_{nN-m}) - h(t, \hat{\xi}_2, \mathbb{0}_{nN-m})\}\| \leq \sqrt{\gamma}\|H(\hat{\xi}_2 - \hat{\xi}_1)\| \tag{6.3.6}$$

for all $t \in \mathbb{Z}$ and $\hat{\xi}_1, \hat{\xi}_2 \in \mathbb{R}^m$.

Before proving Theorem 6.3.1, we claim the solution of the blended dynamics (6.3.1) is bounded. In particular, the boundedness of the solution of the blended dynamics can be shown by

$$\|H\hat{\xi}[t]\| \leq w_\xi[t], \quad \forall t \geq 0 \tag{6.3.7}$$

where $w_\xi \in \mathbb{R}$ is the solution of

$$w_\xi[t+1] = \sqrt{\gamma} w_\xi[t] + \|Hh(t, \mathbb{0}_m, \mathbb{0}_{nN-m})\|, \quad w_\xi[0] = \|H\hat{\xi}[0]\|. \tag{6.3.8}$$

It is clear that $\|H\hat{\xi}[\tau]\| \leq w_\xi[\tau]$ for $\tau = 0$. Let us suppose $\|H\hat{\xi}[\tau]\| \leq w_\xi[\tau]$ for some integer $\tau \geq 0$, then

$$\begin{aligned}
\|H\hat{\xi}[\tau+1]\| &= \|H\{h(\tau, \hat{\xi}[\tau], \mathbb{0}_{nN-m}) - h(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m}) + h(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\}\| \\
&\leq \|H\{h(\tau, \hat{\xi}[\tau], \mathbb{0}_{nN-m}) - h(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\}\| \\
&\quad + \|Hh(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\| \\
&\leq \sqrt{\gamma}\|H\hat{\xi}[\tau]\| + \|Hh(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\| \\
&\leq \sqrt{\gamma} w_\xi[\tau] + \|Hh(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\|
\end{aligned}$$

$$= w_\xi[\tau+1],$$

where the second inequality comes from (6.3.6). By mathematical induction, this proves (6.3.7).

Meanwhile, the solution w_ξ of (6.3.8) is obtained as

$$w_\xi[t] = \sqrt{\gamma}^t w_\xi[0] + \sum_{\tau=0}^{t-1} \sqrt{\gamma}^{t-\tau-1} \|Hh(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\|.$$

From the uniformly bounded assumption of h in (6.3.4), i.e., $\|h(t,\xi,\tilde{\xi})\| \le M_h(\|\mathrm{col}(\xi,\tilde{\xi})\|)$, it follows that

$$\begin{aligned}
\limsup_{t\to\infty} \|\hat{\xi}[t]\| &\le \|H^{-1}\| \limsup_{t\to\infty} \|H\hat{\xi}[t]\| \\
&\le \|H^{-1}\| \limsup_{t\to\infty} \left\{ \sqrt{\gamma}^t \|H\hat{\xi}[0]\| + \frac{\sup_{\tau\ge 0} \|Hh(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\|}{1-\sqrt{\gamma}} \right\} \\
&\le \|H^{-1}\|\|H\| \frac{\sup_{\tau\ge 0} \|h(\tau, \mathbb{0}_m, \mathbb{0}_{nN-m})\|}{1-\sqrt{\gamma}} \\
&\le \|H^{-1}\|\|H\| \frac{M_h(0)}{1-\sqrt{\gamma}} =: M_{\hat{h}}
\end{aligned} \tag{6.3.9}$$

which justifies the claim.

Now, the behavior of the network under the rank-deficient coupling (6.3.3), or equivalently (6.2.8), is studied with the blended dynamics (6.3.5), or equivalently (6.3.1), which describes the evolution of the overall system at every integer time t. For this, we introduce a Lyapunov function

$$V = \big\|H(\xi - \hat{\xi})\big\| + \eta\big\|\tilde{\xi}\big\|$$

where $\eta > L_{h,\tilde{\xi}} \|H\| / \sqrt{\gamma}$. Then, we have

$$\begin{aligned}
V[t+1] &= \big\|H(\xi[t+1] - \hat{\xi}[t+1])\big\| + \eta\big\|\tilde{\xi}[t+1]\big\| \\
&= \big\|H\{h(t,\xi[t],\mathbb{0}_{nN-m}) - h(t,\hat{\xi}[t],\mathbb{0}_{nN-m}) \\
&\qquad + h(t,\xi[t],\tilde{\xi}[t]) - h(t,\xi[t],\mathbb{0}_{nN-m})\}\big\|
\end{aligned}$$

$$
\begin{aligned}
&+\eta\big\|(I+S)^{K-1}\{g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\\
&\quad+g(t,\xi[t],\mathbb{0}_{nN-m})-g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\\
&\quad+g(t,\xi[t],\tilde{\xi}[t])-g(t,\xi[t],\mathbb{0}_{nN-m})\}\big\|\\
\le\ &\big\|H\{h(t,\xi[t],\mathbb{0}_{nN-m})-h(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\}\big\|\\
&+\big\|H\{h(t,\xi[t],\tilde{\xi}[t])-h(t,\xi[t],\mathbb{0}_{nN-m})\}\big\|\\
&+\rho(I+S)^{K-1}\eta\Big\{\|g(t,\xi[t],\mathbb{0}_{nN-m})-g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\|\\
&\quad+\|g(t,\xi[t],\tilde{\xi}[t])-g(t,\xi[t],\mathbb{0}_{nN-m})\|\\
&\quad+\|g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\|\Big\}.
\end{aligned}
$$

The above inequality can be simplified by the following properties:

- by (6.3.6),

$$\big\|H\{h(t,\xi[t],\mathbb{0}_{nN-m})-h(t,\tilde{\xi}[t],\mathbb{0}_{nN-m})\}\big\|\le\sqrt{\gamma}\big\|H(\xi[t]-\hat{\xi}[t])\big\|$$

- by (6.3.4),

$$
\begin{aligned}
\big\|h(t,\xi[t],\tilde{\xi}[t])-h(t,\xi[t],\mathbb{0}_{nN-m})\big\|&\le L_{h,\tilde{\xi}}\|\tilde{\xi}[t]\|\\
\big\|g(t,\xi[t],\tilde{\xi}[t])-g(t,\xi[t],\mathbb{0}_{nN-m})\big\|&\le L_{g,\tilde{\xi}}\|\tilde{\xi}[t]\|,
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\big\|g(t,\xi[t],\mathbb{0}_{nN-m})-g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\big\|&\le L_{g,\xi}\big\|\xi[t]-\hat{\xi}[t]\big\|\\
&\le L_{g,\xi}\big\|H^{-1}\big\|\big\|H\{\xi[t]-\hat{\xi}[t]\}\big\|
\end{aligned}
$$

Using the above inequalities, it is obtained that

$$\begin{aligned} V[t+1] &\leq \sqrt{\gamma}\big\|H(\xi[t]-\hat{\xi}[t])\big\| + L_{h,\tilde{\xi}}\big\|H\big\|\|\tilde{\xi}[t]\| \\ &\quad + \rho(I+S)^{K-1}\eta L_{g,\xi}\|H^{-1}\|\|H(\xi[t]-\hat{\xi}[t])\| \\ &\quad + \rho(I+S)^{K-1}\eta L_{g,\tilde{\xi}}\|\tilde{\xi}[t]\| \\ &\quad + \rho(I+S)^{K-1}\eta\|g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\| \\ &\leq \sqrt{\gamma}V[t] + \rho(I+S)^{K-1}\eta M_1 V[t] \\ &\quad + \rho(I+S)^{K-1}\eta\|g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\|, \end{aligned} \tag{6.3.10}$$

where $M_1 := \max\{L_{g,\xi}\|H^{-1}\|, L_{g,\tilde{\xi}}/\eta\}$.

For the given ϵ, let $K^{\min}$ be a positive integer such that

$$\rho(I+S)^{K^{\min}}\eta M_1 \leq \frac{1-\sqrt{\gamma}}{2} \tag{6.3.11}$$

$$\rho(I+S)^{K^{\min}}\big\|\mathcal{T}^{-1}\big\|\max\left\{\big\|H^{-1}\big\|, \frac{1}{\eta}\right\}\frac{2\eta}{1-\sqrt{\gamma}}M_g(M_{\hat{h}}) \leq \epsilon \tag{6.3.12}$$

Then the following inequality is obtained by substituting (6.3.11) into (6.3.10),

$$\begin{aligned} V[t+1]-V[t] &\leq -\frac{(1-\sqrt{\gamma})}{2}V[t] \\ &\quad + \rho(I+S)^{K-1}\eta\|g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\| \end{aligned} \tag{6.3.13}$$

for all $K > K^{\min}$. By (6.3.4) and (6.3.9),

$$\begin{aligned} \limsup_{t\to\infty}\|g(t,\hat{\xi}[t],\mathbb{0}_{nN-m})\| &\leq M_g(\limsup_{t\to\infty}\|\hat{\xi}[t]\|) \\ &\leq M_g\left(M_{\hat{h}}\right) \end{aligned}$$

Using this and (6.3.13), the ultimate bound of V is obtained as

$$\limsup_{t\to\infty} V[t] \leq \rho(I+S)^{K-1}\frac{2\eta}{1-\sqrt{\gamma}}M_g(M_{\hat{h}}). \tag{6.3.14}$$

Therefore, for each agent $i \in \mathcal{N}$ and $K > K^{\min}$,

$$\begin{aligned}
&\limsup_{t\to\infty} \left\| x_i[t] - (Z_i\hat{z}_i[t] - R_i\Lambda_i\Xi_i L_i\hat{z}[t] + R_i\Lambda_i\Xi_i V_i\hat{r}[t]) \right\| \\
&= \limsup_{t\to\infty} \left\| \mathcal{T}^{-1}\begin{bmatrix}\xi[t]\\ \tilde{\xi}[t]\end{bmatrix} - \mathcal{T}^{-1}\begin{bmatrix}\hat{\xi}[t]\\ \mathbb{0}_{nN-m}\end{bmatrix} \right\| \\
&\le \left\|\mathcal{T}^{-1}\right\| \limsup_{t\to\infty} \left\| \begin{bmatrix}\xi[t]\\ \tilde{\xi}[t]\end{bmatrix} - \begin{bmatrix}\hat{\xi}[t]\\ \mathbb{0}_{nN-m}\end{bmatrix} \right\| \\
&\le \left\|\mathcal{T}^{-1}\right\| \limsup_{t\to\infty} \{ \left\|H^{-1}\right\| \left\|H(\xi[t] - \hat{\xi}[t])\right\| + \left\|\tilde{\xi}[t]\right\| \} \\
&\le \left\|\mathcal{T}^{-1}\right\| \max\left\{ \left\|H^{-1}\right\|, \frac{1}{\eta} \right\} V[t] \\
&\le \rho(I+S)^{K-1} \left\|\mathcal{T}^{-1}\right\| \max\left\{ \left\|H^{-1}\right\|, \frac{1}{\eta} \right\} \frac{2\eta}{1-\sqrt{\gamma}} M_g(M_{\hat{h}}) \\
&\le \epsilon,
\end{aligned}$$

where the last inequality comes from (6.3.12). This completes the proof for (6.3.2) of Theorem 6.3.1. □

6.3.1 Approximation by Blended Dynamics for Simplified Cases

6.3.1.1 Case 1: Every B_i is non-identical and positive definite

If every B_i is positive definite (but not necessarily to be identical), then $\mathrm{r}_i = n$ so that $B_i = R_i\Lambda_i^2 R_i^\top$, and Z_i and L become null for all $i \in \mathcal{N}$. In addition, $V \in \mathbb{R}^{nN\times n}$ because $\mathrm{r}_0 = n$. Without loss of generality, we can define $R_i = I_n$ and $\Lambda_i = \sqrt{B_i}$ and it follows

$$\begin{aligned}
\ker\big((W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\big) &= \ker\big((W-I_N)_{\otimes n}\mathrm{diag}(\sqrt{B_i})\Xi\big) \\
&= \{v \in \mathbb{R}^{nN} : \mathrm{diag}(\sqrt{B_i})\Xi v \in \mathrm{im}(p\otimes I_n)\} \\
&= \left\{ \mathrm{col}(v_1,\cdots,v_N) \in \mathbb{R}^{nN} : \frac{\sqrt{B_1}\Xi_1 v_1}{p_1} = \cdots = \frac{\sqrt{B_N}\Xi_N v_N}{p_N} \right\} \\
&= \left\{ \mathrm{col}(v_1,\cdots,v_N) \in \mathbb{R}^{nN} : \frac{\sqrt{B_1} v_1}{\sqrt{p_1 q_1}} = \cdots = \frac{\sqrt{B_N} v_N}{\sqrt{p_N q_N}} \right\}
\end{aligned}$$

where the last equality comes from $\Xi_i = \sqrt{p_i/q_i} I_n$. This implies that there exists a square matrix $V_0 \in \mathbb{R}^{n\times n}$ such that, for $V = \mathrm{col}(V_1, \cdots, V_N) \in \mathbb{R}^{nN}$, $V_i = \sqrt{p_i q_i}\sqrt{B_i}^{-1} V_0$ for all $i \in \mathcal{N}$. Since every column of V is orthonormal, we have

$$\begin{aligned} V^\top V &= \sum_{i=1}^N V_i^\top V_i \\ &= \sum_{i=1}^N p_i q_i {V_0}^\top \sqrt{B_i}^{-1}\sqrt{B_i}^{-1} V_0 \\ &= {V_0}^\top \Big(\sum_{i=1}^N {B_i}^{-1}\Big) V_0 \\ &= I_n. \end{aligned}$$

By letting $V_0 = \sqrt{(\sum_{i=1}^N B_i^{-1})^{-1}}$, we define a new variable $\hat{s} = V_0 \hat{r}$ so the blended dynamics (6.3.1) is transformed into

$$\begin{aligned} \hat{s}[t+1] &= V_0 \sum_{i=1}^N \sqrt{p_i q_i}\sqrt{\frac{q_i}{p_i}} {V_0}^\top \sqrt{B_i}^{-1}\sqrt{B_i}^{-1} \\ &\qquad \times f_i\Big(t, \sqrt{\frac{p_i}{q_i}}\sqrt{p_i q_i}\sqrt{B_i}\sqrt{B_i}^{-1}\hat{s}[t]\Big) \\ &= \Big(\sum_{i=1}^N B_i^{-1}\Big)^{-1} \sum_{i=1}^N q_i B_i^{-1} f_i(t, p_i \hat{s}[t]). \end{aligned}$$

With sufficiently large K, we can approximate $x_i[t]$ from (6.2.7) by

$$\begin{aligned} R_i \Lambda_i \Xi_i V_i \hat{r} &= \sqrt{\frac{p_i}{q_i}}\sqrt{p_i q_i}\sqrt{B_i}\sqrt{B_i}^{-1} V_0 \hat{r}[t] \\ &= p_i \hat{s}[t]. \end{aligned}$$

6.3.1.2 Case 2: Every B_i is identical and positive semi-definite

If every B_i is identical to some positive semi-definite matrix B_0 for all $i \in \mathcal{N}$, then $R_1 = \ldots = R_N =: R_0$, $Z_1 = \ldots = Z_N =: Z_0$, and $\Lambda_1 = \ldots = \Lambda_N =: \Lambda_0$. This yields $\mathring{B} = I_N \otimes B_0$, $\mathring{R} = I_N \otimes R_0$, $\mathring{B} = I_N \otimes Z_0$, and $\mathring{\Lambda} = I_N \otimes \Lambda_0$. From

these,

$$\begin{aligned} L &= \bar{V}S^{-1}\bar{V}^\top \Xi \mathring{\Lambda}\mathring{R}^\top (W-I_N)_{\otimes n}\mathring{Z} \\ &= \bar{V}S^{-1}\bar{V}^\top \Xi (I_N\otimes\Lambda_0)(I_N\otimes R_0)^\top (W-I_N)_{\otimes n}(I_N\otimes Z_0) \\ &= \bar{V}S^{-1}\bar{V}^\top \Xi \{(W-I_N)\otimes R_0^\top Z_0\} = O_{\bar{\mathrm{r}}\times(nN-\bar{\mathrm{r}})} \end{aligned}$$

where the last equality comes from $R_0^\top Z_0 = O_{\bar{\mathrm{r}}\times(nN-\bar{\mathrm{r}})}$.

In addition, it is obtained that

$$\begin{aligned} (W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi V &= (W-I_N)_{\otimes n}(I_N\otimes B_0)(I_N\otimes\Lambda_0)\Xi V \\ &= \{(W-I_N)\otimes(R_0\Lambda_0)\}\Xi V. \end{aligned}$$

Then, we can define $V\in\mathbb{R}^{N\mathrm{r}_0\times\mathrm{r}_0}$ such that its columns are orthonormal basis of

$$\begin{aligned} &\ker\big((W-I_N)_{\otimes n}\mathring{R}\mathring{\Lambda}\Xi\big) = \ker\big((W-I_N)\otimes(R_0\Lambda_0)\Xi\big) \\ &= \{v\in\mathbb{R}^{\bar{\mathrm{r}}} : (I_N\otimes(R_0\Lambda_0))\Xi v \in \mathrm{im}(p\otimes I_n)\} \\ &= \left\{\mathrm{col}(v_1,\cdots,v_N)\in\mathbb{R}^{\bar{\mathrm{r}}} : \frac{R_0\Lambda_0\Xi_1 v_1}{p_1} = \cdots = \frac{R_0\Lambda_0\Xi_N v_N}{p_N}\right\} \\ &= \left\{\mathrm{col}(v_1,\cdots,v_N)\in\mathbb{R}^{\bar{\mathrm{r}}} : \frac{R_0\Lambda_0 v_1}{\sqrt{p_1q_1}} = \cdots = \frac{R_0\Lambda_0 v_N}{\sqrt{p_Nq_N}}\right\}. \end{aligned}$$

This implies that V can be choosen as $V=\mathrm{col}(\sqrt{p_1q_1},\cdots,\sqrt{p_Nq_N})\otimes V_0\in\mathbb{R}^{N\mathrm{r}_0\times\mathrm{r}_0}$ for any square matrix $V_0\in\mathbb{R}^{\mathrm{r}_0\times\mathrm{r}_0}$ such that the columns of V are orthonormal. This yields $V^\top V = \sum_{i=1}^N p_iq_i{V_0}^\top V_0 = {V_0}^\top V_0 = I_{\bar{\mathrm{r}}}$, which also implies $V_0{V_0}^\top = I_{\bar{\mathrm{r}}}$.

Based on the aforementioned observations, with a coordinate change $\hat{s} := \Lambda_0V_0\hat{r}$, the blended dynamics (6.3.1) can be rewritten as

$$\begin{aligned} \hat{z}_i[t+1] &= Z_0^\top f_i(t, Z_0\hat{z}_i[t] + R_0\hat{s}[t]) \in \mathbb{R}^{n-\mathrm{r}_0}, \\ \hat{s}[t+1] &= \sum_{i=1}^N q_iR_0^\top f_i\left(t, Z_0\hat{z}_i[t] + p_iR_0\hat{s}[t]\right) \in \mathbb{R}^{\mathrm{r}_0}, \end{aligned} \tag{6.3.15}$$

because both Ξ_i and Λ_0 is block diagonal matrices so that they commute and

$$V_0{V_i}^\top\Xi_i^{-1} = \sqrt{p_iq_i}V_0{V_0}^\top\Xi_i^{-1}$$

$$
\begin{aligned}
&= \sqrt{p_i q_i} \times \sqrt{\frac{q_i}{p_i}} V_0 {V_0}^\top \\
&= q_i I_{\mathrm{r}_0}, \\
R_0 \Lambda_0 \Xi_i V_i \hat{r} &= \sqrt{\frac{q_i}{p_i}} R_0 \Lambda_0 \times \sqrt{p_i q_i} V_0 \hat{r} \\
&= p_i R_0 \hat{s}.
\end{aligned}
$$

As a result, the system state $x_i[t]$ at each integer count t can be approximated by

$$Z_0 \hat{z}_i[t] + p_i R_0 \hat{s}[t].$$

6.3.1.3 Case 3: Every B_i is identical and positive definite

If every B_i is identical to some positive definite matrix B_0, then Z_0 becomes null and $R_0 = I_n$ so that the blended dynamics (6.3.15) can be rewritten by

$$\hat{s}[t+1] = \sum_{i=1}^{N} q_i f_i \left(t, p_i \hat{s}[t]\right) \in \mathbb{R}^n. \tag{6.3.16}$$

From this, the behavior of each agent, $x_i[t]$, at each integer count t can be estimated by $p_i \hat{s}[t]$. This is exactly corresponds to the result of [KLLS22], which is introduced in Chapter 4.

Chapter 7

Application to Distributed State Estimation

7.1 Problem Formulation

In this chapter, a distributed state estimation algorithm is proposed based on the blended dynamics approach for the rank-deficient coupling, which is introduced in Chapter 6. The proposed algorithm allows every agent to estimate the full state information of the target plant using only its partial output information. Comparing to the classic state estimation problem, the proposed algorithm overcomes the limitations through the inter-agent communication.

We consider a discrete-time linear time-invariant system

$$\begin{aligned} \chi[t+1] &= A\chi[t] \in \mathbb{R}^n, \\ y[t] &= C\chi[t] \in \mathbb{R}^m, \end{aligned} \tag{7.1.1}$$

where χ is the state and y is the output. Here, we assume that (C, A) pair is detectable so that there exists an asymptotic observer [TSH12]. It is supposed that the output y of the system (7.1.1) is monitored by a network of N agents (or sensors) such that an arbitrary agent i can measure the partial output information $y_i \in \mathbb{R}^{m_i}$ which is given by

$$y_i = C_i\chi \in \mathbb{R}^{m_i}, \quad i \in \mathcal{N}, \tag{7.1.2}$$

where $C = \mathrm{col}(C_1, \cdots, C_N) \in \mathbb{R}^m$, with $\sum_{i=1}^N m_i = m$. It should be emphasized

that each (C_i, A) is not necessarily detectable, while (C, A) pair is detectable. In order to utilize the entire output information in the network, we assume the connectivity of the communication network as the following way.

Assumption 7.1.1. The communication network $\mathcal{G}$ is strongly connected. $\Diamond$

Our goal is to design a distributed observer for the system (7.1.1) and (7.1.2) with the communication network $\mathcal{G}$ composed of N agents, such that the state estimation $\hat{\chi}_i[t]$ of each i-th local observer converges to the plant state $\chi[t]$, i.e.,

$$\lim_{t\to\infty} \left\| \hat{\chi}_i[t] - \chi[t] \right\| = 0, \quad i \in \mathcal{N}. \tag{7.1.3}$$

In particular, the proposed local observer in the every agent is designed to satisfy the following properties:

- distributed operation: each local observer communicates with only its neighbors
- local measurement: each local observer utilizes only its local measurement y_i

This chapter is organized as follows. Section 7.2 is composed of the following two subsections. Subsection 7.2.1 introduces the detectability decomposition which is an extension of the Kalman decomposition. Based on this decomposition, a distributed state observer is proposed and its asymptotic performance is shown in Subsection 7.2.2. To verify the performance of the proposed observer, a toy example and its simulation results are illustrated in Section 7.3.

7.2 Distributed Observer for State Estimation

7.2.1 Detectability Decomposition

Recall that each (C_i, A) pair is not necessarily detectable, while (C, A) pair is detectable. Let us denote $\mathcal{U}_i$ as the undetectable subspace of (C_i, A) pair. From this, the following lemma introduces the detectability decomposition, which is an extension of the Kalman decomposition.

Lemma 7.2.1. [KLS19] Denote r_i be the dimension of $\mathcal{U}_i$, the undetectable subspace of (C_i, A) and let $U_i \in \mathbb{R}^{n\times \mathrm{r}_i}$ and $D_i \in \mathbb{R}^{n\times(n-\mathrm{r}_i)}$ such that the columns of U_i are orthonormal vectors which are bases of $\mathcal{U}_i$ and $\mathcal{T}_i := \begin{bmatrix} D_i & U_i \end{bmatrix} \in \mathbb{R}^{n\times n}$ be an orthonormal matrix. Then, $\mathcal{T}_i$ satisfies that

$$\mathcal{T}_i^\top A \mathcal{T}_i = \begin{bmatrix} A_{i\mathrm{d}} & O \\ A_{i\mathrm{r}} & A_{i\mathrm{u}} \end{bmatrix}, \quad C_i\mathcal{T}_i = \begin{bmatrix} C_{i\mathrm{d}} & O \end{bmatrix}, \tag{7.2.1}$$

where $A_{i\mathrm{d}} \in \mathbb{R}^{(n-\mathrm{r_i})\times(\mathrm{n}-\mathrm{r_i})}$, $A_{i\mathrm{r}} \in \mathbb{R}^{\mathrm{r}_i\times(n-\mathrm{r}_i)}$, $A_{i\mathrm{u}} \in \mathbb{R}^{\mathrm{r_i}\times\mathrm{r_i}}$, and $C_{i\mathrm{d}} \in \mathbb{R}^{\mathrm{r_i}\times(\mathrm{n}-\mathrm{r_i})}$. Moreover, the pair $(C_{i\mathrm{d}}, A_{i\mathrm{d}})$ is detectable and the matrix $A_{i\mathrm{u}}$ is unstable, i.e., $\rho(A_{i\mathrm{u}}) > 1$. ◊

It can be shown that there exists $H_{i\mathrm{d}} \in \mathbb{R}^{(n-\mathrm{r}_i)\times\mathrm{r}_i}$ such that $A_{i\mathrm{d}} - H_{i\mathrm{d}}C_{i\mathrm{d}}$ is stable in the discrete-time sense, i.e., $\rho(A_{i\mathrm{d}} - H_{i\mathrm{d}}C_{i\mathrm{d}}) < 1$ because $(C_{i\mathrm{d}}, A_{i\mathrm{d}})$ is detectable.

In addition, the following lemma states that the undetectable subspace of $(\mathrm{col}(C_1, \cdots, C_N), A)$ pair can be described by $\mathcal{U}_i$, the undetectable subspace of (C_i, A), for $i \in \mathcal{N}$.

Lemma 7.2.2. [KLS19] The undetectable subspace of $(\mathrm{col}(C_1, \cdots, C_N), A)$ is $\cap_{i=1}^N \mathcal{U}_i$. Thus, the (C, A) pair is detectable if and only if $\cap_{i=1}^N \mathcal{U}_i = \{0\}$. ◊

7.2.2 Distributed Observer Design

From the previous section, the plant state is decomposed into the detectable part and the undetectable part. Since each agent could solely estimate the information on the detectable part at most, the information on the undetectable part should be obtained from other agents via the network communication. To maximally utilize the communication information on the undetectable subspace, we adopt the multi-step coupling framework in Chapter 6, i.e., every agent exchanges the information with its neighbors for each fraction count k, while the system updates its state and output at the integer count t.

Let us propose a distributed state observer for the system (7.1.1) and (7.1.2)

whose arbitrary i-th local observer is given by

$$\hat{\chi}_i[t_{k+1}] = \begin{cases} A\hat{\chi}_i[t_k] + H_i(y_i[t] - C_i\hat{\chi}_i[t_k]), & \text{if } k = 0, \quad (7.2.2a) \\ \hat{\chi}_i[t_k] + U_iU_i^\top \sum_{j\in\mathcal{N}_i} w_{ij}(\chi_j[t_k] - \chi_i[t_k]), & \text{if } k = 1,\dots,K-1, \quad (7.2.2b) \end{cases}$$

where $\hat{\chi}_i \in \mathbb{R}^n$ is the estimation for χ of the i-th local observer, w_{ij} is the coupling weight, and $H_i \in \mathbb{R}^{n\times \mathrm{r}_i}$ is the injection gain matrix which is designed by

$$H_i := \begin{bmatrix} D_i & U_i \end{bmatrix} \begin{bmatrix} H_{id} \\ 0 \end{bmatrix},$$

where $H_{i\mathrm{d}}$ is the stabilizing gain for $(C_{i\mathrm{d}}, A_{i\mathrm{d}})$ pair, i.e., $\rho(A_{i\mathrm{d}} - H_{i\mathrm{d}}C_{i\mathrm{d}}) < 1$. Here, (7.2.2a) is a typical state observer composed of the system copy term and the error correction term. Moreover, the coupling dynamics (7.2.2b) can be rewritten as

$$\begin{aligned} \hat{\chi}_i[t_{k+1}] &= \hat{\chi}_i[t_k] + U_iU_i^\top \sum_{j\in\mathcal{N}_i} w_{ij}(\chi_j[t_k] - \chi_i[t_k]) \\ &= (I - \mathcal{P}_i)\hat{\chi}_i[t_k] + \mathcal{P}_i \sum_{j\in\mathcal{N}_i} w_{ij}\chi_j[t_k], \end{aligned}$$

where $\mathcal{P}_i := U_iU_i^\top$ is the projection matrix into $\mathcal{U}_i$, the undetectable space of (C_i, A) pair. This implies that the coupling dynamics (7.2.2b) makes each local observer compensates the lacking information on the undetectable part through the network communication while retaining the state estimation on detectable part.

The coupling weight w_{ij} is chosen such that the weight matrix $W = [w_{ij}]$ is a row-stochastic matrix to synchronize all agents. For example, the Metropolis-Hastings coupling weight w_{ij}^{MH} in Section 4.3.1 or the average coupling weight w_{ij}^{avg} for the average consensus protocol in Section 4.3.3 can be adopted for the coupling weight w_{ij} of (7.2.2). Indeed, [WLMA19] proposed a two-time scaled distributed observer for a discrete-time linear system using the average coupling weigh w_{ij}^{avg}.

Let the estimation error variable $x_i := \hat{\chi}_i - \chi_i$ of the i-th observer. Then, the

error dynamics of x_i is written by

$$x_i[t_{k+1}] = \begin{cases} (A - H_i C_i)x_i[t_k], & \text{if } k = 0, \\ x_i[t_k] + U_i U_i^\top \sum_{j \in \mathcal{N}_i} w_{ij}(x_j[t_k] - x_i[t_k]), & \text{if } k = 1, \dots, K-1. \end{cases} \tag{7.2.3}$$

It should be emphasized that (7.2.3) has the form of multi-step coupling framework under rank-deficient coupling, i.e., (6.1.1). In particular, the rank-deficient coupling matrix B_i becomes $U_i U_i^\top$ in the problem of the distributed observer so that $R_i = U_i$, $Z_i = D_i$, and $\Lambda_i = I$ for all $i \in \mathcal{N}$.

In order to predict the behavior of the network of (7.2.3), the blended dynamics (6.3.1) can be useful tool. First, let us consider V whose columns are the orthonormal bases for the kernel of $(W - I_N)\otimes n\mathring{U}\Xi$, where $\mathring{U} = \mathrm{diag}(U_i)$ and $\Xi = \mathrm{diag}(q_i^{-1/2} I_{\mathrm{r}_i})$ with $\mathrm{r}_i = \mathrm{rank}(U_i U_i^\top)$. Since $\ker\big((W - I_N)_{\otimes n}\big) = \mathrm{im}(\mathbb{1}_N \otimes I_n)$ for row-stochastic W, it follows

$$\begin{aligned} &\ker\big((W - I_N)_{\otimes n}\mathring{U}\Xi\big) \\ &= \{v \in \mathbb{R}^{\bar{\mathrm{r}}} : \mathring{U}\Xi v \in \mathrm{im}(\mathbb{1}_N \otimes I_n)\} \\ &= \{\mathrm{col}(v_1, \cdots, v_N) \in \mathbb{R}^{\bar{r}} : q_1^{-1/2} U_1 v_1 = \cdots = q_N^{-1/2} U_N v_N\} \\ &= \{v \in \mathbb{R}^{\bar{\mathrm{r}}} : \mathring{U}\Xi v = \mathbb{1}_N \otimes c \text{ for some } c \in \cap_{i=1}^N \mathrm{im}(U_i)\}, \end{aligned}$$

where $\bar{\mathrm{r}} = \sum_{i=1}^N \mathrm{r}_i$. From this, the dimension of $\ker\big((W - I_N)_{\otimes n}\mathring{R}\Xi\big)$ (i.e., r_0) is given by

$$\begin{aligned} \mathrm{r}_0 &= \dim \ker\big((W - I_N)_{\otimes n}\mathring{U}\Xi\big) \\ &= \dim \cap_{i=1}^N \mathrm{im}(U_i) \\ &= 0 \end{aligned}$$

where the last equality comes from Lemma 7.2.2, i.e., $\cap_{i=1}^N \mathrm{im}(U_i) = 0$ because (C, A) pair is detectable. This implies that V is null, and thus, $\bar{V}$ can be chosen as the identity matrix.

Based on the aforementioned observations, the blended dynamics for (7.2.3)

is given by, with the state $\hat{r}$ being null,

$$\begin{aligned}\hat{z}_i[t+1] &= D_i^\top (A - H_i C_i)(D_i \hat{z}_i[t] - U_i \Xi_i L_i \hat{z}[t]) \\ &= D_i^\top \begin{bmatrix} D_i & U_i \end{bmatrix} \left\{ \begin{bmatrix} A_{i\mathrm{d}} & 0 \\ A_{i\mathrm{r}} & A_{i\mathrm{u}} \end{bmatrix} - \begin{bmatrix} H_{i\mathrm{d}} \\ 0 \end{bmatrix} \begin{bmatrix} C_{i\mathrm{d}} & 0 \end{bmatrix} \right\} \begin{bmatrix} D_i^\top \\ U_i^\top \end{bmatrix} \\ &\quad \times (D_i \hat{z}_i[t] - U_i \Xi_i L_i \hat{z}[t]) \\ &= (A_{i\mathrm{d}} - H_{i\mathrm{d}} C_{i\mathrm{d}}) \hat{z}_i[t],\end{aligned}$$

for all $i \in \mathcal{N}$. The last equality comes from $D_i^\top U_i = O$. It should be emphasized that $\hat{z}_i$ asymptotic converges to $\mathbb{0}$ because $L_{i\mathrm{d}}$ is chosen such that $\rho(A_{i\mathrm{d}} - H_{i\mathrm{d}} C_{i\mathrm{d}}) < 1$.

Now, by applying Theorem 6.3.1 in the previous chapter, the performance of the proposed observer (7.2.2) is given as follows.

Corollary 7.2.3. For any $\epsilon > 0$, there exists a $K^{\min}$ such that for all $K > K^{\min}$,

$$\limsup_{t\to\infty} \left\| \hat{\chi}_i[t] - \chi[t] \right\| < \epsilon, \quad \forall i \in \mathcal{N}. \qquad \Diamond$$

Corollary 7.2.3 guarantees the performance of the proposed observer by the arbitrary ϵ error level by choosing sufficiently large K. However, it does not guarantees the asymptotic performance of the proposed observer as (7.1.3).

Nevertheless, we can guarantee the asymptotic performance of the proposed observer in the following theorem.

Theorem 7.2.4. For any $\epsilon > 0$, there exists a $K^{\min}$ such that for all $K > K^{\min}$,

$$\lim_{t\to\infty} \left\| \hat{\chi}[t] - \chi[t] \right\| = \mathbb{0}, \quad \forall i \in \mathcal{N}. \qquad \Diamond$$

Proof. In the distributed state observer problem, recall that V is null so that $\bar{V}$ can be chosen as the identity matrix, $R_i = U_i$, $Z_i = D_i$, and $\Lambda_i = I$ for all $i \in \mathcal{N}$.

From this, the coordinate change proposed in Section 6.2 is written by

$$\begin{aligned}\begin{bmatrix} z \\ \tilde{r} \end{bmatrix} &= \begin{bmatrix} \mathring{D}^\top \\ S^{-1}\Xi^{-1}\mathring{U}^\top (W - I_N)_{\otimes n} \end{bmatrix} \bar{x} \\ \bar{x} &= \begin{bmatrix} \mathring{D} - \mathring{U}\Xi L & \mathring{U}\Xi \end{bmatrix} \begin{bmatrix} z \\ \tilde{r} \end{bmatrix}\end{aligned} \tag{7.2.4}$$

where $z \in \mathbb{R}^{nN-\bar{r}}$, $\tilde{r} \in \mathbb{R}^{\bar{r}}$, $\mathring{D} := \mathrm{diag}(D_i)$, $\mathring{U} := \mathrm{diag}(U_i)$, and

$$\begin{aligned} S &= \Xi^{-1}\mathring{U}^\top (W - I_N)_{\otimes n}\mathring{U}\Xi, \\ L &= S^{-1}\Xi\mathring{U}^\top (W - I_N)_{\otimes n}\mathring{D}. \end{aligned}$$

Similarly with (6.2.8), by the coordinate change (7.2.4), the error dynamics of x_i, (7.2.3), is transformed into

$$\begin{aligned} z_i[t+1] &= D_i^\top (A - H_i C_i)(D_i z_i[t] - U_i \Xi_i L_i z[t] + U_i \Xi_i \tilde{r}[t]) \\ &= D_i^\top \begin{bmatrix} D_i & U_i \end{bmatrix} \left\{ \begin{bmatrix} A_{i\mathrm{d}} & 0 \\ A_{i\mathrm{r}} & A_{i\mathrm{u}} \end{bmatrix} - \begin{bmatrix} H_{i\mathrm{d}} \\ 0 \end{bmatrix} \begin{bmatrix} C_{i\mathrm{d}} & 0 \end{bmatrix} \right\} \begin{bmatrix} D_i^\top \\ U_i^\top \end{bmatrix} \\ &\qquad \times (D_i z_i[t] - U_i \Xi_i L_i z[t] + U_i \Xi_i \tilde{r}[t]) \\ &= (A_{i\mathrm{d}} - H_{i\mathrm{d}} C_{i\mathrm{d}}) z_i[t], \\ \tilde{r}[t+1] &= (I + S)^{K-1} S^{-1} \Xi^{-1} \mathring{U}^\top (W - I_N)_{\otimes n} \\ &\qquad \times (\mathring{A} - \mathring{H}\mathring{C})(\mathring{D} z[t] - \mathring{U}\Xi L z[t] + \mathring{U}\Xi \tilde{r}[t]), \end{aligned} \tag{7.2.5}$$

where $\mathring{A} := \mathrm{diag}(A_i)$, $\mathring{H} := \mathrm{diag}(H_i)$, and $\mathring{C} := \mathrm{diag}(C_i)$. It should be emphasized that $\tilde{r}$ term does not appear in the z_i dynamics in (7.2.5) because $D_i^\top U_i = O$. This yields that z_i converges to $\mathbb{0}$ and we can adjust the convergence rate of z_i by choosing appropriate injection gain H_i (or $H_{i\mathrm{d}}$). Moreover, the convergence rate of $\tilde{r}$ could be arbitrarily chosen by getting sufficiently large K because $\rho(I + S) < 1$ by Lemma 7.2.1 and 7.2.2. This implies that, if we choose K such that the convergence rate of $\tilde{r}$ is faster than the convergence rate of z, then the error state

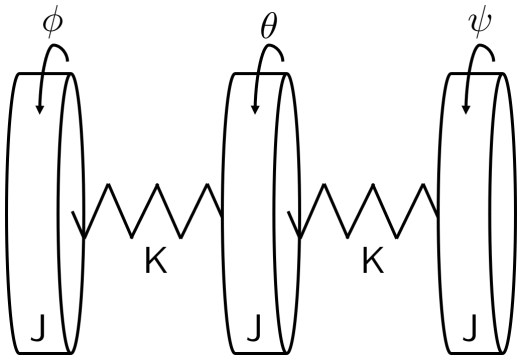

Figure 7.1: Three-inertia system considered in Section 7.3

x_i given by

$$x_i = D_i z_i - U_i \Xi_i L_i z + U_i \Xi_i \bar{V}_i \tilde{r},$$

also converges to $\mathbb{0}$ with the convergence rate dominated by the convergence rate of $(A_{i\mathrm{d}} - H_{i\mathrm{d}} C_{i\mathrm{d}})$, for $i \in \mathcal{N}$. In other wors, this allows

$$\lim_{t \to \infty} \left\| \hat{\chi}[t] - \chi[t] \right\| = \mathbb{0},$$

which complete the proof of Theorem 7.2.4. □

7.3 Simulation Results

For simulation, we consider a three-inertia system in Figure 7.1, which is monitored by three sensors. Each load has same inertia of J and they are connected by flexible connectors whose stiffness coefficient is K. In particular, by denoting the angles of the three inertia as ϕ, θ, and ψ, respectively, the sensors measure the outputs $y_1 = \phi - \theta$, $y_2 = \theta$, and $y_3 = \theta - \psi$. In addition, we assume that each sensor (or agent) is inter-connected by the structure of network which is shown in Figure 7.2. Note that the network is strongly connected by the ring network structure.

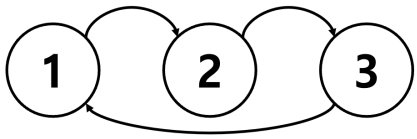

Figure 7.2: Communication network considered in Section 7.3

Let $\chi = \begin{bmatrix} \phi & \dot{\phi} & \theta & \dot{\theta} & \psi & \dot{\psi} \end{bmatrix}$, then the system dynamics is given by

$$\dot{\chi} = A_0 \chi,$$
$$y = C\chi,$$

where

$$A_0 := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{\mathsf{K}}{\mathsf{J}} & 0 & \frac{\mathsf{K}}{\mathsf{J}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{\mathsf{K}}{\mathsf{J}} & 0 & -\frac{2\mathsf{K}}{\mathsf{J}} & 0 & \frac{\mathsf{K}}{\mathsf{J}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{\mathsf{K}}{\mathsf{J}} & 0 & -\frac{\mathsf{K}}{\mathsf{J}} & 0 \end{bmatrix}$$
$$C = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}$$

This system can be discretized with the forward difference method as follows:

$$\chi[t+1] = (I + \Delta_t A_0)\chi[t]$$
$$=: A\chi[t],$$

where Δ_t is the sampling time.

It should be emphasized that any pair of (C_i, A) is not detectable, but the pair of (C, A) is detectable. In addition, each U_i whose columns are bases of the

undetectable subspace $\mathcal{U}_i$ for $i \in \mathcal{N}$ can be obtained as

$$U_1 = U_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

From these, we can take D_1, D_2, and D_3 such that $\begin{bmatrix} D_i & U_i \end{bmatrix}$, $i \in \mathcal{N}$ is an orthonormal matrix which satisfies (7.2.1) as follows:

$$D_1 = D_3 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{6} & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} & 0 & 0 \\ 0 & 0 & 0 & -2/\sqrt{6} \end{bmatrix},$$

$$D_2 = \begin{bmatrix} 0 & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1/\sqrt{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & 0 \\ 0 & 0 & 0 & 1/\sqrt{2} \end{bmatrix}.$$

In addition, the submatrices of the undetectable part, i.e., $A_{i\mathrm{u}}$ for $i \in \mathcal{N}$, are given by

$$A_{1\mathrm{u}} = A_{3\mathrm{u}} = \begin{bmatrix} 1 & \Delta_t \\ 0 & 1 \end{bmatrix},$$

$$A_{2\mathrm{u}} = \begin{bmatrix} 1 & \Delta_t \\ -\Delta_t \mathsf{K}/\mathsf{J} & 1 \end{bmatrix}.$$

With $\mathsf{K} = 1$, $\mathsf{J} = 1$, and $\Delta_t = 0.1$, we choose the design parameter K as 2 in

the simulation. In addition, the injection gains L_1, L_2, and L_3 are chosen such that, with $L_i = D_i L_{i\mathrm{d}}$, the eigenvalues of $A_{i\mathrm{d}} - L_{i\mathrm{d}} C_{i\mathrm{d}}$ are 0.85, 0.8, 0.75, and 0.7 for all $i \in \mathcal{N}$. The simulation results are shown in Figure 7.3. It shows that the estimation of every agent (or local observer) asymptotically converges to the plant state.

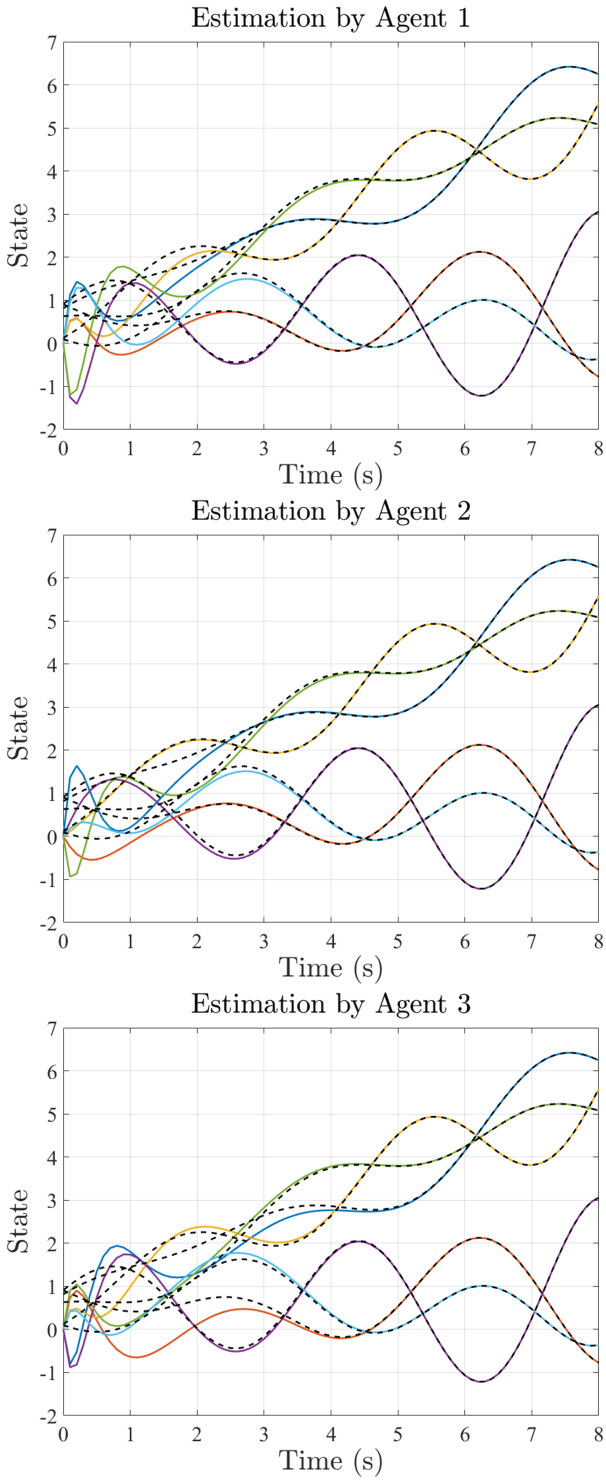

Figure 7.3: State estimation by the local observer (7.2.2) of each agent for the three-inertia system (7.1.1): Black dashed curves represent the true states of the plant and colored solid curves represent their estimations.

Chapter 8

Conclusions and Further Issues

8.1 Conclusions

The main objective of this dissertation is to provide a distributed design methodology by utilizing the discrete-time blended dynamics which predicts the behavior of a discrete-time heterogeneous multi-agent system under the multi-step coupling. The details are listed as follows.

i) In Chapter 3, we have reviewed an initial studies on synchronization of heterogeneous multi-agent system. Strong coupling can be considered as an alternative to achieve practical synchronization among the heterogeneous multi-agent system. Based on this observation, the continuous-time blended dynamics is proposed, which allows to predict an overall behavior of the network. In the sense that this blended dynamics approach only requires the stability of the blended dynamics, unstable agents are allowed as long as their instability is compensated by other agents' stability. In particular, since the blended dynamics is a simple average of individual node dynamics, it has been successfully employed as a design tool for many distributed algorithms by designing a desired algorithm as the blended dynamics first, and then, splitting it into different node dynamics with the diffusive coupling. However, while all the results are in the continuous-time domain, it is required to implement the designed algorithm in the discrete-time domain so that it operates on digital devices in practice. A naive idea such as using simple dis-

cretization methods (e.g., forward difference) cannot be employed for the implementation of the blended dynamics approach because increasing coupling gain yields instability of the network unless the sampling time is decreased. This motivates the development of the discrete-time version of the blended dynamics approach which is introduced in the next chapter.

ii) In Chapter 4, we have considered a discrete-time heterogeneous multi-agent system where a weighted averaging action is repeated for many times before progressing through the heterogeneous individual node dynamics (i.e., multi-step coupling). This multi-step coupling concept corresponds to the strong coupling in the sense that the synchronization is taken more care of than the progress through the node dynamics. Moreover, by proposing the discrete-time blended dynamics, we have illustrated the behavior of the multi-agent system can be approximated by the the solution of the blended dynamics. This approach maintains the advantages of the continuous-time case, such as the plug-and-play operation, and that the individual node dynamics need not be stable as long as the blended dynamics is stable. Moreover, while the continuous-time approach predicted collective synchronization behavior of the multi-agent system, this discrete-time approach estimates not only emergent but also individually scaled behavior, i.e., each agent behaves similarly to the solution of the blended dynamics with an agent-wise scaling factor.

iii) In Chapter 5, to emphasize the benefit of the discrete-time blended dynamics approach where the blended dynamics can have more variety depending on the coupling matrix, we have proposed a distributed algorithm for estimation of PageRank scores by employing the multi-step coupling framework introduced in Chapter 4. In the proposed algorithm, each node estimates its relative importance which is possibly agent-wise different so that overall nodes are not synchronized in the network. Moreover, the proposed algorithm has an initialization-free benefit while most of distributed PageRank algorithms have assumed the initialization process. This also allows the algorithm to adopt plug-and-play feature, i.e., some nodes and their associate link can join or leave the network during the operation of the algorithm. Finally, we

have verified the aforementioned advantages of the proposed algorithm by simulation for a real data with large scale.

iv) In Chapter 6, the result of Chapter 4 has been extended. We have considered a heterogeneous multi-agent system under rank-deficient multi-step coupling and stated that its behavior can be approximated by the solution obtained from the proposed blended dynamics for the rank-deficient coupling. In particular, based on the analysis in Chapter 4, various properties of the proposed approach are also discussed. In addition, we have illustrated the blended dynamics for simple cases in the formulation of the problem.

v) In Chapter 7, in order to emphasize the practical utility of the proposed approach in Chapter 6, we have proposed a distributed state estimation algorithm for a linear time-invariant discrete-time plant whose partial measurements are monitored by sensors based on the rank-deficient multi-step coupling framework. We have designed the local observer for each agent such that it estimates the plant state on its detectable part while compensating the lacking information on undetectable space by network communication. Even though the result in Chapter 6 only guarantees the practical convergence among all agents in the network, it has been shown that the proposed observer guarantees the asymptotic performance. Finally, the performance of the proposed observer has been verified by the simulation of toy example.

8.2 Further Issues

So far we have proposed the discrete-time version of the blended dynamics as a design method for distributed algorithms as well as its numerous applications. However, the results of this dissertation might open up many questions for future research.

8.2.1 Extension to Asynchronous Communication

A distributed algorithm should operate in a distributed manner, i.e., each agent utilizes its local measurement only, and communicates with its neighboring

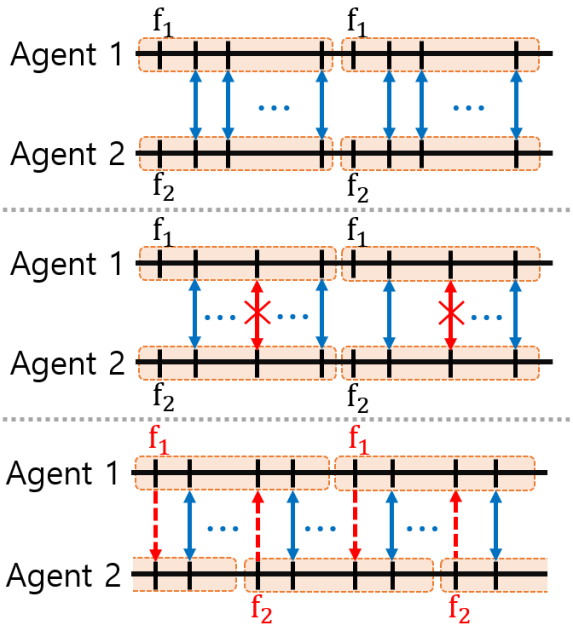

Figure 8.1: Comparison between synchronous operation and asynchronous operation for the multi-step coupling framework

agents. In addition, the design process of the algorithm should be decentralized in a way that the parameters of the algorithm are computed using the local knowledge accessible by each agent. This might yields an asynchronous operation of the communication protocol in the distributed algorithm because each agent could not be synchronized to a common clock shared by all other agents. For example, [FA05] studies a synchronization problem for the asynchronous multi-agent system where the order in which states of agents are updated is not fixed and the selection of previous values of the states used in the updates is also not fixed.

Recall that a weighted averaging coupling action (or communication) is repeated for multiple times for each progress of the node dynamics in the multi-step coupling framework. In the asynchronous setting, the communication failure or delay among agents could occur in the network, which is depicted as middle figure in Figure 8.1. Furthermore, an asynchronous update of the node dynamics for each agent is possible, i.e., some agents progress their node dynamics while other agents progress their coupling dynamics at the same fraction discrete-time index

t_k (shown in bottom of Figure 8.1).

Compared to results of the synchronous synchronization, [FA05] claims that, under bounded communication delay and an infinitely often updated agent, the overall network still achieves asymptotic synchronization while the communication delay or stochastic update progress varies where the agent converges. This makes challenging to predict the behavior of each agent by iterating asynchronous coupling dynamics. Moreover, the asynchronous update of the node dynamics and the coupling dynamics makes difficult to approximate the behavior of the network by the discrete-time blended dynamics so that it might require a modified blended dynamics.

8.2.2 Generalization of Multi-step Approach

In the continuous-time blended dynamics approach, [LS20] transformed a strongly coupled heterogeneous multi-agent system into the standard singularly perturbed form [Kha02]. The slow dynamics is a reduced-order multi-agent system consisting of a weighted average of the vector fields of all agents and some sub-dynamics of agents, which is defined as the blended dynamics.

Meanwhile, the multi-step coupled multi-agent system is proposed in this dissertation as a counterpart of the strongly coupled multi-agent system. In the sense that the strong coupling implies the communication by the synchronization protocol is much more taken care than the progress of node dynamics, the key idea to implement the strong coupling in the discreet-time domain is repeating the coupling progress for the multiple times for each progress of node dynamics. Indeed, by utilizing a similar multi-time scale approach, [WLMA19] proposed a discrete-time distributed state observer which is motivated by a continuous-time distributed observer designed through the continuous-time blended dynamics approach in [KLS19].

This observation might pave a road to develop a discrete-time approach for the singular perturbation theory. In particular, by adopting the fractional discrete-time index t_k in the multi-step coupling, it is expected that both fast and slow dynamics can be handled as the former and the latter proceed with the time-index based on the fraction count k and the integer count t, respectively.

Since the high-gain observer is a representative application example of the singular perturbation theory, the multi-step approach can be adopted to implement the high-gain observer in the discrete-time domain. For example, consider the following continuous-time plant

$$\begin{aligned} \dot{x}_1 &= x_2 \in \mathbb{R}, \\ \dot{x}_2 &= -x_1 - 2x_2 + ax_1^2 x_2 + b\sin 2t \in \mathbb{R}, \\ y &= x_1, \end{aligned} \tag{8.2.1}$$

where x_1 and x_2 are states of the plant, a and b are uncertain parameters, and y is output. It is well-known that the high-gain observer which robustly estimates the plant state can be designed as follows [Kha15]:

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + \frac{2}{\epsilon}(y - \hat{x}_1) \in \mathbb{R}, \\ \dot{\hat{x}}_2 &= -\hat{x}_1 - 2\hat{x}_2 + \hat{a}\hat{x}_1^2\hat{x}_2 + \hat{b}\sin 2t + \frac{1}{\epsilon^2}(y - \hat{x}_1) \\ &=: \hat{f}(t, \hat{x}) + \frac{1}{\epsilon^2}(y - \hat{x}_1) \in \mathbb{R}, \end{aligned}$$

where $\hat{x}_1$ and $\hat{x}_2$ are estimations, $\hat{a}$ and $\hat{b}$ are nominal values for a and b, and $\epsilon \ll 1$ is a design parameter.

We claim that the multi-step approach can be utilized to implement the discrete-time observer which robustly estimates the plant state. In specific, by assuming that ϵ is chosen as such that $K_\epsilon := 1/\epsilon$ is an integer, the observer can be designed as follows:

$$\hat{x}_1[t_{k+1}] = \begin{cases} \hat{x}_1[t_k] + \Delta_t \hat{x}_2[t_k], & \text{if } k = 0, \\ \hat{x}_1[t_k] + 2\Delta_t(y[t] - \hat{x}_1[t_k]), & \text{if } \mathrm{mod}(k, K_\epsilon) = 0, \\ \hat{x}_1[t_k], & \text{otherwise}, \end{cases} \tag{8.2.2}$$

$$\hat{x}_2[t_{k+1}] = \begin{cases} \hat{x}_2[t_k] + \Delta_t \hat{f}(t_k, \hat{x}[t_k]), & \text{if } k = 0, \\ \hat{x}_2[t_k] + \Delta_t(y[t] - \hat{x}_1[t_k]), & \text{otherwise}. \end{cases}$$

where Δ_t is the sampling time which is chosen as sufficiently small and the frac-

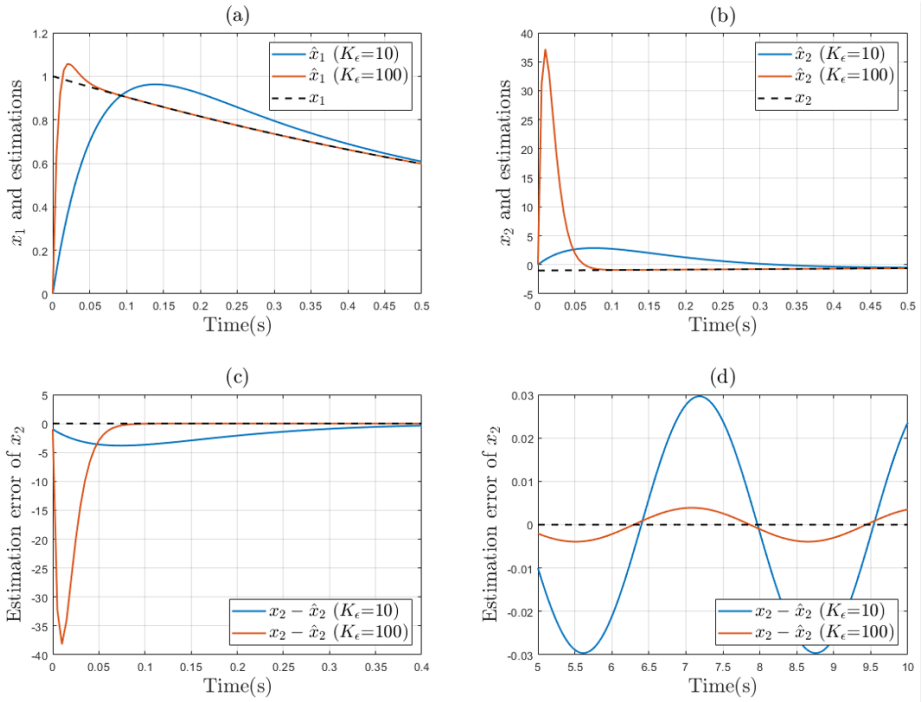

Figure 8.2: State estimation by the proposed observer (8.2.2) for the uncertain plant (8.2.1): Black dashed curve represents the true states of the plant and colored solid curves represent the estimations.

tional discrete-time index t_k is defined as

$$t_k = t + \frac{k}{K_\epsilon^2 + 1}$$

with the fraction count k varying from 0 to K_ϵ^2, e.g., t_k advances 0_0, 0_1, …, $0_{K_\epsilon^2}$, 1_0, 1_1, …. Note that plant updates its state and output for each integer count t so that they are fixed for the fraction count $k = 1, 2, \ldots, K_\epsilon^2$.

The simulation results for the proposed observer is shown in Figure 8.2. The proposed observer is designed for two distinct $K_\epsilon = 10$ and $K_\epsilon = 100$. The results show that, as K_ϵ increases, an estimation speed gets faster while its transient behavior exhibits a larger peak ("peaking phenomenon").

The proposed observer has benefit in a way that it could operate with low power by repeating an error correction into the multiple steps. While [AM15, AMT16] also proposed the low power high-gain observer in the continuous-time

domain, the proposed observer is designed in the discrete-time so that it could be useful to be implemented in practice. However, this work has a limitation of which, as K_ϵ increases, reducing the sampling time Δ_t with the same ratio is required in the proposed observer.

8.2.3 Extension to Nonlinear Coupling

In the multi-step coupling framework, a weighted averaging action is repeated before the progress by an individual node dynamics. In the sense that this dissertation only handles a linear coupling, it seems that nonlinear couplings including Kuramoto model [ABV$^+$05] might be a natural next step. One possible approach to extend the discrete-time blended dynamics theorems to the nonlinear couplings might come from [LS17] which studies a practical synchronization of the network under a nonlinear coupling called "funnel coupling". In particular, it could be useful material in the sense that it contains comparison between the result of the funnel coupling and the strong diffusive coupling.

BIBLIOGRAPHY

[ABV+05] Juan A Acebrón, Luis L Bonilla, Conrad J Pérez Vicente, Félix Ritort, and Renato Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of modern physics*, 77(1):137, 2005.

[AM15] Daniele Astolfi and Lorenzo Marconi. A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, 60(11):3059–3064, 2015.

[AMT16] Daniele Astolfi, Lorenzo Marconi, and Andrew Teel. Low-power peaking-free high-gain observers for nonlinear systems. In *Proceedings of the 15th European Control Conference (ECC)*, pages 1424–1429. IEEE, 2016.

[AS20] Sulaiman A Alghunaim and Ali H Sayed. Linear convergence of primal–dual gradient methods and their performance in distributed optimization. *Automatica*, 117:109003, 2020.

[BAMJ11] Carlos Baquero, Paulo Sérgio Almeida, Raquel Menezes, and Paulo Jesus. Extrema propagation: Fast distributed estimation of sums and network sizes. *IEEE Transactions on Parallel and Distributed Systems*, 23(4):668–675, 2011.

[Ber09] Dennis S Bernstein. Matrix mathematics. In *Matrix Mathematics*. Princeton university press, 2009.

[Bla08] Paul E Black. big-o notation, dictionary of algorithms and data structures. *US National Institute of Standards and Technology*, 2008.

[BP98] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

[Bul19] Francesco Bullo. *Lectures on network systems.* Kindle Direct Publishing, 1 edition, 2019.

[CXMR07] Peng Chen, Huafeng Xie, Sergei Maslov, and Sidney Redner. Finding scientific gems with google's PageRank algorithm. *Journal of Informetrics*, 1(1):8–15, 2007.

[Die17] Reinhard Diestel. *Graph Theory.* Springer-Verlag, Berlin, 2017.

[FA05] Lei Fang and Panos J Antsaklis. Information consensus of asynchronous discrete-time multi-agent systems. In *Proceedings of American Control Conference*, pages 1883–1888. IEEE, 2005.

[Fie73] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[Gle15] David F Gleich. PageRank beyond the web. *SIAM REVIEW*, 57(3):321–363, 2015.

[GR01] Chris Godsil and G. Royle. *Algebraic Graph Theory.* Springer, 2001.

[Hal97] Jack K Hale. Diffusive coupling, dissipation, and synchronization. *Journal of Dynamics and Differential Equations*, 9(1):1–52, 1997.

[HJ19] Roger A Horn and Charles R Johnson. *Matrix Analysis.* Cambridge University Press, 2019.

[HP14] Frank Harary and Edgar M Palmer. *Graphical Enumeration.* Elsevier, 2014.

[IT10] Hideaki Ishii and Roberto Tempo. Distributed randomized algorithms for the PageRank computation. *IEEE Transactions on Automatic Control*, 55(9):1987–2002, 2010.

[IT14] Hideaki Ishii and Roberto Tempo. The PageRank problem, multiagent consensus, and web aggregation: A systems and control viewpoint. *IEEE Control Systems Magazine*, 34(3):34–53, 2014.

[ITB12] Hideaki Ishii, Roberto Tempo, and Er-Wei Bai. A web aggregation approach for distributed randomized PageRank algorithms. *IEEE Transactions on Automatic Control*, 57(11):2703–2717, 2012.

[Jak18] Dušan Jakovetić. A unification and generalization of exact distributed first-order methods. *IEEE Transactions on Signal and Information Processing over Networks*, 5(1):31–46, 2018.

[Kha02] Hassan K Khalil. *Nonlinear systems (3rd ed.)*. Prentice-Hall, 2002.

[Kha15] Hassan K Khalil. *Nonlinear control*, volume 406. Pearson New York, 2015.

[KLLS22] Jeong Woo Kim, Jin Gyu Lee, Donggil Lee, and Hyungbo Shim. A design method of distributed algorithms via discrete-time blended dynamics theorem. *arXiv preprint arXiv:2210.05142*, 2022.

[KLS19] Taekyoo Kim, Chanhwa Lee, and Hyungbo Shim. Completely decentralized design of distributed observer for linear systems. *IEEE Transactions on Automatic Control*, 65(11):4664–4678, 2019.

[KLS20] Taekyoo Kim, Donggil Lee, and Hyungbo Shim. Decentralized design and plug-and-play distributed control for linear multi-channel systems. *arXiv preprint arXiv:2011.09735*, 2020.

[KSS11] Hongkeun Kim, Hyungbo Shim, and Jin Heon Seo. Output consensus of heterogeneous uncertain linear multi-agent systems. *IEEE Transactions on Automatic Control*, 56(1):200–206, 2011.

[KYS+16] Jaeyong Kim, Jongwook Yang, Hyungbo Shim, Jung-Su Kim, and Jin Heon Seo. Robustness of synchronization of heterogeneous agents by strong coupling and a large number of agents. *IEEE Transactions on Automatic Control*, 61(10):3096–3102, 2016.

[LBNVdS05] Xiaoming Liu, Johan Bollen, Michael L Nelson, and Herbert Van de Sompel. Co-authorship networks in the digital library research community. *Information Processing and Management*, 41(6):1462–1480, 2005.

[LC14] Jinlong Lei and Han-Fu Chen. Distributed randomized PageRank algorithm based on stochastic approximation. *IEEE Transactions on Automatic Control*, 60(6):1641–1646, 2014.

[LFLG15] Shaobao Li, Gang Feng, Xiaoyuan Luo, and Xinping Guan. Output consensus of heterogeneous linear discrete-time multiagent systems with structural uncertainties. *IEEE Transactions on Cybernetics*, 45(12):2868–2879, 2015.

[LKS20] Jin Gyu Lee, Junsoo Kim, and Hyungbo Shim. Fully distributed resilient state estimation based on distributed median solver. *IEEE Transactions on Automatic Control*, 65(9):3935–3942, 2020.

[LLKS18] Donggil Lee, Seungjoon Lee, Taekyoo Kim, and Hyungbo Shim. Distributed algorithm for the network size estimation: Blended dynamics approach. In *Proceedings of the 57th IEEE Conference on Decision and Control (CDC)*, pages 4577–4582, 2018.

[LS98] Winfried Lohmiller and Jean-Jacques E Slotine. On contraction analysis for non-linear systems. *Automatica*, 34(6):683–696, 1998.

[LS17] Jin Gyu Lee and Hyungbo Shim. Synchronized behavior of heterogeneous multi-agent systems under funnel coupling. In *Proceedings of the 2nd International Symposium on Swarm Behavior and Bio-Inspired Robotics*, pages 266–269, 2017.

[LS19] Jin Gyu Lee and Hyungbo Shim. A distributed algorithm that finds almost best possible estimate under non-vanishing and time-varying measurement noise. *IEEE Control Systems Letters*, 4(1):229–234, 2019.

[LS20] Jin Gyu Lee and Hyungbo Shim. A tool for analysis and synthesis of heterogeneous multi-agent systems under rank-deficient coupling. *Automatica*, 117:108952, 2020.

[LS21] Jin Gyu Lee and Hyungbo Shim. Design of heterogeneous multi-agent system for distributed computation. In Zhong-Ping Jiang, Christophe Prieur, and Alessandro Astolfi, editors, *Trends in Nonlinear and Adaptive Control*, volume 488 of *Lecture Notes in Control and Information Sciences*, pages 83–108. Springer, 2021.

[LS22] Seungjoon Lee and Hyungbo Shim. Blended dynamics approach to distributed optimization: Sum convexity and convergence rate. *Automatica*, 141:110290, 2022.

[LZD17] Constantino M Lagoa, Luca Zaccarian, and Fabrizio Dabbene. A distributed algorithm with consistency for PageRank-like linear algebraic systems. *IFAC-PapersOnLine*, 50(1):5172–5177, 2017.

[MRC15] Jie Mei, Wei Ren, and Jie Chen. Distributed consensus of second-order multi-agent systems with heterogeneous unknown inertias and control gains under a directed graph. *IEEE Transactions on Automatic Control*, 61(8):2019–2034, 2015.

[MS16] Aritra Mitra and Shreyas Sundaram. An approach for distributed state estimation of LTI systems. In *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing*, pages 1088–1093, 2016.

[MS18] Aritra Mitra and Shreyas Sundaram. Distributed observers for lti systems. *IEEE Transactions on Automatic Control*, 63(11):3689–3704, 2018.

[NL18] Angelia Nedić and Ji Liu. Distributed optimization for control. *Annual Review of Control, Robotics, and Autonomous Systems*, 1:77–103, 2018.

[NO09] Angelia Nedic and Asuman Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009.

[OPA15] Kwang-Kyo Oh, Myoung-Chul Park, and Hyo-Sung Ahn. A survey of multi-agent formation control. *Automatica*, 53:424–440, 2015.

[OSFM07] Reza Olfati-Saber, J Alex Fax, and Richard M Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[PL17] Elena Panteley and Antonio Loría. Synchronization and dynamic consensus of heterogeneous networked systems. *IEEE Transactions on Automatic Control*, 62(8):3758–3773, 2017.

[PWN05] Alexey Pavlov, Nathan van de Wouw, and Henk Nijmeijer. Convergent systems: analysis and synthesis. In *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems*, pages 131–146. Springer, 2005.

[QL17] Guannan Qu and Na Li. Harnessing smoothness to accelerate distributed optimization. *IEEE Transactions on Control of Network Systems*, 5(3):1245–1260, 2017.

[QL19] Guannan Qu and Na Li. Accelerated distributed nesterov gradient descent. *IEEE Transactions on Automatic Control*, 65(6):2566–2581, 2019.

[Rad11] Filippo Radicchi. Who is the best player ever? a complex network analysis of the history of professional tennis. *PloS one*, 6(2):e17249, 2011.

[RB05] Wei Ren and Randal W Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50(5):655–661, 2005.

[RBA05] Wei Ren, Randal W Beard, and Ella M Atkins. A survey of consensus problems in multi-agent coordination. In *Proceedings of American Control Conference*, pages 1859–1864. IEEE, 2005.

[Sak17] Kazunori Sakurama. Formation control of multi-agent systems with the scale freedom. In *Proceedings of the 2nd International Symposium on Swarm Behavior and Bio-Inspired Robotics*, pages 270–271, 2017.

[SCHJ12] Iman Shames, Themistoklis Charalambous, Christoforos N Hadjicostis, and Mikael Johansson. Distributed network size estimation and average degree estimation and control in networks isomorphic to directed graphs. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1885–1892. IEEE, 2012.

[SHM14] Valentin Schwarz, Gabor Hannak, and Gerald Matz. On the convergence of average consensus with generalized Metropolis-Hasting weights. In *Proceedings of IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, pages 5442–5446, 2014.

[SI19] Atsushi Suzuki and Hideaki Ishii. PageRank computation via web aggregation in distributed randomized algorithms. In *Proceedings of the 58th IEEE Conference on Decision and Control (CDC)*, pages 1856–1861. IEEE, 2019.

[SLWY15] Wei Shi, Qing Ling, Gang Wu, and Wotao Yin. Extra: An exact first-order algorithm for decentralized consensus optimization. *SIAM Journal on Optimization*, 25(2):944–966, 2015.

[TRK18] Duc N Tran, Björn S Rüffer, and Christopher M Kellett. Convergence properties for discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, 64(8):3415–3422, 2018.

[TSH12] Harry L Trentelman, Anton A Stoorvogel, and Malo Hautus. *Control theory for linear systems*. Springer Science & Business Media, 2012.

[VL05] Fabien Viger and Matthieu Latapy. Efficient and simple generation of random simple connected graphs with prescribed degree sequence. In *Proceedings of the International Computing and Combinatorics Conference*, pages 400–449. Springer, Berlin, 2005.

[Wie10] Peter Wieland. *From static to dynamic couplings in consensus and synchronization among identical and non-identical systems*. Logos Verlag Berlin GmbH, 2010.

[WLMA19] Lili Wang, Ji Liu, A Stephen Morse, and Brian DO Anderson. A distributed observer for a discrete-time linear system. In *Proceedings of the 58th IEEE Conference on Decision and Control (CDC)*, pages 367–372. IEEE, 2019.

[WSA11] Peter Wieland, Rodolphe Sepulchre, and Frank Allgöwer. An internal model principle is necessary and sufficient for linear output synchronization. *Automatica*, 47(5):1068–1074, 2011.

[XLF17] Xiang Xu, Lu Liu, and Gang Feng. Consensus of discrete-time linear multiagent systems with communication, input and output delays. *IEEE Transactions on Automatic Control*, 63(2):492–497, 2017.

[YCM20] Yamin Yan, Zhiyong Chen, and Richard H Middleton. Autonomous synchronization of heterogeneous multiagent systems. *IEEE Transactions on Control of Network Systems*, 8(2):940–950, 2020.

[YLY16] Kun Yuan, Qing Ling, and Wotao Yin. On the convergence of decentralized gradient descent. *SIAM Journal on Optimization*, 26(3):1835–1854, 2016.

[YSA19] Hyeonjun Yun, Hyungbo Shim, and Hyo-Sung Ahn. Initialization-free privacy-guaranteed distributed algorithm for economic dispatch problem. *Automatica*, 102:86–93, 2019.

[YTQ16] Keyou You, Roberto Tempo, and Li Qiu. Distributed algorithms for computation of centrality measures in complex networks. *IEEE Transactions on Automatic Control*, 62(5):2080–2094, 2016.

[ZBE12] Nazar Zaki, Jose Berengueres, and Dmitry Efimov. Detection of protein complexes using a protein ranking algorithm. *Proteins: Structure, Function, and Bioinformatics*, 80(10):2459–2468, 2012.

[ZLF18] Dan Zhang, Lu Liu, and Gang Feng. Consensus of heterogeneous linear multiagent systems subject to aperiodic sampled-data and dos attack. *IEEE Transactions on Cybernetics*, 49(4):1501–1511, 2018.

국문초록

Design of Distributed Algorithms via Discrete-time Blended Dynamics Theorem

이산시간 혼합 동역학 정리를 통한 분산 알고리즘의 설계

본 논문에서는 다양한 분산 알고리즘을 설계하는데 사용할 수 있는 이산시간에서 동작하는 혼합 동역학(blended dynamics) 이론을 개발한다. 혼합 동역학 이론은 이기종 다개체 시스템의 행동을 예측할 수 있다. 따라서, 만약 특정한 연산을 수행하는 혼합 동역학이 주어진다면, 이기종 개체의 개별 동역학을 설계하는 방법을 쉽게 얻을 수 있다. 연속시간 이론의 경우, 혼합 동역학에 의한 예측은 이웃한 개체들 사이의 높은 결합 이득(high coupling gain)에 의해서 가능했다. 본 논문의 이산시간 이론에서는 이와 상응하는 개념으로서 다단계 결합(multi-step coupling)을 제안한다. 이러한 이산시간 접근법은 플러그 앤 플레이(plug-and-play) 연산이 가능하거나 혼합 동역학이 안정하다면 개별 동역학이 안정하지 않아도 된다는 연속시간 이론의 장점들을 그대로 가진다.

한편, 연속시간 이론과 비교했을 때, 이산시간 혼합 동역학은 결합 행렬(coupling matrix)에 따라 더 많은 변형을 가질 수 있다. 이러한 장점은 네트워크 안에서 개별 노드의 상대적 중요도를 추정하는 PageRank를 분산적으로 추정하는 문제 등의 예제들을 통해 보여된다. 특히, 대부분의 다른 PageRank 분산 알고리즘들이 초기화 과정(initialization process)을 가정하는 반면에, 다단계 결합 접근법을 기반으로 제안된 알고리즘은 초기화 과정이 필요없다는 장점을 가진다.

더 나아가, 앞선 이산시간 혼합 동역학 이론의 결과를 상태정보의 교환이 제한된 환경(이러한 결합을 랭크 부족 결합(rank-deficient coupling)이라고 부른다)에서의 이기종 다개체 시스템에 대하여 확장한다. 이를 위해서, 다개체 시스템을 각 개체가 결합 동역학(coupling dynamics)을 다수 반복할수록 사라지는 동역학과 사라지지 않는 동역학을 분리하는 좌표 변환이 소개된다. 이러한 사라지지 않는 동역학을 기반으로 랭크 부족 결합에 대한 혼합 동역학을 유도하여 이기종 네트워크의 행동을 예측할 수 있다.

다시 한번 혼합 동역학의 실용적인 유용성을 강조하기 위하여, 다수의 센서가 부분적인 관측치를 통해 대상 플랜트의 상태변수를 분산적으로 추정하는 문제에 앞서

소개된 확장 결과를 적용한다. 분산 관측기(distributed observer)는 각 개체가 측정가능한(detectable) 영역에서의 상태정보를 추정하되 측정가능하지 않은(undetectable) 영역의 부족한 정보는 네트워크 통신을 통해 보완하도록 설계된다. 앞선 혼합 동역학의 결과가 오직 근사적 수렴만을 보장하지만, 제안된 관측기는 점근적인 성능을 보장함을 보인다.

주요어: 이산시간 이기종 다개체 시스템, 혼합 동역학, 다단계 결합, 다시간 척도, 열린 다개체 시스템
학　번: 2015–22779