論文 / 著書情報
Article / Book Information

題目(和文)	モデル低次元化に基づく大規模動的ネットワークの制御理論
Title(English)	Control theory for large-scale dynamical network systems based on model reduction techniques
著者(和文)	定本知徳
Author(English)	Tomonori Sadamoto
出典(和文)	学位:博士（工学）, 学位授与機関:東京工業大学, 報告番号:甲第9887号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:井村　順一,天谷　賢治,早川　朋久,中尾　裕也,山北　昌毅
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9887号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Tokyo Institute of Technology

Doctoral Thesis

Control theory for large-scale dynamical network systems based on model reduction techniques

Author:
Tomonori Sadamoto

Supervisor:
Dr. Jun-ichi Imura

Graduate School of Information Science and Engineering

February 2015

Abstract

This thesis provides a line of work for development of control theory for large-scale dynamical network systems such as distributed parameter systems and electric power networks. An observer and a controller for a large-scale network system are necessarily required to be low-dimensional compared with systems of interest and to guarantee an a priori performance of the whole network system. We consider constructing observers and controllers not only satisfying the above two requirements, but also having additional properties suitable for large-scale network systems. More specifically, we propose a novel low-dimensional observer to estimate average behavior of network systems from a macroscopic point of view where a set of states capturing the average behavior is systematically determined. Furthermore, we propose low-dimensional hierarchical distributed control where compositional controllers can be designed in a distributed manner. In contrast to existing distributed controller design methods where all compositional controllers have to be designed simultaneously, the distributed design property enables us to implement a control system in particular for a large-scale network system involving a number of subsystems. These proposed observers/controllers are expected to be useful for applications in various research fields, e.g., weather prediction and data-assimilation in meteorological engineering, and supply-demand balancing of power systems in electric power engineering.

Acknowledgements

First and foremost, I must thank my advisor Jun-ichi Imura for his guidance and support and for providing me this precious study opportunity as a doctoral student in his laboratory. All of my research topics in this thesis began by his insightful suggestions, and carried out by his proper advice. I truly appreciate that he gave me opportunity to pursue my own study and other projects.

I am deeply grateful to Associate Professor Takayuki Ishizaki for his enthusiastic support, proper advise and invaluable discussion that not only reveal the weaknesses in my study and make my research of great achievement. Without his support and widespread advice from mathematics to scientific writing and presentation, this work does not exist.

I am also very grateful to Associate Professor Tomohisa Hayakawa, Professor Kazuhiro Nakadai in Tokyo Institute of Technology, and Professor Karl Henrik Johansson, Professor Henrik Sandberg and Bart Besselink in KTH Royal Institute of Technology, for valuable comments and suggestions.

I am indebted to my previous advisor Associate Professor Masaki Yamakita who taught me the fundamentals of control theory. He encouraged me to pursue my study as a doctoral student.

I am very grateful to Noriko Sugimoto for her support of my life in laboratory. Furthermore, I am very grateful to Masakazu Koike, Osamu Sugiyama, Masayasu Suzuki and Masaki Inoue for their support, suggestions and delightful study life in laboratory. Moreover, I thank to current and past members of Imura, Hayakawa and Nakadai Laboratory many great discussions and all the good times. Furthermore, my thanks for my family will never be enough. Finally, I greatly appreciate Asako Yamada and her family for their deep understanding and heartwarming encouragement.

List of Publications

Journal Papers

- Tomonori Sadamoto, Takayuki Ishizaki, Jun-ichi Imura: Low-Dimensional Functional Observer Design of Large-Scale Systems, Transactions of the Society of Instrument and Control Engineers, vol. 50, no. 6, pp.478-486, 2014. (in Japanese) [Chapter 2]

- Tomonori Sadamoto, Takayuki Ishizaki, Jun-ichi Imura: Average State Observers for Large-Scale Networked Linear Systems, Transactions of the Institute of Systems, Control and Information Engineers. (in Japanese) (in Press) [Chapter 3]

- Takayuki Ishizaki, Tomonori Sadamoto, Jun-ichi Imura: Hierarchical Distributed Stabilization of Power Networks, The European Physical Journal Special Topics on Resilient power grids and extreme events, September 2014 (in Press) [Chapter 4]

- Tomonori Sadamoto, Kenji Kashima, Hiroshi Morita, Hiroyuki Mizuno: Nonlinear Reduced Order Modeling of Plasticization Cylinders, Transactions of the Institute of Systems, Control and Information Engineers, vol. 26, no. 5, pp.174-181, 2013. (in Japanese) [Chapter 6]

Reviewed Conference Papers

- Tomonori Sadamoto, Takayuki Ishizaki, Jun-ichi Imura: Low-Dimensional Functional Observer Design for Linear Systems via Observer Reduction Approach, Proc. of Conference on Decision and Control, pp.776-781, 2013. [Chapter 2]

- Tomonori Sadamoto, Takayuki Ishizaki, Jun-ichi Imura: Projective State Observers for Large-Scale Linear Systems, Proc. of European Control Conference, pp.2969-2974, 2014. [Chapter 3]

- Tomonori Sadamoto, Takayuki Ishizaki, Jun-ichi Imura: Hierarchical Distributed Control for Networked Linear Systems, Proc. of Conference on Decision and Control, 2014 (to appear) [Chapter 4]

- Tomonori Sadamoto, Kenji Kashima, Hiroshi Morita, Hiroyuki Mizuno: Nonlinear Reduced Order Modeling of Plasticization Cylinders, in Proc. of American Control Conference, pp. 129-134, 2014. [Chapter 6]

Contents

Abstract i

Acknowledgements ii

List of Publications iii

Contents v

Notations viii

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Contributions and Organization . . . 3
- 1.3 Organization . . . 5

2 Low-Dimensional Functional Observer Design 7
- 2.1 Introduction . . . 7
- 2.2 Problem Formulation . . . 8
  - 2.2.1 Design Problem of Low-dimensional Observers . . . 8
  - 2.2.2 Problem Formulation via Observer Reduction . . . 9
- 2.3 Low-dimensional Functional Observer Design . . . 11
  - 2.3.1 Error Analysis of Low-dimensional Functional Observers . . . 11
  - 2.3.2 Design Algorithm of Low-dimensional Functional Observers . . . 16
- 2.4 Numerical Simulation . . . 17
  - 2.4.1 Power Network Systems . . . 17
  - 2.4.2 Low-dimensional Functional Observer Design for Power Network Systems . . . 19
- 2.5 Chapter Summary . . . 24

3 Average State Observers 25
- 3.1 Introduction . . . 25
- 3.2 Fundamentals of Average State Observers . . . 26
  - 3.2.1 Review of Functional State Observers . . . 26
  - 3.2.2 Error Analysis of Average State Observers . . . 27
- 3.3 Design of Average State Observers . . . 29

3.3.1 A Road Map for Systematic Design . . . 29
3.3.2 Extension to Unstable Systems . . . 35
3.3.3 Design Algorithm of Average State Observers . . . 36
3.4 Numerical Simulation . . . 37
3.5 Chapter Summary . . . 39

4 Hierarchical Distributed Control 40
4.1 Introduction . . . 40
4.2 Problem Formulation . . . 41
4.2.1 Review of Decentralized Control . . . 41
4.2.2 Hierarchical Distributed Control Problem . . . 42
4.3 Hierarchical Distributed Control Systems . . . 44
4.3.1 Design of Hierarchical Distributed Controllers . . . 44
4.3.2 Integration with Hierarchical Distributed Observers . . . 48
4.3.3 Discussion on Hierarchical Clustering and Sensor and Actuator Allocation Towards Scalable Implementation . . . 49
4.4 Numerical Example . . . 51
4.4.1 Power Network Model . . . 51
4.4.2 Hierarchical Distributed Control of Power Networks . . . 53
4.5 Chapter Summary . . . 55

5 Low-dimensional Hierarchical Distributed Controller Design 56
5.1 Introduction . . . 56
5.2 Problem Formulation . . . 57
5.3 Main Results . . . 61
5.3.1 Analysis of Controller Reduction . . . 61
5.3.2 Stability Preservation of Low-dimensional Hierarchical Distributed Controller . . . 65
5.3.3 Design Algorithm of Low-dimensional Hierarchical Distributed Controller . . . 68
5.4 Numerical Example . . . 69
5.5 Chapter Summary . . . 71

6 Low-dimensional Nonlinear Modelling of Plasticization Cylinders 72
6.1 Introduction . . . 72
6.2 Nonlinear Modelling of Plasticization Cylinders . . . 73
6.2.1 Barrel, Outer air, Inner fluid and Water-cooling cylinder . . . 74
6.2.2 Modelling heaters . . . 76
6.2.3 Network System Given by Spatial discretization . . . 76
6.3 Experimental Results . . . 79
6.3.1 Configuration of Real Systems . . . 79
6.3.2 Model Validation by Experiment . . . 80
6.4 Model Order Reduction of Nonlinear Network System . . . 82
6.4.1 Error analysis . . . 83
6.4.2 Simulation results . . . 87
6.5 Chapter summary . . . 88

7 Conclusion 89

7.1 Summary of results 89
7.2 Future works 90

Bibliography **92**

Notations

$\mathbb{R}$	set of real numbers
$\mathbb{R}_+(\mathbb{R}_-)$	set of nonnegative (nonpositive) real numbers
$\mathbb{C}$	set of complex numbers
I_n	unit matrix of size $n \times n$
$0_{m\times n}$ (0_n)	zero matrix of size $m \times n$ $(n \times n)$
$\mathbf{1}_n$	n-dimensional one vector, i.e., $\mathbf{1}_n := [1, \ldots, 1]^\mathsf{T} \in \mathbb{R}^n$
e_i^n	the ith column of I_n
$e_\mathcal{I}^n$	$e_\mathcal{I}^n := [e_{i_1}^n, \ldots, e_{i_m}^n]$ for $i \in \mathcal{I} := \{i_1, \ldots, i_m\}$
$\lvert\mathcal{I}\rvert$	the cardinality of a set $\mathcal{I}$
$M \prec 0_n$ $(M \succ 0_n)$	negative (positive) definiteness of a symmetric matrix $M \in \mathbb{R}^{n\times n}$
$M \preceq 0_n$ $(M \succeq 0_n)$	negative (positive) semidefiniteness of a symmetric matrix $M \in \mathbb{R}^{n\times n}$
$\mathrm{im}(M)$	range space spanned by the column vectors of a matrix M
$\mathrm{tr}(M)$	trace of a matrix M
$\mathrm{dg}(M_i)_{i\in\mathcal{N}}$	the block-diagonal matrix having matrices $M_1, \ldots, M_N$ on its diagonal blocks for $\mathcal{N} = \{1, \ldots, N\}$
$M_{\frac{1}{2}}$	Cholesky factor of semipositive matrix $M \succeq 0_n$, i.e., $M_{\frac{1}{2}} 0_n$ such that $M = M_{\frac{1}{2}}^\mathsf{T} M_{\frac{1}{2}}$
$\lVert M\rVert_F$	the Frobenius norm of a matrix M
$\lVert M\rVert$	the induced 2-norm of a matrix M
$\lVert v(t)\rVert_{\mathcal{L}_2}$	the $\mathcal{L}_2$-norm of a square integrable function $v(t) \in \mathbb{R}^n$, i.e., $\lVert v(t)\rVert_{\mathcal{L}_2} := \left(\int_0^\infty v^\mathsf{T}(t)v(t)dt\right)^{\frac{1}{2}}$
$\lVert G(s)\rVert_{\mathcal{H}_\infty}$	the $\mathcal{H}_\infty$-norm of a stable proper transfer matrix G, i.e., $\lVert G(s)\rVert_{\mathcal{H}_\infty} := \sup_{\omega\in\mathbb{R}} \lVert G(j\omega)\rVert$
$\lVert G(s)\rVert_{\mathcal{H}_2}$	the $\mathcal{H}_2$-norm of a stable strictly proper transfer matrix G, i.e, $\lVert G(s)\rVert_{\mathcal{H}_2} := \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{tr}(G(j\omega)G^\mathsf{T}(-j\omega))d\omega\right)^{\frac{1}{2}}$

Chapter 1

Introduction

1.1 Background

As technology advances, systems to be dealt with become more complex and larger in scale. For example in meteorological engineering [1, 2, 3, 4, 5, 6], we use system models constructed by spatial discretization [7] of distributed parameter systems such as thermal diffusion systems [8] and Navier-Stokes equations [7, 9]. For prediction and data-assimilation with satisfactory accuracy, these systems necessarily contain hundreds of thousands of equations. In addition, in electric power engineering [10, 11, 12, 13, 14], we are required to maintain supply-demand balance of power network systems involving more than one million consumers and a number of power plants towards efficient use of renewable energy resources. In view of this, observers and controllers for large-scale network systems tend to play important roles in various research field. Throughout this thesis, we call *network systems* as dynamical systems evolving on networks, e.g., electric power networks, spatially discretized systems and gene regulatory networks [15, 16].

In this thesis, we consider the following problem: What is a desirable observer and controller suitable for handling large-scale network systems? Necessary requirements for observers/controllers for large-scale network systems include:

(i) lower-dimensionality compared to systems of interest, and

(ii) existence of an a priori performance evaluation.

The requirement (i) is natural and indispensable from the viewpoint of computational costs for implementation [17, 18]. The significance of the requirement (ii) is as follows: One naive approach to construct a low-dimensional observer/controller is to design a observer/controller for a low-dimensional model that approximates the behavior of the

system of interest in an appropriate sense; see e.g., [4, 19] for observer design. To construct such approximate model, model reduction techniques are available developed in literature; see, e.g., [17, 18, 20]. However, in this approach, it is difficult to construct low-dimensional observers/controllers achieving a prescribed performance, e.g., the $\mathcal{L}_2$-norm of estimation error and the convergence rate. This is because the relation between the approximation error and the performance is not explicitly taken into account. In view of this, the theoretical guarantee of performances (i.e., the requirement (ii)) is needed for observers/controllers.

Now, are the above two requirements sufficient for observers/controllers for large-scale network systems? To consider this, we review several existing design methods satisfying the two requirements. These methods can be categorized into two groups: those designing a centralized low-dimensional observer/controller and those designing multiple decentralized/distributed observers and controllers.

As related works for the former methods, [21, 22, 23] provide low-dimensional controller design based on numerical optimization. However, the optimization problem is not computationally friendly in general because that is non-convex due to rank constraints. In literature, e.g. [24, 25, 26, 27], low-dimensional controller design based on model reduction techniques, which are computationally tractable, is proposed. For example, [24] has shown that low-dimensional controller design via controller reduction where a low-dimensional controller approximates an original controller for the system of interest can be solved by frequency weighted model reduction techniques such as weighted-balanced truncation [28] or weighted-optimal Hankel norm approximation [29]. See, e.g., [25] for fundamentals. For low-dimensional observer design, to the best of our knowledge, there are no methods with explicit consideration of influences of the approximation error on the estimation error. Furthermore, since it is difficult to exactly estimate all of system states by using low-dimensional observers in general, it is significant to determine a few estimation signals that are suitable for capturing behavior of large-scale network systems.

Next, let us quickly review decentralized/distributed estimation and control where individual components are low-dimensional while they achieve some performances of whole large-scale network systems. In [30], decentralized observer has been proposed in a systematic fashion, yet the resultant observers are often conservative from the viewpoint of estimation performance. This is due to the fact that the interaction among subsystems is not dealt with quantitatively. See, e.g.,[31] for survey. Furthermore, in [32] for example, the authors have developed a method synthesizing a distributed dynamic output feedback controller achieving $\mathcal{H}_\infty$ performance for network systems on the basis of dissipativity theory [33]. The optimal distributed controller can be obtained by solving

a couple of LMIs simultaneously. In addition, [34] characterizes a class of convex problems in decentralized optimal control. See, e.g. [35, 36] for applications and surveys. However, these methods require us to design distributed observers/controllers in a centralized manner. Thus, in particular for large-scale network systems (i.e., that involves a number of subsystems), these methods do not fully fit for practical applications.

1.2 Contributions and Organization

Against the background mentioned in the previous subsection, in chapter 2, we propose a design method of low-dimensional linear functional observers to estimate a given set of states via observer reduction approach where a low-dimensional observer approximates a given original observer for systems of interest. We first clarify that we have to take into account not only an initial state estimation error but also external input signals. Furthermore, analyzing the approximation error arising from the above two error factors by means of model reduction techniques, in particular, structured model reduction in [37], we clarify the relation between the approximation error and the $\mathcal{L}_2$-performance of the estimation error. Moreover, we devise a systematic low-dimensional observer construction algorithm satisfying a prescribed $\mathcal{L}_2$-performance of the estimation error.

This type of observer is useful for estimation of a limited number of states such as load power in a particular area of electric power network systems. However, for example in electric power network systems involving a number of consumers and various power plants, we are required to estimate the load power of overall power network systems (not in a particular area) with small computational costs. This is because the power network system becomes possibly unstable due to installation of a large amount of renewable energy resources. However, it is difficult to exactly estimate all of system states by using low-dimensional observers in general.

In view of this, in Chapter 3, we propose a novel framework of low-dimensional observers called *average state observer*. The observer estimates averaged states, which represents average behavior of systems from a macroscopic point of view, instead of estimating all of system states. To explain this idea, let us consider the following example: From a microscopic point of view, the behavior of fluid arises from complex interaction among a huge number of molecules. On the other hand, from a macroscopic point of view, we observe only a kind of average behavior of molecules. This fact implies that the estimation of average behavior is essential to capture the fluid behavior (i.e., large-scale network systems) from a macroscopic point of view. Since it is nontrivial to find a set of states capturing average behavior in general, we cannot determine a signal to be estimated in advance. To overcome this difficulty, we utilize the concept of clustered model

reduction developed in [38, 39]. Furthermore, deriving a tractable representation of the error system, we provide a design procedure for average state observers with systematic determination of a set of states capturing average behavior of systems. Moreover, we show a theoretical $\mathcal{L}_2$-error bound of estimation error by average state observers. The average state observer is suitable for large-scale network systems in the sense that it not only satisfies the requirements (i) and (ii) in the previous subsection, but also estimates nontrivial average behavior of systems from a macroscopic point of view.

In Chapter 4, we propose hierarchical distributed control for network systems. The hierarchical distributed control system consists of several layers in which subsystems and hierarchically clustered subsystems are controlled by distributed controllers. The proposed method enables us to construct each compositional distributed controller without taking into account the other distributed controllers and the overall network system. A notion of *distributed design* is introduced in [40], where a performance limitation of controllers that are designed in a distributed manner is discussed by confining the class of systems to handle. In addition, a distributed design method in terms of the $\mathcal{L}_1$-induced norm has been developed for positive linear systems [41]. However, since this method fully utilizes a specific property of positive systems, generalization to a broader class of systems is not straightforward. In contrast, our proposed method is applicable for general linear systems and has an advantage that an $\mathcal{L}_2$-performance of the closed-loop system improves as improving a performance of distributed controllers that stabilize disjoint subsystems individually. Towards systematic design, we utilize state-space expansion that enables us to deal with the state variables associated with disjoint subsystems and those associated with interference among hierarchically clustered subsystems in a tractable manner. Moreover, by the integration of a hierarchical distributed observer having good compatibility with the structured controller, we build a framework to implement an observer-based hierarchical distributed control.

The proposed hierarchical distributed observers/controllers satisfies the requirement (ii) in the previous subsection and has a property allowing us distributed design. However, the designed hierarchical distributed (observer-based) controller in each layer necessarily have the same dimension as that of the system to be controlled. Thus, the designed hierarchical distributed controllers do not fully comply with practical application for large-scale systems from a viewpoint of computational costs for implementation (i.e., requirement (i) is not satisfied).

In Chapter 5, we propose a design method of low-dimensional hierarchical distributed controllers for large-scale network systems via a controller reduction approach. Supposing that a hierarchical distributed controller is given by the method in the previous chapter, we find a low-dimensional hierarchical distributed controller approximating the

original one for any sets of locally stabilizing distributed controllers. Since the existing controller reduction techniques, e.g., [24, 25], cannot explicitly take into account the interconnection structure among controllers, this controller reduction problem cannot be solved by the straightforward use of those techniques. Thus, we explicitly utilize the hierarchical distributed structure of the closed-loop system. More specifically, taking into account the inherent hierarchy of information transmission which can be represented as the block-triangular structure of a coordinate transformed closed-loop system, we show that the approximation error of compositional controllers in upper layers does not affect those in lower layers. Next, using biorthogonal projection [17], we clarify the relation between the approximation error of compositional controllers and the performance degradation of the closed-loop system.

Finally, in Chapter 6, as a first step towards low-dimensional observers/controllers design for nonlinear large-scale network systems, we show the importance and necessity of nonlinear low-dimensional modelling through an example of a real industrial application. More specifically, we first construct a nonlinear thermal-diffusion network model of the plasticization cylinder, which is an important component in plasticization process, while taking into account the temperature-dependent properties of heaters (which makes the model nonlinear). However, the network system, which is spatially discretized model, becomes an inevitably high-dimensional nonlinear system. Thus, in the second half of this chapter, we reduce the dimension of nonlinear network systems. More specifically, utilizing particular structures of the nonlinear network system arising from radiation to the air of heaters, we show that the reduced nonlinear network system preserves the stability with an a priori approximation error bound. The nonlinear reduced order model is expected to be useful for quality management and improvement of plastic products.

1.3 Organization

In Chapter 2, we propose a design method of low-dimensional linear functional observers on the basis of model reduction techniques. This method can not only preserves stability of the low-dimensional observer but also guarantee an a priori estimation error bound. Moreover, owing to the indipendency of the original observer design from observer reduction, the method is compatible with the standard feedback gain determination methods, such as pole placement techniques. The efficiency of the proposed method is shown through a numerical example of electric power network systems.

In Chapter 3, we propose a novel framework of low-dimensional observers called *average state observer*, which estimates averaged states instead of estimating all of system states. First, we give a mathematical formulation of the average state observer. Furthermore, we

derive the error system clarifying that not only an initial state estimation error but also the initial value response of systems and the input signal are relevant to the observation error. On the basis of this error analysis, we devise a systematic design procedure for average state observers with determination of a set of system states capturing average behavior. The efficiency of the proposed average state observers is shown through a numerical example for a reaction-diffusion system evolving over a complex network.

In Chapter 4, we propose a design method of hierarchical distributed controllers for network systems. The hierarchical distributed controller has an advantage that an $\mathcal{L}_2$-performance of the closed-loop system is guaranteed for all sets of locally stabilizing distributed controllers. Towards systematic design, we first introduce state-space expansion to independently deal with the state variables associated with disjoint subsystems and those associated with the interference among hierarchically clustered subsystems. By the hierarchical distributed controller, whose compositional controllers can be designed individually, it is shown that an $\mathcal{L}_2$-performance of closed-loop systems improves as just improving an $\mathcal{L}_2$-performance of local controllers. Moreover, by the integration of a hierarchical distributed observer having good compatibility with the structured controller, we build a framework to implement an observer-based hierarchical distributed control. The efficiency of the proposed control system is shown through an example of power network systems.

In Chapter 5, we propose a design method of low-dimensional hierarchical distributed controllers for large-scale network systems. Towards systematic design, we solve a controller reduction problem where the low-dimensional controller approximates the given hierarchical distributed controller in the previous chapter for any sets of locally stabilizing distributed controllers while preserving the same hierarchical distributed structure as that of the original controller. Finally, we demonstrate the efficiency of the proposed method through a numerical example of power network systems.

In Chapter 6, as a first step towards low-dimensional observers/controllers design for nonlinear large-scale network systems, we construct a low-dimensional nonlinear model of the plasticization cylinder that has a spatially distributed nonlinear dynamics. First, we derive nonlinear distributed parameter models of the plasticization cylinder on the basis of the physical first principles. Next, we reduce the model complexity by utilizing the particular structures of the nonlinearity arising from temperature-dependency of radiation of heaters. Whereas the original nonlinear model is an 808-dimensional spatially discretized model, we obtain an 28-dimensional model while guaranteeing a practically satisfactory accuracy. Furthermore, we show the validity of the resultant model by experiment and numerical simulation.

Chapter 2

Low-Dimensional Functional Observer Design

2.1 Introduction

In this chapter, we propose a novel method of designing low-dimensional observer that satisfies a specified estimation error precision. The proposed method is based on observer reduction approach where a low-dimensional observer approximates an observer for the system of interest. We first clarify that we have to take into account not only an initial state estimation error but also external input signals in low-dimensional observer design problems in general. In view of this, we define an evaluation function as weighted sum of estimation errors with respect to the initial state error and the input signal. This evaluation function represents estimation performance degradation arising from approximation of the original observer. Analyzing this evaluation function based on model reduction techniques, in particular, structured model reduction in [37], we clarify the relation between the approximation error and the estimation performance. Furthermore, we derive an *a priori* $\mathcal{L}_2$-error bound on the performance degradation with a systematic design procedure. The proposed method has an advantage that existing observer design methods, e.g., pole placement, can be employed to design original observers because the original observer can be designed independently of its reduction. Finally, we show the efficiency of the proposed method through a numerical example of electric power network systems.

This chapter is organized as follows: In Section 2.2, we formulate a design problem of low-dimensional functional observers via an observer reduction approach. In Section 2.3, we devise a design method of low-dimensional observers by biorthogonal projection. Furthermore, we show an *a priori* $\mathcal{L}_2$-error bound on the performance degradation. In

Section 2.4, we show the efficiency of the proposed method through a numerical example of power network systems. Finally, Section 2.5 concludes this chapter.

2.2 Problem Formulation

2.2.1 Design Problem of Low-dimensional Observers

In this chapter, we consider n-dimensional linear systems described by

$$\Sigma : \begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du \\ z = Sx \end{cases} \tag{2.1}$$

where $x(0) = x_0 \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m_u}$, $C \in \mathbb{R}^{m_y\times n}$, $D \in \mathbb{R}^{m_y\times m_u}$, $S \in \mathbb{R}^{m_z\times n}$ and $y \in \mathbb{R}^{m_y}$ is an measurement output, $z \in \mathbb{R}^{m_z}$ is a signal to be estimated. We first describe the problem formulation and the main result for stable Σ. The extension to unstable systems is described in Remark 2.5 in Section 2.3.1. In addition, we assume that the observability matrix $[C^\mathsf{T}, (CA)^\mathsf{T}, \ldots, (CA^{n-1})^\mathsf{T}]^\mathsf{T}$ is of full rank.

For Σ in (2.1), we consider a minimal dimensional observer in [42] given by

$$\Sigma_\mathrm{o} : \begin{cases} \dot{\xi} = F\xi + Hu + Gy \\ z_\mathrm{o} = L\xi + My \end{cases} \tag{2.2}$$

where $n_\mathrm{o} := n - m_y$, $F \in \mathbb{R}^{n_\mathrm{o}\times n_\mathrm{o}}$, $H \in \mathbb{R}^{n_\mathrm{o}\times m_u}$, $G \in \mathbb{R}^{n_\mathrm{o}\times m_y}$, $L \in \mathbb{R}^{m_z\times n_\mathrm{o}}$ and $M \in \mathbb{R}^{m_z\times m_y}$. For simplicity, we assume that $\xi(0) = 0$.

Let us review properties of observers Σ_o. The signal $z(t)$ of Σ depends on an initial state x_0 and an input signal $u(t)$. In this sense, we describe the signal z as $z(t; x_0, u)$. Similarly to this, the measurement output y can be described as $y(t; x_0, u)$. Since $y(t; x_0, u)$ and $u(t)$ are applied into the observer Σ_o, the estimated signal z_o in (2.2) depends on x_0 and u, which implies that z_o can be described as $z_\mathrm{o}(t; x_0, u)$. It is known in [43] that the transfer function of Σ from u to z_o coincides with that of Σ_o from u to z. In other words, the estimation error of $z(t; x_0, u)$ by $z_\mathrm{o}(t; x_0, u)$ does not depend on u, i.e.,

$$e_z(t; x_0) := z(t; x_0, u) - z_\mathrm{o}(t; x_0, u) \tag{2.3}$$

and this estimation error $e_z(t; x_0)$ converges to zero if F, H, G, L and M are designed in a suitable sense. Thus, we can see that the observer Σ_o exactly cancels the influence of u on the estimation error e_z while Σ_o makes the estimation error small depending on the initial state of systems x_0.

Next, for Σ in (2.1), we consider $\hat{n}_{\mathrm{o}}$-dimensional observers described by

$$\hat{\Sigma}_{\mathrm{o}} : \begin{cases} \dot{\hat{\xi}} = \hat{F}\hat{\xi} + \hat{H}u + \hat{G}y \\ \hat{z}_{\mathrm{o}} = \hat{L}\hat{\xi} + \hat{M}y \end{cases} \tag{2.4}$$

where $\hat{\xi}(0) = 0$, $\hat{F} \in \mathbb{R}^{\hat{n}_{\mathrm{o}} \times \hat{n}_{\mathrm{o}}}$, $\hat{H} \in \mathbb{R}^{\hat{n}_{\mathrm{o}} \times m_u}$, $\hat{G} \in \mathbb{R}^{\hat{n}_{\mathrm{o}} \times m_y}$, $\hat{L} \in \mathbb{R}^{m_z \times \hat{n}_{\mathrm{o}}}$ and $\hat{M} \in \mathbb{R}^{m_z \times m_y}$. Without loss of generality, we assume that $\hat{n}_{\mathrm{o}} \leq n_{\mathrm{o}}$. In what follows, $\hat{\Sigma}_{\mathrm{o}}$ is called a low-dimensional functional observer because the estimated signal $\hat{z}_{\mathrm{o}}$ is a function of the state $\hat{\xi}$.

Similarly to the case of Σ_{o}, let us consider the estimation error by $\hat{\Sigma}_{\mathrm{o}}$. We can describe $\hat{z}_{\mathrm{o}}(t)$ as $\hat{z}_{\mathrm{o}}(t; x_0, u)$ because the estimated signal $\hat{z}_{\mathrm{o}}$ by the low-dimensional functional observer $\hat{\Sigma}_{\mathrm{o}}$ depends on x_0 and u. The transfer function of $\hat{\Sigma}_{\mathrm{o}}$ from u to $\hat{z}_{\mathrm{o}}$ differs from that of Σ from u to z in general. Thus, the estimation error $z - \hat{z}_{\mathrm{o}}$ depends not only on x_0 but also u, i.e.,

$$\hat{e}_z(t; x_0, u) := z(t; x_0, u) - \hat{z}_{\mathrm{o}}(t; x_0, u). \tag{2.5}$$

This fact implies that we should take into account not only x_0 but also u in low-dimensional observer design.

2.2.2 Problem Formulation via Observer Reduction

In this section, we formulate the problem to design a low-dimensional functional observer $\hat{\Sigma}_{\mathrm{o}}$ in (2.4) as a problem approximating an observer Σ_{o} in (2.2). More specifically, using biorthogonal projection [17], we give design parameters in $\hat{\Sigma}_{\mathrm{o}}$ by

$$\hat{F} = PFP^{\dagger}, \quad \hat{H} = PH, \quad \hat{G} = PG, \quad \hat{L} = LP^{\dagger}, \quad \hat{M} = M \tag{2.6}$$

where $P \in \mathbb{R}^{\hat{n}_{\mathrm{o}} \times n_{\mathrm{o}}}$ and $P^{\dagger} \in \mathbb{R}^{n_{\mathrm{o}} \times \hat{n}_{\mathrm{o}}}$ satisfying $PP^{\dagger} = I_{\hat{n}_{\mathrm{o}}}$, and F, H, G, L and M are given such that e_z in (2.3) converges with a desirable convergence rate.

We first define an approximation error of Σ_{o} by $\hat{\Sigma}_{\mathrm{o}}$ as

$$\Delta(t; x_0, u) := e_z(t; x_0) - \hat{e}_z(t; x_0, u) \tag{2.7}$$

Note that Δ depends on not only u but also the initial state of Σ x_0 because the transfer function of Σ_{o} from $y(t; x_0, u)$ depending on x_0 to z_{o} differs from that of $\hat{\Sigma}_{\mathrm{o}}$ from $y(t; x_0, u)$ to $\hat{z}_{\mathrm{o}}$ in general. Since the response of linear systems coincides with sum of the initial response and the input response, we define the approximation error arising from u as

$$\Delta_u(t) := \Delta(t; 0, u) \tag{2.8}$$

and that arising from x_0 as

$$\Delta_{x_0}(t) := \Delta(t; x_0, 0). \tag{2.9}$$

Without loss of generality, we assume that $\|x_0\| = 1$ because $\Delta(t; cx_0, 0) \equiv c\Delta_{x_0}(t)$ holds for any $c \in \mathbb{R}_+$. To take into account both error factors of Δ_u and Δ_{x_0}, let us consider an evaluation function as

$$J(\Delta_{x_0}, \Delta_u) := \sqrt{w_{x_0}\|\Delta_{x_0}(t)\|_{\mathcal{L}_2}^2 + w_u\|\Delta_u(t)\|_{\mathcal{L}_2}^2} \tag{2.10}$$

where $w_{x_0} \in \mathbb{R}_+$ and $w_u \in \mathbb{R}_+$ satisfying $w_{x_0} + w_u = 1$ are design parameters to tune the weight of approximation error factors Δ_{x_0} and Δ_u. For example, J with $w_{x_0} = 1$ (resp. $w_u = 1$) evaluates the approximation error arising from an initial state x_0 (resp. an input u). In this sense, the function J represents performance degradation to evaluate the approximation error of $\hat{z}_\mathrm{o}$ by z_o.

In this setting, let us formulate an observer reduction problem to design $\hat{\Sigma}_\mathrm{o}$ which makes J in (2.10) small. For simplicity, we formulate a problem for unit impulse input signals u, i.e., $u(t) = u_0\delta(t)$ for any $u_0 \in \mathbb{R}^{m_u}$ such that $\|u_0\| = 1$.

Problem 2.1. *For a given Σ in (2.1), give Σ_o in (2.2) such that $e_z(t; x_0)$ in (2.3) converges with a desirable convergence rate. Define J in (2.10) for given $w_{x_0} \in \mathbb{R}_+$ and $w_u \in \mathbb{R}_+$. Then, find $\hat{n}_\mathrm{o}$ and a low-dimensional functional observer $\hat{\Sigma}_\mathrm{o}$ in (2.4), (2.6) such that*

$$J(\Delta_{x_0}, \Delta_u) \leq \epsilon \tag{2.11}$$

for a given $\epsilon > 0$ and any $u(t) = u_0\delta(t)$ where $u_0 \in \mathbb{R}^{m_u}$ satisfying $\|u_0\| = 1$.

Remark 2.1. In the line of work [42, 43, 44, 45, 46, 47], full or partial state observers can be designed from the view point of *exactly canceling* the effect of external input signals with respect to state estimation error. However, it is difficult to design low-dimensional observers based on the above design methods in general because the state-space of observers must include states having even little influence on state estimation. In contrast, our approach introduce the notion of *approximation* to observer design. In other words, truncating approximately uncontrollable state-space, we consider constructing a further low-dimensional functional observer.

Note that Problem 2.1 turns out to be a problem to find biorthogonal projection matrices P and $P^\dagger$ in (2.6) satisfying (2.11) while determining $\hat{n}_\mathrm{o}$ on the assumption that F, H, G, L and M are given in a suitable sense.

2.3 Low-dimensional Functional Observer Design

2.3.1 Error Analysis of Low-dimensional Functional Observers

In this subsection, we investigate the relation between the performance degradation function $J(\Delta_{x_0}, \Delta_u)$ in (2.10) and the choice of the biorthogonal projection defined by P and $P^\dagger$. First, we show the following lemma that will be needed for an error analysis below.

Lemma 2.2. *For a given stable $\Sigma_{\rm o}$ in (2.2), let $V \succ 0_{n_{\rm o}}$ be given such that*

$$F^{\mathsf{T}}V + VF \prec 0_{n_{\rm o}}. \tag{2.12}$$

Then, there exists $\gamma \in (0, \infty)$ satisfying

$$\mathcal{S}_\gamma(F, L, V) \prec 0_{n+n_{\rm o}} \tag{2.13}$$

where

$$\mathcal{S}_\gamma(F, L, V) := \begin{bmatrix} F^{\mathsf{T}}V + VF + \gamma^{-1}L^{\mathsf{T}}L & VF \\ F^{\mathsf{T}}V & -\gamma V \end{bmatrix}. \tag{2.14}$$

Furthermore, let $P = WV_{\frac{1}{2}}$ and $P^\dagger = V_{\frac{1}{2}}^{-1}W^{\mathsf{T}}$ where $W \in \mathbb{R}^{\hat{n}_{\rm o} \times n_{\rm o}}$ such that $WW^{\mathsf{T}} = I_{\hat{n}_{\rm o}}$. Then, it follows that

$$\left\| LP^\dagger(sI_{\hat{n}_{\rm o}} - PFP^\dagger)^{-1}PFV_{\frac{1}{2}}^{-1} \right\|_{\mathcal{H}_\infty} < \gamma \tag{2.15}$$

for any W.

Proof. First, we prove the existence of $\gamma \in (0, \infty)$ satisfying (2.13). Note that (2.13) is equivalent to

$$F^{\mathsf{T}}V + VF + \gamma^{-1}(L^{\mathsf{T}}L + VFV) \prec 0_{n_{\rm o}}. \tag{2.16}$$

From (2.12), there exists $\beta > 0$ such that $F^{\mathsf{T}}V + VF \prec -\beta I_{n_{\rm o}}$. Hence, for any positive η

$$\gamma = \frac{1}{\beta}\left(\lambda_{\max}(L^{\mathsf{T}}L + VFV) + \eta\right)$$

satisfies (2.16) where $\lambda_{\max}(X)$ denotes the largest eigenvalue of semipositive matrix $X = X^{\mathsf{T}} \succeq 0_{n_{\rm o}}$. Thus, the existence of $\gamma > 0$ satisfying (2.13) is proven. Next, we show (2.15) by the Bounded Real Lemma in [48]. More specifically, on the basis of the structured model reduction techniques in [37], we show

$$\mathcal{F}_\gamma(PFP^\dagger, PFV_{\frac{1}{2}}^{-1}, LP^\dagger; I_{\hat{n}_{\rm o}}) \prec 0_{n+\hat{n}_{\rm o}} \tag{2.17}$$

holds with a storage function $V(\hat{\xi}) = \hat{\xi}^{\mathsf{T}}\hat{\xi}$ where

$$\mathcal{F}_\gamma(A, B, C; V) := \begin{bmatrix} A^{\mathsf{T}}V + VA + \gamma^{-1}C^{\mathsf{T}}C & VB \\ B^{\mathsf{T}}V & -\gamma I_{m_u} \end{bmatrix}.$$

Note that $\mathcal{S}_\gamma(F, L, V)$ can be written as

$$\tilde{V}^{\mathsf{T}}\mathcal{S}_\gamma(V_{\frac{1}{2}}FV_{\frac{1}{2}}^{-1}, LV_{\frac{1}{2}}^{-1}, I_{n_\mathrm{o}})\tilde{V}$$

where $\tilde{V} := \mathrm{dg}(V_{\frac{1}{2}}, V_{\frac{1}{2}})$. Hence, (2.13) is equivalent to

$$\mathcal{S}_\gamma(V_{\frac{1}{2}}FV_{\frac{1}{2}}^{-1}, LV_{\frac{1}{2}}^{-1}, I_{n_\mathrm{o}}) \prec 0_{n+n_\mathrm{o}}. \tag{2.18}$$

Furthermore

$$\mathcal{F}_\gamma(PFP^\dagger, PFV_{\frac{1}{2}}^{-1}, LP^\dagger; I_{\hat{n}_\mathrm{o}})$$

can be rewritten as

$$\tilde{W}\mathcal{S}_\gamma(V_{\frac{1}{2}}FV_{\frac{1}{2}}^{-1}, LV_{\frac{1}{2}}^{-1}, I_{n_\mathrm{o}})\tilde{W}^{\mathsf{T}}$$

where $\tilde{W} := \mathrm{dg}(W, I_n)$. Note that W is of full row rank. Thus, (2.18) yields (2.17). □

Lemma 2.2 shows that there always exist a positive-definite matrix V and bounded $\gamma > 0$. Furthermore, it follows for any W that $PFP^\dagger$ is stable and a projection-based reduced model admits the $\mathcal{H}_\infty$-bound shown in (2.15). This $\mathcal{H}_\infty$-bound plays an important role for the error analysis in the following theorem, which gives a solution to the observer reduction problem defined in Section 2.2.2.

Theorem 2.3. *Consider Problem 2.1. Let*

$$\mathcal{A} := \begin{bmatrix} F & GC \\ 0 & A \end{bmatrix}, \quad \mathcal{B} := \begin{bmatrix} GD + H \\ B \end{bmatrix} \tag{2.19}$$

and $\mathcal{K} \succeq 0_{n+n_\mathrm{o}}$ *be given such that*

$$\mathcal{A}\mathcal{K} + \mathcal{K}\mathcal{A}^{\mathsf{T}} + w_u\mathcal{B}\mathcal{B}^{\mathsf{T}} + w_{x_0}\mathrm{dg}(0, I_n) = 0. \tag{2.20}$$

Let $V \succ 0_{n_\mathrm{o}}$ *and* $\gamma > 0$ *such that*

$$\mathcal{S}_\gamma(F, L, V) \prec 0_{n+n_\mathrm{o}} \tag{2.21}$$

where $\mathcal{S}_\gamma$ in (2.14). Furthermore, for a given $\theta \in \mathbb{R}_+$, suppose that $W \in \mathbb{R}^{\hat{n}_\mathrm{o} \times n_\mathrm{o}}$ satisfies $WW^\mathsf{T} = I_{\hat{n}_\mathrm{o}}$ and

$$\mathrm{im}\left((LV_{\frac{1}{2}}^{-1})^\mathsf{T}\right) \subseteq \mathrm{im}(W^\mathsf{T}),\ \sqrt{\mathrm{tr}(\Phi) - \mathrm{tr}(W\Phi W^\mathsf{T})} \leq \theta \tag{2.22}$$

where

$$\Phi := V_{\frac{1}{2}} \mathcal{K}_{1:n_\mathrm{o}} V_{\frac{1}{2}}^\mathsf{T} \in \mathbb{R}^{n_\mathrm{o} \times n_\mathrm{o}} \tag{2.23}$$

and $\mathcal{K}_{1:n_\mathrm{o}} \in \mathbb{R}^{n_\mathrm{o} \times n_\mathrm{o}}$ denotes the principal submatrix of $\mathcal{K}$ corresponding to the first n rows and columns. Then, $\hat{\Sigma}_\mathrm{o}$ with $P = WV_{\frac{1}{2}}$ and $P^\dagger = V_{\frac{1}{2}}^{-1} W^\mathsf{T}$ in (2.6) satisfies

$$J(\Delta_{x_0}, \Delta_u) \leq \gamma\theta \tag{2.24}$$

for any unit impulse input u and $x_0 \in \mathbb{R}^n$ such that $\|x_0\| = 1$.

Proof. Define $\mathcal{S} := [-L,\ S - MC]$, $\mathcal{X}_0 := [0, x_0^\mathsf{T}]^\mathsf{T}$ and

$$\mathcal{P} := \mathrm{dg}(P, I_n),\quad \mathcal{P}^\dagger := \mathrm{dg}(P^\dagger, I_n).$$

Letting $\mathcal{X} := [\xi^\mathsf{T}, x^\mathsf{T}]^\mathsf{T}$, we have

$$\Sigma_{e_z} : \begin{cases} \dot{\mathcal{X}} = \mathcal{A}\mathcal{X} + \mathcal{B}u \\ e_z = \mathcal{S}\mathcal{X} - MDu \end{cases}$$

with $\mathcal{X}(0) = \mathcal{X}_0$. Similarly to this, letting $\hat{\mathcal{X}} := [\hat{\xi}^\mathsf{T}, x^\mathsf{T}]^\mathsf{T}$, we have

$$\hat{\Sigma}_{\hat{e}_z} : \begin{cases} \dot{\hat{\mathcal{X}}} = \mathcal{P}\mathcal{A}\mathcal{P}^\dagger\hat{\mathcal{X}} + \mathcal{P}\mathcal{B}u \\ \hat{e}_z = \mathcal{S}\mathcal{P}^\dagger\hat{\mathcal{X}} - MDu \end{cases}$$

with $\hat{\mathcal{X}}(0) = \mathcal{P}\mathcal{X}_0$. Consider the similarity transformation of the error system defined by Σ_{e_z} and $\hat{\Sigma}_{\hat{e}_z}$ with

$$T = \begin{bmatrix} -\mathcal{P} & I_{n+\hat{n}_\mathrm{o}} \\ I_{n+n_\mathrm{o}} & 0 \end{bmatrix},\quad T^{-1} = \begin{bmatrix} 0 & I_{n+n_\mathrm{o}} \\ I_{n+\hat{n}_\mathrm{o}} & \mathcal{P} \end{bmatrix}.$$

Then, we have

$$T\,\mathrm{dg}(\mathcal{A}, \mathcal{P}\mathcal{A}\mathcal{P}^\dagger)T^{-1} = \begin{bmatrix} \mathcal{P}\mathcal{A}\mathcal{P}^\dagger & -\mathcal{P}\mathcal{A}\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}} \\ 0 & \mathcal{A} \end{bmatrix},\ T\begin{bmatrix} \mathcal{B} \\ \mathcal{P}\mathcal{B} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{B} \end{bmatrix},\ T\begin{bmatrix} \mathcal{X}_0 \\ \mathcal{P}\mathcal{X}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{X}_0 \end{bmatrix}$$

and $[\mathcal{S},\ -\mathcal{S}\mathcal{P}^\dagger]T^{-1} = [-\mathcal{S}\mathcal{P}^\dagger,\ \mathcal{S}\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}}]$ where

$$\overline{\mathcal{P}} := \left[\overline{W}V_{\frac{1}{2}}, 0\right],\ \ \overline{\mathcal{P}}^\dagger := \left[\overline{W}V_{\frac{1}{2}}^{-\mathsf{T}}, 0\right]^\mathsf{T}$$

for $\overline{W} \in \mathbb{R}^{(n-\hat{n}_\mathrm{o})\times n}$ such that $[W^\mathsf{T}, \overline{W}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{n\times n}$ is unitary. Note that

$$\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}} = \mathcal{J}\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}}, \quad \mathcal{J} := \mathrm{dg}(I_{n_\mathrm{o}}, 0)$$

holds from the block structure of $\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}}$. Thus, it follows for any $u_0 \in \mathbb{R}^{m_u}$ satisfying $\|u_0\| = 1$ that

$$\|\Delta_u(t)\|_{\mathcal{L}_2} \le \|\Theta(s)\mathcal{J}\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}}(sI_{n+n_\mathrm{o}} - \mathcal{A})^{-1}\mathcal{B}\|_{\mathcal{H}_2}$$

where

$$\Theta(s) := \mathcal{S}\mathcal{P}^\dagger(sI_{n+\hat{n}_\mathrm{o}} - \mathcal{P}\mathcal{A}\mathcal{P}^\dagger)^{-1}\mathcal{P}\mathcal{A} + \mathcal{S}.$$

Since $LV_{\frac{1}{2}}^{-1}\overline{W}^\mathsf{T} = 0$ follows from (2.22), we have

$$\Theta(s)\mathcal{J}\overline{\mathcal{P}}^\dagger\overline{\mathcal{P}} = \left[-\theta(s)V_{\frac{1}{2}}^{-1}, 0\right]\mathrm{dg}(\overline{W}^\mathsf{T}\overline{W}V_{\frac{1}{2}}, 0)$$

where

$$\theta(s) := LP^\dagger(sI_{\hat{n}_\mathrm{o}} - PFP^\dagger)^{-1}PF.$$

Hence, it follows that

$$\|\Delta_u(t)\|_{\mathcal{L}_2}^2 \le \|\theta(s)V_{\frac{1}{2}}^{-1}\|_{\mathcal{H}_\infty}^2\ \|[\overline{W}V_{\frac{1}{2}}, 0](sI_{n+n_\mathrm{o}} - \mathcal{A})^{-1}\mathcal{B}\|_{\mathcal{H}_2}^2.$$

Furthermore, from simple calculation, we have

$$\mathcal{X}_0\mathcal{X}_0^\mathsf{T} \preceq \mathrm{dg}(0, I_n)$$

for any $x_0 \in \mathbb{R}^n$ satisfying $\|x_0\| = 1$. Thus, it follows that

$$\|\Delta_x(t)\|_{\mathcal{L}_2}^2 \le \|\theta(s)V_{\frac{1}{2}}^{-1}\|_{\mathcal{H}_\infty}^2\ \|[\overline{W}V_{\frac{1}{2}}, 0](sI_{n+n_\mathrm{o}} - \mathcal{A})^{-1}[0, I_n]^\mathsf{T}\|_{\mathcal{H}_2}^2.$$

Hence, we have

$$J(\Delta_{x_0}, \Delta_u) \le \gamma\sqrt{\mathrm{tr}(\overline{W}V_{\frac{1}{2}}(\mathcal{K}_{1:n}^{(1)} + \mathcal{K}_{1:n}^{(2)})V_{\frac{1}{2}}^\mathsf{T}\overline{W}^\mathsf{T})}$$

where $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ are the solutions of Lyapunov equations described as

$$\begin{aligned}\mathcal{A}\mathcal{K}^{(1)} + \mathcal{K}^{(1)}\mathcal{A}^\mathsf{T} + w_{x_0}\mathrm{dg}(0, I_n) &= 0\\ \mathcal{A}\mathcal{K}^{(2)} + \mathcal{K}^{(2)}\mathcal{A}^\mathsf{T} + w_u\mathcal{B}\mathcal{B}^\mathsf{T} &= 0\end{aligned}$$

respectively. Note that from Lyapunov theorem in [49] there always exist semipositive definite solutions $\mathcal{K}^{(1)}$ and $\mathcal{K}^{(2)}$ because $\mathcal{A}$ is stable. Furthermore, the uniqueness of the Lyapunov equation solution yields

$$\mathcal{K} = \mathcal{K}^{(1)} + \mathcal{K}^{(2)}.$$

In addition, it follows from (2.22) that

$$\sqrt{\mathrm{tr}(\overline{W}\Phi\overline{W}^{\mathsf{T}})} = \sqrt{\mathrm{tr}(\overline{W}V_{\frac{1}{2}}\mathcal{K}_{1:n_{\mathrm{o}}}V_{\frac{1}{2}}^{\mathsf{T}}\overline{W}^{\mathsf{T}})} \leq \theta.$$

Noting that $\|\theta(s)V_{\frac{1}{2}}^{-1}\|_{\mathcal{H}_\infty} < \gamma$ follows from Lemma 2.2, we have

$$J(\Delta_x, \Delta_u) \leq \gamma\sqrt{\mathrm{tr}(\overline{W}\Phi\overline{W}^{\mathsf{T}})} \leq \gamma\theta.$$

Hence, the claim follows. □

Theorem 2.3 provides an appropriate biorthogonal projection to approximate the minimal dimensional observer Σ_{o} in (2.2). This result is novel in the sense that, the performance degradation is evaluated as (2.24) while explicitly taking into account the dynamics of Σ in (2.1) and Σ_{o} in (2.2). Note that θ in (2.24) can be used as a design criterion to regulate the approximating quality of the resultant low-dimensional functional observer.

Remark 2.4. To find $\hat{n}_{\mathrm{o}}$ and $W \in \mathbb{R}^{\hat{n}_{\mathrm{o}}\times n_{\mathrm{o}}}$ such that $WW^{\mathsf{T}} = I_{\hat{n}_{\mathrm{o}}}$ and (2.22) for a given θ, we can use the following procedure: First, we find the set $\{(\lambda_i, v_i)\}_{i\in\{1,\ldots,n_{\mathrm{o}}\}}$ of all eigenpairs of Φ in (2.23). Note that we assume that $\lambda_i \geq \lambda_{i+1}$ and $\|v_i\| = 1$ without loss of generality. Next, we find the smallest $m \in \{1, \ldots, n_{\mathrm{o}}\}$ such that

$$\theta^2 \geq \lambda_{m+1} + \cdots + \lambda_{n_{\mathrm{o}}} \tag{2.25}$$

and construct $V_m = [v_1, \ldots, v_m] \in \mathbb{R}^{n_{\mathrm{o}}\times m}$. Finally, by the Gram-Schmidt process, we derive W such that

$$\mathrm{im}(W^{\mathsf{T}}) = \mathrm{im}([V_m, (LV_{\frac{1}{2}}^{-1})^{\mathsf{T}}]).$$

Then, the dimension of the low-dimensional functional observer turns out to be $\hat{n}_{\mathrm{o}} = \mathrm{rank}([V_m, (LV_{\frac{1}{2}}^{-1})^{\mathsf{T}}])$. Moreover, the resultant W satisfies $WW^{\mathsf{T}} = I_{\hat{n}_{\mathrm{o}}}$ and (2.22).

Remark 2.5. We can extend Theorem 2.3 for unstable Σ as follows: Let n_{s} be the number of stable poles of $\mathcal{A}$ and

$$U \in \mathbb{R}^{n_{\mathrm{s}}\times(n+n_{\mathrm{o}})}, \quad \overline{U} \in \mathbb{R}^{(n+n_{\mathrm{o}}-n_{\mathrm{s}})\times(n+n_{\mathrm{o}})} \tag{2.26}$$

be given such that $U\mathcal{A}U^{\mathsf{T}} \in \mathbb{R}^{n_{\rm s}\times n_{\rm s}}$ is stable and

$$U\mathcal{A}\overline{U}^{\mathsf{T}} = 0, \quad U^{\mathsf{T}}U + \overline{U}^{\mathsf{T}}\overline{U} = I_{n+n_{\rm o}}.$$

In additiion, let $\mathcal{K}_U \succeq 0_{n_{\rm s}}$ be given such that

$$\mathrm{sym}(U\mathcal{A}U^{\mathsf{T}}\mathcal{K}_U) + w_u U\mathcal{B}\mathcal{B}^{\mathsf{T}}U^{\mathsf{T}} + w_{x_0}U\mathrm{dg}(0, I_n)U^{\mathsf{T}} = 0$$

where $\mathrm{sym}(X) := X + X^{\mathsf{T}}$ and define $\mathcal{K}$ as $\mathcal{K} := U^{\mathsf{T}}\mathcal{K}_U U$. If W satisfies (2.22) and

$$\mathrm{im}([V_{\frac{1}{2}}, 0]\overline{U}^{\mathsf{T}}) \subseteq \mathrm{im}(W^{\mathsf{T}}) \tag{2.27}$$

then, the result same as in Theorem 2.3 is assured. This is proven as follows:

First, U satisfies

$$\mathcal{U}\mathcal{A}\mathcal{U}^{\mathsf{T}} = \begin{bmatrix} \overline{U}\mathcal{A}\overline{U}^{\mathsf{T}} & \overline{U}\mathcal{A}U^{\mathsf{T}} \\ 0 & U\mathcal{A}U^{\mathsf{T}} \end{bmatrix}, \quad \mathcal{U} := \begin{bmatrix} \overline{U} \\ U. \end{bmatrix}$$

Hence, taking the similarity transformation by $\mathcal{U}$, it follows from (2.27) that

$$\|\overline{W}\mathcal{V}(sI_{n+n_{\rm o}} - \mathcal{A})^{-1}\mathcal{B}\|_{\mathcal{H}_2} = \sqrt{\mathrm{tr}(\overline{W}\mathcal{V}\mathcal{K}^{(1)}\mathcal{V}^{\mathsf{T}}\overline{W}^{\mathsf{T}})}$$

where $\mathcal{V} := [V_{\frac{1}{2}}, 0] \in \mathbb{R}^{n_{\rm o}\times(n+n_{\rm o})}$ and $\mathcal{K}^{(1)}$ is given by $\mathcal{K}^{(1)} := U^{\mathsf{T}}\mathcal{K}_U^{(1)}U$ with $\mathcal{K}_U^{(1)} \succeq \mathcal{O}_{n_{\rm s}}$ satisfying

$$\mathrm{sym}(U\mathcal{A}U^{\mathsf{T}}\mathcal{K}_U^{(1)}) + w_u U\mathcal{B}\mathcal{B}^{\mathsf{T}}U^{\mathsf{T}} = 0.$$

Note that there always exists a unique semipositive definite solution $\mathcal{K}_U^{(1)}$ because $U\mathcal{A}U^{\mathsf{T}}$ is stable. Similarly to this, defiine $\mathcal{K}^{(2)}$ as $\mathcal{K}^{(2)} := U^{\mathsf{T}}\mathcal{K}_U^{(2)}U$ where $\mathcal{K}_U^{(2)}$ is the solution of

$$\mathrm{sym}(U\mathcal{A}U^{\mathsf{T}}\mathcal{K}_U^{(2)}) + w_{x_0}U\mathrm{dg}(0, I_n)U^{\mathsf{T}} = 0.$$

Then, $\mathcal{K} = \mathcal{K}^{(1)} + \mathcal{K}^{(2)}$ follows.

2.3.2 Design Algorithm of Low-dimensional Functional Observers

In this subsection, we propose a procedure to construct a low-dimensional functional observer $\hat{\Sigma}_{\rm o}$ in (2.4) being a solution of Problem 2.1. More specifically, a procedure to solve Problem 2.1 is summarized as follows:

(a) For Σ in (2.1), construct $\Sigma_{\rm o}$ in (2.2) such that $e_z \in \mathbb{R}^{m_z}$ in (2.3) converges with a desirable convergence rate.

(b) Find U and $\overline{U}$ such that (2.26) and $U\mathcal{A}U^{\mathsf{T}}$ is stable.

(c) Find $\gamma \in \mathbb{R}_+$ such that (2.21) by using V satisfying (2.12).

(d) For a given $\epsilon > 0$, let $\theta = \epsilon/\gamma$.

(e) For given $w_{x_0} \in \mathbb{R}_+$ and $w_u \in \mathbb{R}_+$ and θ, find $W \in \mathbb{R}^{\hat{n}_{\rm o} \times n_{\rm o}}$ satisfying $WW^{\mathsf{T}} = I_{\hat{n}_{\rm o}}$, (2.22) and (2.27). More specifically, giving V_m by the procedure shown in Remark 2.4, we construct W satisfying

$$\mathrm{im}(W^{\mathsf{T}}) = \mathrm{im}([V_m, (LV_{\frac{1}{2}}^{-1})^{\mathsf{T}}, [V_{\frac{1}{2}}, 0]\overline{U}^{\mathsf{T}}])$$

by the Gram-Schmidt process.

(f) Construct biorthogonal projection matrices $P = WV_{\frac{1}{2}}$ and $P^{\dagger} = V_{\frac{1}{2}}^{-1}W^{\mathsf{T}}$.

(g) Construct $\hat{\Sigma}_{\rm o}$ by (2.4) and (2.6).

The designed observer $\hat{\Sigma}_{\rm o}$, which is a further reduced observer than the original minimal dimensional observer $\Sigma_{\rm o}$, guarantees the estimation performance given by (2.10). The efficiency of this design procedure is demonstrated through a numerical example in the following section.

2.4 Numerical Simulation

2.4.1 Power Network Systems

In this section, we deal with a power network system in [10, 11, 50] composed of generators and loads. Let N be the number of generators. For $i \in \{1, \ldots, N\}$, we describe the dynamics of the i-th generator by

$$\Sigma_i^g : \begin{cases} \dot{\phi}_i = \tilde{A}_i\phi_i + \frac{1}{M_i}\tilde{b}p_i^g + \breve{b}u_i \\ \delta_i^g = \tilde{c}\phi_i \end{cases} \tag{2.28}$$

with

$$\tilde{A}_i := \begin{bmatrix} A_i^f & \frac{-1}{M_i}bc \\ k_i b b^{\mathsf{T}} & A_i^c \end{bmatrix}, \quad \tilde{b} := \begin{bmatrix} b \\ 0_{2\times 1} \end{bmatrix}, \quad \breve{b} := \tilde{c}^{\mathsf{T}}, \quad \tilde{c} := \begin{bmatrix} c & 0_{1\times 2} \end{bmatrix}$$

and

$$A_i^f := \begin{bmatrix} 0 & 1 \\ 0 & -\frac{D_i}{M_i} \end{bmatrix}, \quad A_i^c := \begin{bmatrix} -\frac{1}{T_i} & \frac{1}{T_i} \\ 0 & -k_i R_i \end{bmatrix}, \quad b := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad c := \begin{bmatrix} 1 & 0 \end{bmatrix}$$

where M_i, D_i, T_i, k_i and R_i are positive constants that represent inertia constant, damping coefficient, turbine time constant, governor time constant, and velocity tuning rate, respectively and $\phi_i \in \mathbb{R}^4$ denotes the state of a prime mover and a governor respectively, $p_i^g \in \mathbb{R}$ denotes the output electric power difference of a generator, $\delta_i^g \in \mathbb{R}$ denotes the phase angle difference of a generator, $u_i \in \mathbb{R}$ denotes the desired value of phase angle difference. The generator Σ_i^g in (2.28) has a structure that the governor described by (A_i^c, b, c) and the prime mover described by (A_i^f, b, c) are feedback interconnected.

Next, let L be the number of loads. In this thesis, we model loads as rotational spring-mass-damper systems in [10]. More specifically, for $i \in \{1, \ldots, L\}$, we give the dynamics of the ith load by

$$\Sigma_i^l : \begin{cases} \dot{\psi}_i = A_i^f \psi_i + \frac{1}{M_i} b p_i^l \\ \delta_i^l = c\psi_i \end{cases} \tag{2.29}$$

where $\psi_i \in \mathbb{R}^2$ denotes phase angle difference and angular velocity difference, $p_i^l \in \mathbb{R}$ denotes injected electric power difference, δ_i^l denotes phase angle difference. Furthermore, the interconnection among generators and loads is given by

$$p = -Y\delta \tag{2.30}$$

where

$$p := [p_1^g, \ldots, p_N^g, p_1^l, \ldots, p_L^l]^\mathsf{T}, \quad \delta := [\delta_1^g, \ldots, \delta_N^g, \delta_1^l, \ldots, \delta_L^l]^\mathsf{T}$$

and $Y \in \mathbb{R}^{(N+L)\times(N+L)}$ denotes an admittance matrix, which is a graph Laplacian matrix in [15] because the injected electric power p depends on the difference of phase angle δ among interconnected generators and loads. Define a state variable as

$$x := [\phi_1^\mathsf{T}, \ldots, \phi_N^\mathsf{T}, \psi_1^\mathsf{T}, \ldots, \psi_L^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^n \tag{2.31}$$

where $n := 4N + 2L$. In addition, we take an measurement output y and z in (2.1) as the phase angle and the angular velocity of all generators, and the angle of the first to the L_zth loads, respectively, i.e.,

$$\begin{aligned} y &:= \mathrm{dg}([I_2, 0_{2\times2}], \ldots, [I_2, 0_{2\times2}])[\phi_1^\mathsf{T}, \ldots, \phi_N^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{2N} \\ z &:= \mathrm{dg}(c, \ldots, c)[\psi_1^\mathsf{T}, \ldots, \psi_{L_z}^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^{L_z} \end{aligned}$$

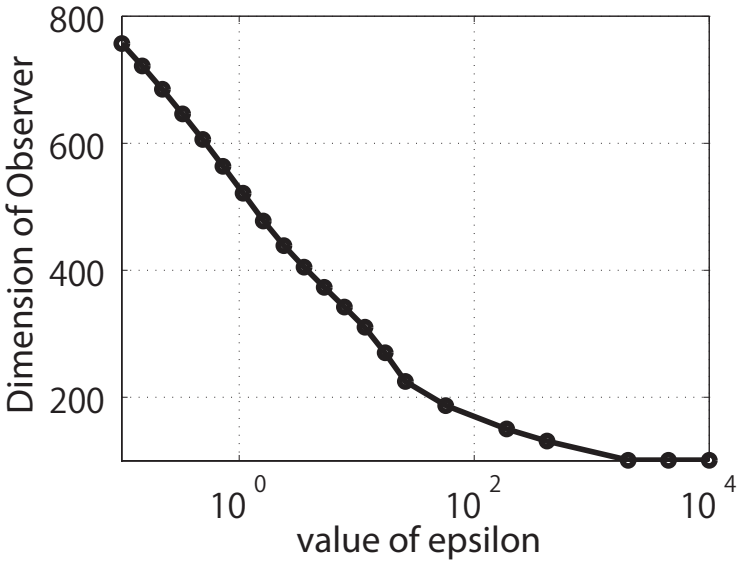


FIGURE 2.1: Order of resulting models versus values of ϵ.

Finally, Σ in (2.1) can be described as follows:

$$\begin{aligned}
A &= \mathrm{dg}(\tilde{A}_1, \ldots, \tilde{A}_N, A_1^f, \ldots A_L^f) - \tilde{B} Y \tilde{C} \\
\tilde{B} &= \mathrm{dg}\Big(\tfrac{1}{M_1}, \ldots, \tfrac{1}{M_N}, \tfrac{1}{M_1}, \ldots, \tfrac{1}{M_L}\Big) \mathrm{dg}(I_N \otimes \tilde{b}, I_L \otimes b), \quad \tilde{C} = \mathrm{dg}(I_N \otimes \tilde{c}, I_L \otimes c) \\
B &= \begin{bmatrix} \mathbf{1}_N \otimes \breve{b} \\ 0_{2L \times 1} \end{bmatrix}, \quad C = [\mathrm{dg}([I_2, 0_{2\times 2}], \ldots, [I_2, 0_{2\times 2}])\ 0_{2N \times 2L}], \quad D = 0_{2N \times 1}, \\
S &= \begin{bmatrix} 0_{L_z \times 4N}\ \mathrm{dg}(c, \ldots, c)\ 0_{L_z \times 2(L - L_z)} \end{bmatrix}
\end{aligned}$$

It should be noted that the dynamics of this power network system is invariant with respect to the bias of electric power angles, i.e.,

$$Av = 0, \quad v := [\mathbf{1}_N^\mathsf{T} \otimes \tilde{c} \ \ \mathbf{1}_L^\mathsf{T} \otimes c]^\mathsf{T} \in \mathbb{R}^n \tag{2.32}$$

holds because Y in (2.30) is a graph Laplacian matrix. Thus, at least one eigenvalue of A is on the imaginary axis, which yields that Σ is semi-stable.

2.4.2 Low-dimensional Functional Observer Design for Power Network Systems

In this subsection, we demonstrate the efficiency of the proposed method to construct low-dimensional functional observers. In what follows, we deal with an electric power network system composed of $N = 50$ generators and $L = 400$ loads, which yields $n = 1000$. In addition, we take a signal to be estimated $z \in \mathbb{R}^{100}$ as the phase angles of the first to $L_z = 100$ loads. Let parameters of Σ are given as follows: For all i,

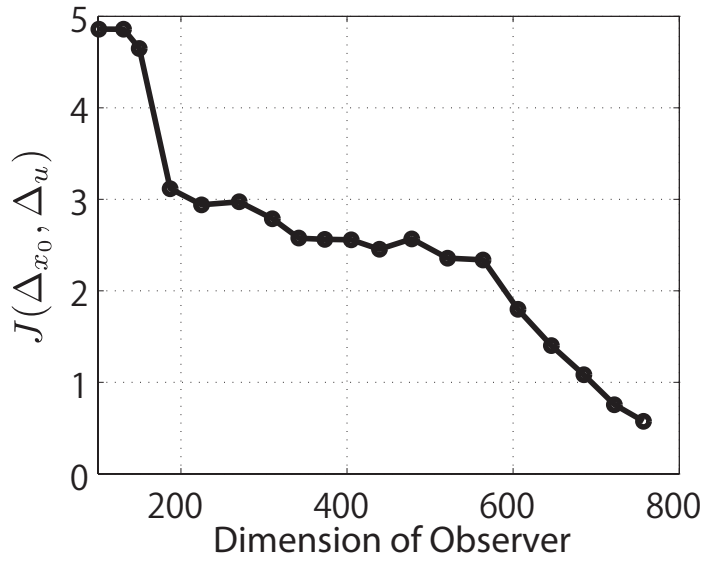


FIGURE 2.2: Performance function versus $\hat{n}_{\mathrm{o}}$.

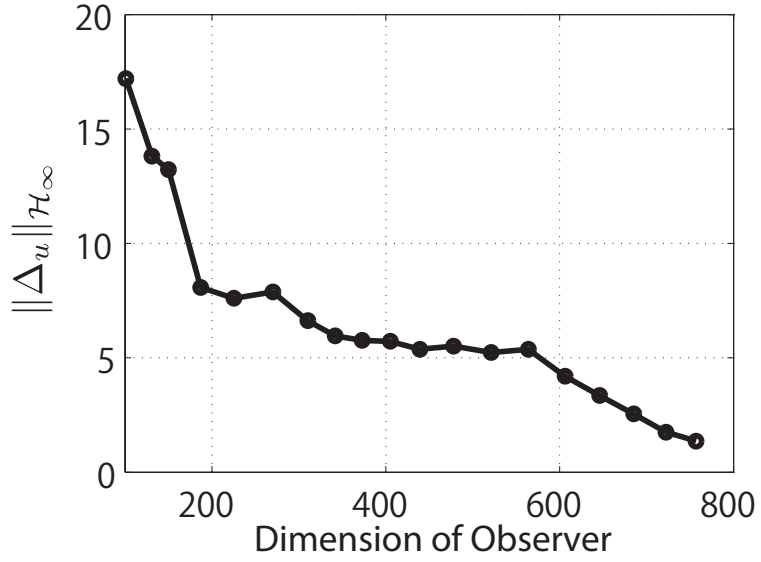


FIGURE 2.3: $\|\Delta_u(s)\|_{\mathcal{H}_\infty}$ versus $\hat{n}_{\mathrm{o}}$.

we take $k_i = 0.01$, $R_i = 0.05$, $D_i = 1.5$, $T_i = 0.2$ and $M_i = 10$. In addition, the admittance matrix Y in (2.30) is given as the graph Laplacian of a complex network model called Holme-Kim in [51] and nonzero values in off-diagonal elements of Y are randomly chosen from $[0, 1]$. Then, A in (2.1) has one eigenvalue on the imaginary axis and the corresponding eigenvector is denoted by v in (2.32). The other eigenvalues of A are located in the left half open complex plane.

Taking design parameters w_{x_0} and w_u in (2.10) as $w_{x_0} = 0.1, w_u = 0.9$, we first design a set of low-dimensional functional observers $\hat{\Sigma}_{\mathrm{o}}$ in (2.4) by the procedure shown in Section 2.3.2 for several given values of ϵ in (2.11). More specifically, let $\overline{U}$ in (2.26) be

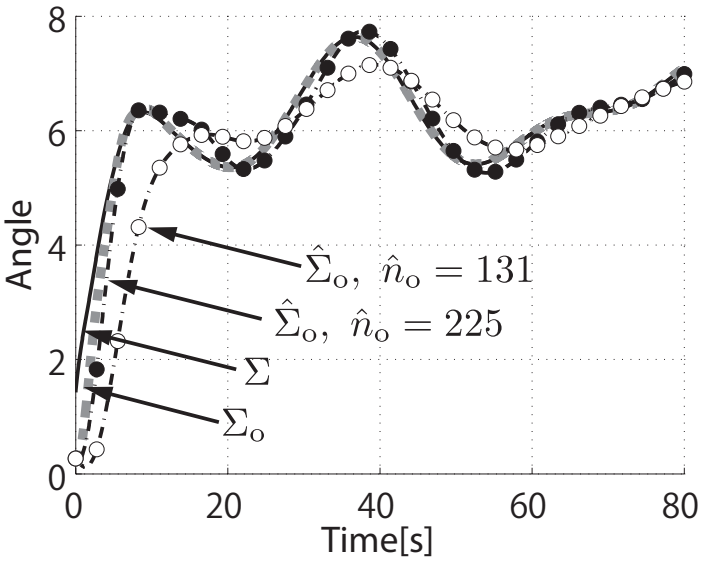


FIGURE 2.4: Responses of angle of the first load.

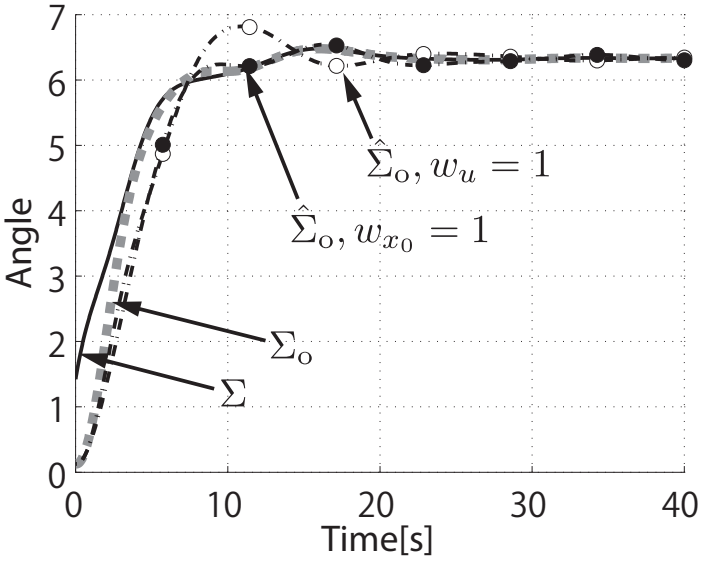


FIGURE 2.5: Initial value responses of angle of the first load.

v in (2.32). In the step (c), we find V by solving (2.12) and find $\gamma > 0$ satisfying (2.21). Next, design $\hat{\Sigma}_{\rm o}$ along with steps (d)-(g) for several given values of ϵ.

In FIGURE. 2.1, we plot the resultant dimension of low-dimensional functional observers $\hat{n}_{\rm o}$ versus the values of ϵ. From this figure, we can see that $\hat{n}_{\rm o}$ is decreasing as ϵ is increasing. This is because θ in (2.25) is given by $\theta = \epsilon/\gamma$ with a constant value of γ depending on $\Sigma_{\rm o}$. Thus, the smallest m satisfying (2.25) decreases as ϵ increases.

Furthermore, we plot in FIGURE. 2.2 the resultant performance degradation $J(\Delta_{x_0}, \Delta_u)$ in (2.10) with respect to $\hat{n}_{\rm o}$. From this figure, we can see that the estimation performance improves as taking a smaller value of ϵ. These results show that there is a tradeoff

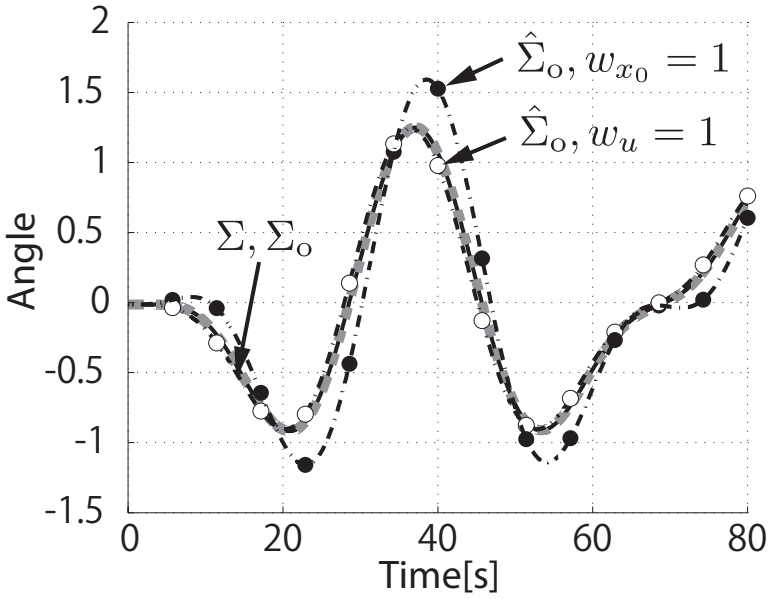


FIGURE 2.6: Input responses of angle of the first load.

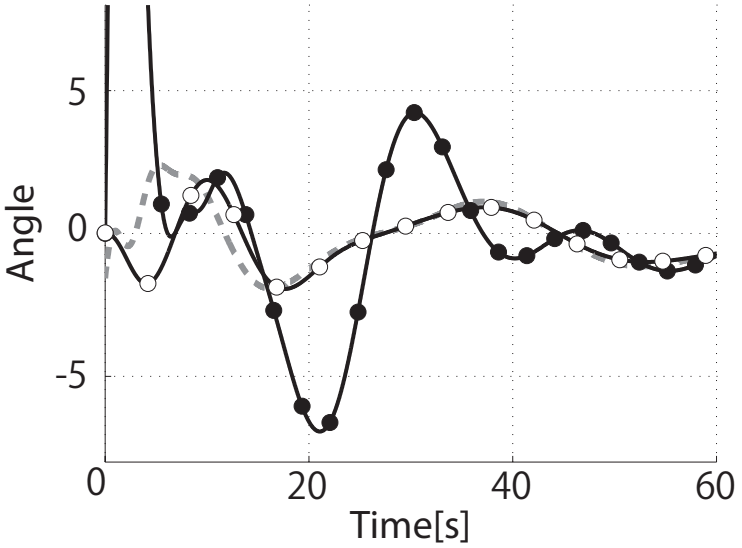


FIGURE 2.7: Comparison of the first element of $z, \hat{z}_{\rm o}$ with that of response of designed observer by existing approach.

between the low-dimensionality and performances. Furthermore, the design parameter ϵ can regulate the tradeoff.

In the previous section, for unit impulse input u, we have analyzed the approximation error and devised a design procedure based on the error analysis. Next, we examine the resultant low-dimensional functional observer makes the approximation error small for other input signals $u(t)$. In FIGURE. 2.3, we plot $\|\Delta_u(s)\|_{\mathcal{H}_\infty}$, which is the $\mathcal{H}_\infty$-norm of the transfer function of $\Delta(t)$ from u, with respect to $\hat{n}_{\rm o}$. FIGURE. 2.3 shows that $\|\Delta_u(s)\|_{\mathcal{H}_\infty}$ is decreasing as the dimension $\hat{n}_{\rm o}$ is increasing. This result implies that the proposed design method is effective for any square-integrable input signals u.

Furthermore, we compare the low-dimensional functional observer $\hat{\Sigma}_{\mathrm{o}}$ with the minimal dimensional observer Σ_{o} by comparing the resultant trajectories of $\hat{z}_{\mathrm{o}}$ and z_{o}. Note that the dimension of Σ_{o} is $\hat{n}_{\mathrm{o}} = 900$. We plot the first element of z, z_{o} and $\hat{z}_{\mathrm{o}}$ in (2.1), (2.2) and (2.4) in FIGURE. 2.4 where we give $x_0 \in \mathbb{R}^{1000}$ randomly and $u(t) = 0.02 + \sin(0.1t)\cos(0.1t)\sin(0.05t)$ containing multiple frequencies. We omit the other elements because they behave similarly to the first elements. We can see from FIGURE. 2.4 that the trajectories of $\hat{z}_{\mathrm{o}}(t)$ depicted as the dotted lines with circles are getting closer to $z_{\mathrm{o}}(t)$ of Σ_{o} as the dimension of $\hat{\Sigma}_{\mathrm{o}}$ is increasing.

Furthermore, taking $\hat{n}_{\mathrm{o}} = 190$, we show a result for several values of w_{x_0} and w_u in $J(\Delta_{x0}, \Delta_u)$ in (2.10). In FIGURE. 2.5 (resp. FIGURE. 2.6), we plot initial responses and (resp. input responses) of the first element of z, z_{o} and $\hat{z}_{\mathrm{o}}$ in (2.1), (2.2) and (2.4). We take x_0 and u as the same above. From these figures, the estimation performance of initial responses (resp. input responses) improves when the weight for initial states w_{x_0} (resp. that for inputs w_u) gets larger. This result implies that we can tune the estimation performance by regulating weight w_{x_0} and w_u in (2.10) without changing the dimension of low-dimensional functional observers.

Finally, we compare the proposed method with an existing method in [19]. Let the dimension of low-dimensional functional observer by proposed method be $\hat{n}_{\mathrm{o}} = 190$. The design procedure in [19] is summarized as follows: First, we construct a reduced order model approximating the transfer function of Σ in (2.1) from u to y by the balanced truncation [17]. Then, the resultant model is 190-dimensional system and the $\mathcal{H}_\infty$-norm of the model reduction error results in 1×10^{-14}. Next, we construct a Luenberger-type observer for this approximant with an estimated signal compatible with z.

In FIGURE. 2.7, the glay dotted line, the black solid line with white circles and that with black circles depict the first element of $z, \hat{z}_{\mathrm{o}}$ and the estimated signal of the low-dimensional observer designed by the above method, respectively. We can see from this figure that estimated signal by the low-dimensional observer designed by the existing method is fluctuated even though the model reduction error is sufficiently small. This is because the low-dimensional observer designed by the existing method does not take into account the influence of x_0 on estimated signals. This result implies that as mentioned in Section 2.2, we have to take into account both error factors arising from u and x_0 in low-dimensional observer design problems.

2.5 Chapter Summary

In this chapter, we have proposed a method of designing low-dimensional functional observers that satisfy a prescribed estimation performance. The proposed method is based on an observer reduction approach where a low-dimensional observer approximates an observer for the system of interest. We have clarified that we have to take into account not only an initial state estimation error but also external input signals. In view of this, we define an evaluation function which deals with both error factors arising from an initial state estimation error and external input signals. Analyzing this evaluation function based on model reduction techniques, we have derived an a priori $\mathcal{L}_2$-error bound on the evaluation function with a systematic design procedure. The proposed method has an advantage that existing observer design methods can be employed to design original observers because original observer can be designed independently of its reduction. Finally, we have shown the efficiency of the proposed method through an numerical example of electric power network systems.

Chapter 3

Average State Observers

3.1 Introduction

In this chapter, we propose a novel framework of low-dimensional observers called *average state observer*, which estimates averaged states (i.e., average behavior of the target network system from a macroscopic point of view) instead of estimating all of the system states. Since it is nontrivial to find a set of states capturing average behavior in general, we cannot determine the estimation signal capturing average behavior of systems in advance. To overcome this difficulty, we utilize the clustered model reduction technique developed in [38, 39]. Furthermore, deriving a tractable representation of the error system, we provide a design procedure for average state observers with systematic determination of a set of states capturing average behavior of the system. Moreover, we show a theoretical $\mathcal{L}_2$-error bound of the estimation error by an average state observer. The efficiency of the proposed average state observer is shown through a numerical example for a reaction-diffusion system evolving over a directed complex network.

The organization of this chapter is as follows: In Section 3.2, we first formulate average state observers. Furthermore, deriving a tractable representation of the error system, we clarify differences between the design of classical linear functional state observers and that of average state observers. In Section 3.3.1, we show a road map for systematic design of average state observers. Then, in Section 3.3.3, we devise a design procedure of average state observers that can estimate average behavior of the system from a macroscopic point of view. In Section 3.4, we show the efficiency of the proposed methods through a numerical example of a reaction-diffusion system on a directed network. Finally, concluding remarks are described in Section 3.5.

3.2 Fundamentals of Average State Observers

3.2.1 Review of Functional State Observers

In this chapter, we deal with the n-dimensional linear system described by

$$\Sigma : \begin{cases} \dot{x} = Ax + Bu, \\ y = Cx + Du, \end{cases} \qquad x(0) = x_0, \tag{3.1}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m_u}$, $C \in \mathbb{R}^{m_y\times n}$, and $D \in \mathbb{R}^{m_y\times m_u}$. To simplify the arguments, we first show results for stable systems Σ. The extension to unstable Σ is provided in Section 3.3.2. In Σ described by (3.1), $y \in \mathbb{R}^{m_y}$ denotes a measurement output signal. Furthermore, we give a signal to be estimated by

$$z = Sx \tag{3.2}$$

where $S \in \mathbb{R}^{m_z\times n}$. In general, we do not know a set of states capturing average behavior of Σ in advance. In view of this, we suppose that S in (3.2) is not given in advance.

In this notation, let us consider the n-dimensional observer, which is a Luenberger-type observer, described by

$$O : \begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + H(y - \hat{y}), \\ \hat{y} = C\hat{x} + Du, \end{cases} \qquad \hat{x}(0) = \hat{x}_0, \tag{3.3}$$

where the observer gain $H \in \mathbb{R}^{n\times m_y}$ is a design parameter. Similarly to (3.2), we define the estimation signal of z by

$$\hat{z} = S\hat{x}. \tag{3.4}$$

In what follows, O is referred to as a functional state observer [43, 47] because z and $\hat{z}$ are defined as functions of x and $\hat{x}$.

Define the state error by $e := x - \hat{x}$. It is well-known that the error system with this functional state observer can be described by

$$\mathcal{E} : \begin{cases} \dot{e} = (A - HC)e, \\ \Delta = Se, \end{cases} \qquad e(0) = e_0 \tag{3.5}$$

where

$$\Delta := z - \hat{z}, \quad e_0 := x_0 - \hat{x}_0$$

denote the estimation error and the initial state error, respectively. Note that the estimation error turns out to be a function of e_0, i.e.,

$$\Delta = \Delta(t; e_0). \tag{3.6}$$

Usually, we design the observer gain H such that this error system has a desirable behavior. To regulate a convergence rate of estimation errors, for a given constant $\delta \in \mathbb{R}_+$, one can design H such that

$$\sup_{e_0 \neq 0} \frac{\|\Delta(t; e_0)\|_{\mathcal{L}_2}}{\|e_0\|} \leq \delta. \tag{3.7}$$

In the next subsection, we define a lower-dimensional functional observer as a generalization of this n-dimensional functional state observer.

3.2.2 Error Analysis of Average State Observers

In this subsection, for Σ in (3.1), we define an $\hat{n}$-dimensional functional state observer as a generalized one of O in (3.3). More specifically, on the basis of the notion of orthogonal projection [17], we give an $\hat{n}$-dimensional functional state observer by

$$O_P : \begin{cases} \dot{\hat{x}} = PAP^{\mathsf{T}}\hat{x} + PBu + H(y - \hat{y}), \\ \hat{y} = CP^{\mathsf{T}}\hat{x} + Du, \end{cases} \quad \hat{x}(0) = \hat{x}_0 \tag{3.8}$$

where the observer gain $H \in \mathbb{R}^{\hat{n} \times m_y}$ and $P \in \mathbb{R}^{\hat{n} \times n}$ satisfying $PP^{\mathsf{T}} = I_{\hat{n}}$ are design parameters. We suppose $\hat{n} \leq n$ without loss of generality. Similarly to (3.4), we define the estimation signal of z in (3.2) by

$$\hat{z} = SP^{\mathsf{T}}\hat{x}. \tag{3.9}$$

In this chapter, we call this functional state observer O_P as an *average state observer*. Moreover, we define the estimation error of z in (3.2) by $\hat{z}$ as

$$\Delta_P := z - \hat{z}. \tag{3.10}$$

To analyze this estimation error for the average state observer, we derive a tractable representation of the error system as follows:

Theorem 3.1. *Let Σ in (3.1) be given. Consider z in (3.2), O_P in (3.8) and $\hat{z}$ in (3.9). Then, Δ_P defined by (3.10) obeys*

$$\mathcal{E}_P : \begin{cases} \dot{\xi} &= \mathcal{A}\xi + \mathcal{B}u, \\ \Delta_P &= \mathcal{S}\xi, \end{cases} \qquad \xi(0) = \begin{bmatrix} e_0 \\ x_0 \end{bmatrix} \tag{3.11}$$

where

$$e_0 := Px_0 - \hat{x}_0 \tag{3.12}$$

and

$$\mathcal{A} := \begin{bmatrix} PAP^\mathsf{T} - HCP^\mathsf{T} & (PA - HC)(I_n - P^\mathsf{T}P) \\ 0 & A \end{bmatrix}$$
$$\mathcal{B} := \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \mathcal{S} := \begin{bmatrix} SP^\mathsf{T} & S(I_n - P^\mathsf{T}P) \end{bmatrix}.$$

Proof. Taking a state as $\hat{\mathcal{X}} := [\hat{x}^\mathsf{T}\ x^\mathsf{T}]^\mathsf{T}$, we have

$$\begin{cases} \dot{\hat{\mathcal{X}}} &= \hat{\mathcal{A}}\hat{\mathcal{X}} + \hat{\mathcal{B}}u \\ \Delta_P &= \hat{\mathcal{S}}\hat{\mathcal{X}} \end{cases}, \ \hat{\mathcal{X}}(0) = \begin{bmatrix} \hat{x}_0 \\ x_0 \end{bmatrix}$$

where

$$\hat{\mathcal{A}} := \begin{bmatrix} PAP^\mathsf{T} - HCP^\mathsf{T} & HC \\ 0 & A \end{bmatrix}, \quad \hat{\mathcal{B}} := \begin{bmatrix} PB \\ B \end{bmatrix}, \quad \hat{\mathcal{S}} := [-SP^\mathsf{T}\ S].$$

Define

$$T := \begin{bmatrix} -I_{\hat{n}} & P \\ & I_n \end{bmatrix} = T^{-1}.$$

From the similarity transformation of $T\hat{\mathcal{A}}T^{-1}, T\hat{\mathcal{B}}$ and $\hat{\mathcal{S}}T^{-1}$, the claim follows. □

In Theorem 3.1, we can see that $\mathcal{E}_P$ in (3.11) corresponds to a generalized representation of the error system $\mathcal{E}$ in (3.5). This is because, if $P = I_n$, we have

$$\mathcal{A} = \begin{bmatrix} A - HC & 0 \\ 0 & A \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \mathcal{S} = \begin{bmatrix} S & 0 \end{bmatrix}$$

and

$$\xi(0) = \begin{bmatrix} e_0 \\ x_0 \end{bmatrix}, \quad e_0 = x_0 - \hat{x}_0.$$

Moreover, we can see that Δ_P in (3.10) is a function of not only e_0 but also x_0 and u, i.e.,

$$\Delta_P = \Delta_P(t; e_0, x_0, u). \tag{3.13}$$

This is clearly contrasted with Δ in (3.6) for the traditional functional state observer. An intuitive explanation on these three error factors is as follows:

(i) the error arising from the initial state error e_0 in (3.12),

(ii) the error amplified by the initial state response of Σ, i.e., $y = Ce^{At}x_0$, and

(iii) the error amplified by the dynamical discrepancy of Σ and O_P with respect to the external input u.

We conclude that we have to take into account the error factors (i), (ii) and (iii) for the design of average state observers.

3.3 Design of Average State Observers

3.3.1 A Road Map for Systematic Design

In what follows, we aim to design P and H in (3.8) that suppress the estimation error due to (i), (ii) and (iii) as much as possible. Since the dynamics of the error system is linear, we can represent individual error factors as

$$\Delta_P(t; e_0, 0, 0), \quad \Delta_P(t; 0, x_0, 0), \quad \Delta_P(t; 0, 0, u).$$

For the first one, similarly to (3.7), we consider regulating the convergence rate of the initial state error by introducing the criterion of

$$\sup_{e_0 \neq 0} \frac{\|\Delta_P(t; e_0, 0, 0)\|_{\mathcal{L}_2}}{\|e_0\|} \leq \delta \tag{3.14}$$

with a given constant $\delta \in \mathbb{R}_+$.

To see the influence of x_0 and u on Δ_P more explicitly, supposing that $e_0 = 0$, we derive the Laplace domain representation of the estimation error arising from the second and third factors as

$$\hat{\Delta}_P(s; x_0, u) := \Xi_{P,H}(s) X_P(s; x_0, u) \tag{3.15}$$

where

$$\Xi_{P,H}(s) := C_\Xi (sI_{\hat{n}} - A_\Xi)^{-1} B_\Xi + D_\Xi \tag{3.16}$$

with

$$A_\Xi := PAP^\mathsf{T} - HCP^\mathsf{T}, \quad B_\Xi := (PA - HC)\overline{P}^\mathsf{T}, \quad C_\Xi := SP^\mathsf{T}, \quad D_\Xi := S\overline{P}^\mathsf{T}$$

and

$$X_P(s; x_0, u) := \overline{P}(sI_n - A)^{-1}[x_0 + Bu(s)] \tag{3.17}$$

with an orthogonal complement $\overline{P} \in \mathbb{R}^{(n-\hat{n})\times n}$ of $P \in \mathbb{R}^{\hat{n}\times n}$ such that

$$P^\mathsf{T}P + \overline{P}^\mathsf{T}\overline{P} = I_n. \tag{3.18}$$

From (3.15), the estimation error arising from the second and third error factors is expected to be small if the norms of $\Xi_{P,H}$ and X_P are sufficiently small.

However, simultaneous design of P and H is difficult because $\Xi_{P,H}$ involves the design parameters in a nonlinear fashion. To overcome this difficulty, we utilize the following facts:

- The parameter H appears in the system $\Xi_{P,H}$, but not in X_P.
- The system X_P involves the parameter $\overline{P}$ (or equivalently P), but not H.
- By a suitable choice of P, we can achieve $D_\Xi = 0$ in (3.16), which may directly decrease the norm of $\Xi_{P,H}$.

On the basis of these facts, we first find P that minimizes the norm of X_P while making $D_\Xi = 0$, and then find H that minimizes the norm of $\Xi_{P,H}$. Taking this road map for the average state observer design, we give the following theorem:

Theorem 3.2. *Let Σ in (3.1) be given. Let z in (3.2). For a constant $\alpha \in \mathbb{R}_+$, define $\Phi \succeq 0_n$ such that*

$$A\Phi + \Phi A^\mathsf{T} + BB^\mathsf{T} + \alpha I_n = 0. \tag{3.19}$$

Furthermore, take $P \in \mathbb{R}^{\hat{n}\times n}$ satisfying

$$\mathrm{im}(S^\mathsf{T}) \subseteq \mathrm{im}(P^\mathsf{T}), \quad PP^\mathsf{T} = I_{\hat{n}} \tag{3.20}$$

and

$$\mathrm{tr}(\Phi) - \mathrm{tr}(P\Phi P^\mathsf{T}) \leq \epsilon. \tag{3.21}$$

If there exist

$$\gamma > 0, \quad X \succ 0_{\hat{n}}, \quad Y \in \mathbb{R}^{\hat{n}\times m_y}$$

such that $X \prec \delta^2 I_{\hat{n}}$ *and*

$$\begin{bmatrix} \mathrm{sym}(XPAP^{\mathsf{T}} - YCP^{\mathsf{T}}) + PS^{\mathsf{T}}SP^{\mathsf{T}} & * \\ \overline{P}A^{\mathsf{T}}P^{\mathsf{T}}X - \overline{P}C^{\mathsf{T}}Y^{\mathsf{T}} & -\gamma I_{n-\hat{n}} \end{bmatrix} \prec 0_n \tag{3.22}$$

where $\mathrm{sym}(M) := M + M^{\mathsf{T}}$ *and* $\overline{P} \in \mathbb{R}^{(n-\hat{n})\times n}$ *satisfys (3.18), then* O_P *in (3.8) with*

$$H = X^{-1}Y \tag{3.23}$$

satisfies (3.14) for any $x_0 \in \mathbb{R}^n$ *and* $\hat{x}_0 \in \mathbb{R}^{\hat{n}}$*, and*

$$\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2 + \alpha\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2 \leq \gamma\epsilon \tag{3.24}$$

for any $\hat{x}(0)$*, unit impulse inputs* u *and any* $x_0 \in \mathbb{R}^n$ *such that* $\|x_0\| = 1$*, where* e_0 *and* Δ_P *are defined as in (3.12) and (3.13).*

Proof. First, we evaluate $\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2$ for any unit impulse input u, i.e., $u(t) = u_0\delta(t)$ for any $u_0 \in \mathbb{R}^{m_u}$. Based on the error system in (3.15), we have

$$\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2 \leq \|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty}^2 \|\overline{P}(sI-A)^{-1}B\|_{\mathcal{H}_2}^2$$

where $\Xi_{P,H}$ in (3.16). Substituting (3.23) into (3.22), we have

$$\begin{bmatrix} \mathrm{sym}(XPAP^{\mathsf{T}} - XHCP^{\mathsf{T}}) + PS^{\mathsf{T}}SP^{\mathsf{T}} & * \\ \overline{P}A^{\mathsf{T}}P^{\mathsf{T}}X - \overline{P}C^{\mathsf{T}}H^{\mathsf{T}}X & -\gamma I_{n-\hat{n}} \end{bmatrix} \prec 0_n. \tag{3.25}$$

It yields from the first condition in (3.20) that

$$D_\Xi = S\overline{P}^{\mathsf{T}} = 0.$$

Thus, from the bounded-real lemma [52], (3.25) yields

$$\|\Xi_{P,H}(s)\|_{\mathcal{H}_\infty} < \gamma.$$

Note that there always exists a unique $\Phi^{(1)} \succeq 0_n$ satisfying

$$A\Phi^{(1)} + \Phi^{(1)}A^{\mathsf{T}} + BB^{\mathsf{T}} = 0$$

for the stable system Σ. Utilizing

$$\|\overline{P}(sI-A)^{-1}B\|_{\mathcal{H}_2}^2 = \mathrm{tr}(\overline{P}\Phi^{(1)}\overline{P}^{\mathsf{T}}),$$

we have

$$\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2 \le \gamma \mathrm{tr}(\overline{P}\Phi^{(1)}\overline{P}^{\mathsf{T}}).$$

Next, we evaluate $\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2$. Note that there always exists $\Phi^{(2)} \succeq 0_n$ satisfying

$$A\Phi^{(2)} + \Phi^{(2)}A^{\mathsf{T}} + I_n = 0.$$

Thus, using

$$x_0 x_0^{\mathsf{T}} \preceq I_n \tag{3.26}$$

for any $x_0 \in \mathbb{R}^n$ such that $\|x_0\| = 1$, we have

$$\|\overline{P}(sI - A)^{-1}x_0\|_{\mathcal{H}_2}^2 \le \mathrm{tr}(\overline{P}\Phi^{(2)}\overline{P}^{\mathsf{T}}).$$

Therefore

$$\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2 + \alpha\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2 \le \gamma \mathrm{tr}\left(\overline{P}(\Phi^{(1)} + \alpha\Phi^{(2)})\overline{P}^{\mathsf{T}}\right).$$

From the Lyapunov theorem [49], Φ given in (3.19) satisfies $\Phi = \Phi^{(1)} + \alpha\Phi^{(2)}$. In addition, the condition in (3.21) yields

$$\mathrm{tr}\left(\overline{P}(\Phi^{(1)} + \alpha\Phi^{(2)})\overline{P}^{\mathsf{T}}\right) \le \epsilon.$$

Thus, (3.24) follows. Finally, we show (3.14). To describe the time evolution of $\eta(t) := \Delta_P(t;e_0,0,0)$, we consider the system given by

$$\left\{ \begin{array}{l} \dot{e} = (PAP^{\mathsf{T}} - HCP^{\mathsf{T}})e \\ \eta = SP^{\mathsf{T}}e \end{array} \right. , \ e(0) = e_0.$$

From $(1,1)$ block of the left-hand equation in (3.22), this system has a Lyapunov function $V(e) := e^{\mathsf{T}}Xe$ such that

$$\dot{V}(e(t)) < -\eta^{\mathsf{T}}(t)\eta(t).$$

Integrating this inequality over $[0,\infty)$ and utilizing $V(e(\infty)) = 0$, we have

$$\|\eta(t)\|_{\mathcal{L}_2}^2 < V(e(0)) \le \|e_0\|^2\|X\|_2$$

for any $e_0 \in \mathbb{R}^{\hat{n}}$. Hence, (3.14) follows from $X \prec \delta^2 I_{\hat{n}}$. □

Remark 3.3. As shown in the proof of this theorem, we measure the effect of u in terms of the $\mathcal{H}_2$-norm, with similar results available for the case of the $\mathcal{H}_\infty$-norm. One possible approach to find P in the $\mathcal{H}_\infty$-norm evaluation is available on the basis of *Hessenberg transformation*; see [53] for details.

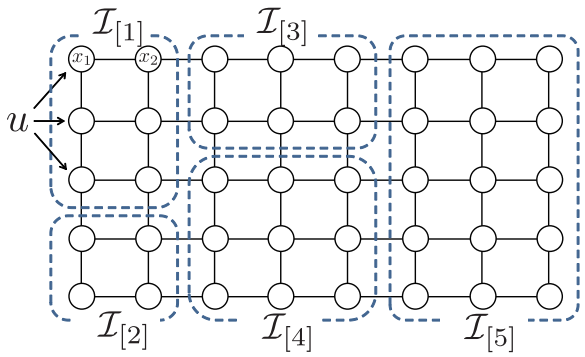


FIGURE 3.1: Illustrative example: Thermal diffusion system evolving on $2D$ lattice network.

In Theorem 3.2, we show an explicit error bound for average state observers. Next, to make the signal z represent averaged states, we introduce the following notion of clustering in [38, 39]:

Definition 3.4. The family of an index set $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ for $\mathbb{L} := \{1,\dots,L\}$ is called a *cluster set*, whose element is referred to as a cluster, if each element $\mathcal{I}_{[l]}$ is a disjoint subset of $\{1,\dots,n\}$ and it satisfies

$$\bigcup_{l\in\mathbb{L}} \mathcal{I}_{[l]} = \{1,\dots,n\}.$$

Then, an *aggregation matrix* compatible with $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ is defined by

$$P := \mathrm{dg}(p_{[1]},\dots,p_{[L]})\Pi \in \mathbb{R}^{L\times n} \tag{3.27}$$

with the permutation matrix

$$\Pi := [e^n_{\mathcal{I}_{[1]}},\dots,e^n_{\mathcal{I}_{[L]}}]^{\mathsf{T}} \in \mathbb{R}^{n\times n}, \quad e^n_{\mathcal{I}_{[l]}} \in \mathbb{R}^{n\times|\mathcal{I}_{[l]}|} \tag{3.28}$$

and $p_{[l]} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}$ such that $\|p_{[l]}\| = 1$.

Note that $\hat{n} = L$. Furthermore, we give S in (3.2) by

$$S = P. \tag{3.29}$$

In this setting, the condition in (3.20) is automatically satisfied. Furthermore, z and $\hat{z}$ in (3.2) and (3.9) are clearly given by

$$z = Px, \quad \hat{z} = \hat{x}.$$

Note that, if P has the specific structure shown in (3.27), then, the estimated signal $\hat{z}$ can be interpreted as a *weighted average* of states of Σ. In particular, if $p_{[l]}$ is in the

form of

$$p_{[l]} = \frac{[1, \ldots, 1]}{\|[1, \ldots, 1]\|} \in \mathbb{R}^{1 \times |\mathcal{I}_{[l]}|}, \tag{3.30}$$

then $\hat{z}$ corresponds to an average state in the sense that

$$z_l = \frac{1}{\sqrt{|\mathcal{I}_{[l]}|}} \sum_{i \in \mathcal{I}_{[l]}} x_i, \quad l \in \mathbb{L} \tag{3.31}$$

where z_l (resp. x_i) is the lth element of z (resp. the ith element of x). Owing to the block-diagonal structure of P, signal to be estimated z has physical (i.e., intuitive) meaning. For example, we consider Σ in (3.8) representing a thermal diffusion network system

$$\dot{x}_i = a_{i,i} x_i + \sum_{j \neq i} a_{i,j}(x_i - x_j) + b_i u, \quad i \in \{1, \ldots, n\} \tag{3.32}$$

evolving on $2D$ lattices shown in FIGURE. 3.1 where $a_{i,i}$ is a reaction diffusive coefficient and $a_{i,j}$ is a nonzero diffusion coefficient if the ith and jth nodes are connected. In addition, $x_i \in \mathbb{R}$ denotes the temperature of the ith node. Furthermore, suppose that a cluster set $\{\mathcal{I}_{[l]}\}_{l \in \mathbb{L}}$ is given as shown in FIGURE. 3.1. Then, a signal z_l in (3.31) represents an average temperature of the nodes belonging to $\mathcal{I}_{[l]}$.

A method to achieve usual averaging, i.e., normalized by $|\mathcal{I}_{[l]}|$ not by $\sqrt{|\mathcal{I}_{[l]}|}$, is described in Remark 3.6. In what follows, for simplicity, we only consider the case of (3.30). For systematic construction of a cluster set, we introduce the following lemma:

Lemma 3.5. *Let Σ in (3.1) be given, and define $\Phi \succeq 0_n$ such that (3.19) for a constant $\alpha \in \mathbb{R}_+$. Let $\theta \geq 0$ be given. For each $l \in \mathbb{L}$, if there exists a row vector $\phi_{[l]} \in \mathbb{R}^{1 \times n}$ such that*

$$\left\| (e^n_{\mathcal{I}_{[l]}})^{\mathsf{T}} \Phi_{\frac{1}{2}} - p_{[l]}^{\mathsf{T}} \phi_{[l]} \right\|_F \leq |\mathcal{I}_{[l]}|^{\frac{1}{2}} \theta \tag{3.33}$$

where $p_{[l]} \in \mathbb{R}^{1 \times |\mathcal{I}_{[l]}|}$ satisfying $\|p_{[l]}\| = 1$, then it follows that

$$\mathrm{tr}(\Phi) - \mathrm{tr}(P \Phi P^{\mathsf{T}}) \leq \theta^2 \left(\textstyle\sum_{l=1}^{L} |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1) \right) \tag{3.34}$$

where P is defined by (3.27).

Proof. See [39]. □

This lemma shows that, if we find a cluster set $\{\mathcal{I}_{[l]}\}_{l \in \mathbb{L}}$ satisfying (3.33), then ϵ in Theorem 3.2 can be taken as

$$\epsilon = \theta^2 \left(\textstyle\sum_{l=1}^{L} |\mathcal{I}_{[l]}|(|\mathcal{I}_{[l]}| - 1) \right).$$

Remark 3.6. The estimated signal z given by S and P in (3.29) and (3.27) with (3.30) implies an average state in the sense of (3.31). Instead of S in (3.29), if we take

$$S = DP, \quad D := \mathrm{dg}\left(\frac{1}{\sqrt{|\mathcal{I}_1|}}, \ldots, \frac{1}{\sqrt{|\mathcal{I}_L|}}\right) \in \mathbb{R}^{L\times L} \tag{3.35}$$

then z implies an average state in the sense of

$$z_l = \frac{1}{|\mathcal{I}_{[l]}|}\sum_{i\in\mathcal{I}_{[l]}} x_i, \quad l \in \mathbb{L}. \tag{3.36}$$

3.3.2 Extension to Unstable Systems

In this subsection, we consider designing an average state observer O_P in (3.8) for unstable Σ in (3.1). Suppose that A in (3.1) has n_{u} unstable eigenvalues. To estimate states diverging dependently on unstable modes of Σ, average state observers necessarily have the same unstable modes. Let

$$V \in \mathbb{R}^{n_{\mathrm{u}}\times n}, \quad \overline{V} \in \mathbb{R}^{(n-n_{\mathrm{u}})\times n}$$

be given such that all of n_{u} eigenvalues of VAV^{T} is unstable and

$$\overline{V}AV^{\mathsf{T}} = 0, \quad V^{\mathsf{T}}V + \overline{V}^{\mathsf{T}}\overline{V} = I_n. \tag{3.37}$$

Instead of P in (3.27), we take

$$P := \left[\mathrm{dg}\left(p_{[1]}, \ldots, p_{[L]}\right)\Pi,\ V^{\mathsf{T}}\right] \in \mathbb{R}^{\hat{n}\times n} \tag{3.38}$$

where $\hat{n} = L + n_{\mathrm{u}}$, $p_{[l]} \in \mathbb{R}^{1\times|\mathcal{I}_{[l]}|}$ in (3.30) and Π in (3.28). In this setting, we have the following lemma as an extented one of Lemma 3.5:

Lemma 3.7. *Let unstable Σ in (3.1) be given. Define V and $\overline{V}$ in (3.37). Let $\Phi_{\overline{V}} \succeq 0_{n-n_{\mathrm{u}}}$ be given such that*

$$(\overline{V}A\overline{V}^{\mathsf{T}})\Phi_{\overline{V}} + \Phi_{\overline{V}}(\overline{V}A\overline{V}^{\mathsf{T}})^{\mathsf{T}} + \overline{V}BB^{\mathsf{T}}\overline{V}^{\mathsf{T}} = 0 \tag{3.39}$$

and define $\Phi \in \mathbb{R}^{n\times n}$ such that $\Phi = \overline{V}^{\mathsf{T}}\Phi_{\overline{V}}\overline{V}$. Suppose that there exists a row vector $\phi_{[l]} \in \mathbb{R}^{1\times n}$ such that (3.33). Then, P in (3.38) satisfies (3.34).

Proof. It follows from (3.37) that

$$\mathcal{V}A\mathcal{V}^{\mathsf{T}} = \begin{bmatrix} VAV^{\mathsf{T}} & VA\overline{V}^{\mathsf{T}} \\ 0 & \overline{V}A\overline{V}^{\mathsf{T}} \end{bmatrix}, \quad \mathcal{V} := \begin{bmatrix} V \\ \overline{V} \end{bmatrix}. \tag{3.40}$$

Taking a similarity transformation by $\mathcal{V}$, we have

$$\|X_P(s)\|_{\mathcal{H}_2} = \|\overline{PV}^{\mathsf{T}}(sI_{n-n_{\mathrm{u}}} - \overline{V}A\overline{V}^{\mathsf{T}})\overline{V}B\|_{\mathcal{H}_2} = \|\overline{P}\Phi_{\overline{V}}\|_F$$

where $X_P(s)$ in (3.17) and $\overline{P} \in \mathbb{R}^{(n-\hat{n})\times n}$ is given satisfying (3.18). Note that the solution $\Phi_{\overline{V}}$ in (3.39) exists because $\overline{V}A\overline{V}^{\mathsf{T}}$ is stable. We omit the rest of this proof because it is the same as that of Lemma 3.5. □

This lemma allows us to construct P for unstable Σ. Furthermore, by taking

$$S = \mathrm{dg}\left(p_{[1]}, \ldots, p_{[L]}\right)\Pi,$$

a signal to be estimated $z = Sx$ results in average state defined as (3.31). Thus, O_P in (3.8) with H in (3.23) satisfies (3.24) for any $\hat{x}(0)$, unit impulse inputs u and any $x_0 \in \mathbb{R}^n$ such that $\|x_0\| = 1$.

3.3.3 Design Algorithm of Average State Observers

In this subsection, we propose a procedure to construct an average state observer that estimates average behavior of the system Σ.

First, we describe a cluster construction procedure. On the premise that $\theta \geq 0$ is given and $\Phi_{\frac{1}{2}}$ is obtained, we can find such a cluster set in the following manner: Suppose that a set of clusters $\{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}$ are already constructed and let

$$\mathcal{J} := \{1, \ldots, n\} \backslash \{\mathcal{I}_{[1]}, \ldots, \mathcal{I}_{[l]}\}.$$

Next, we consider constructing a new cluster $\mathcal{I}_{[l+1]}$. We first choose an index $j \in \mathcal{J}$, and take $\mathcal{I}_{[l+1]}$ such that

$$\mathcal{I}_{[l+1]} = \{i \in \mathcal{J} \backslash \{j\} \mid \|\phi_i - \phi_j\| \leq \theta\} \tag{3.41}$$

where $\phi_i \in \mathbb{R}^{1\times n}$ denotes the ith row vector of $\Phi_{\frac{1}{2}}$. Then, we can straightforwardly verify that this newly constructed cluster satisfies (3.33).

On the basis of this procedure, for a given design parameter $\rho \in \mathbb{R}_+$, we summarize a systematic design procedure of $\hat{n}$-dimensional average state observer O_P in (3.8) such

that (3.14) and

$$\|\Delta_P(t;0,0,u)\|_{\mathcal{L}_2}^2 + \alpha\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2 \leq \rho \tag{3.42}$$

as follows:

1. Find V and $\overline{V}$ such that (3.37). These matrices can be computed by the Real-Schur Decomposition.
2. Give $\theta \in \mathbb{R}_+$.
3. Find $\{\mathcal{I}_{[l]}\}_{l\in\mathbb{L}}$ and $\phi_{[l]} \in \mathbb{R}^{1\times n}$ such that (3.33) with $p_{[l]}$ in (3.30).
4. Construct $P \in \mathbb{R}^{\hat{n}\times n}$ such that (3.38) along with the above procedure.
5. For a given $\delta \in \mathbb{R}_+$, solve (3.22) while minimizing γ.
6. If (3.22) is infeasible or (3.42) does not hold, then take larger δ and smaller θ and back to 2).
7. Compute H by (3.23) and construct an average state observer O_P in (3.8).

Finally, it should be noted that since the number of decision variables of LMI given by (3.22) is $\frac{1}{2}\hat{n}(\hat{n}-1) + \hat{n}m_y$, this design procedure is computationally tractable if $\hat{n}$ is small.

Remark 3.8. If we do not know the direction of x_0 in advance, then we can not evaluate $\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2$ by explicit use of the information of x_0. In view of this, in Theorem 3.2, $\|\Delta_P(t;0,x_0,0)\|_{\mathcal{L}_2}^2$ is evaluated by using a conservative bound as in (3.26), which is satisfied for any $x_0 \in \mathbb{R}^n$ satisfying $\|x_0\| = 1$. Thus, if we take a larger α, then the performance of designed observers based on such evaluation tends to become conservative compared with the case of explicit use of the information of x_0.

3.4 Numerical Simulation

In this section, we show the efficiency of the proposed average state observer through a numerical example. We deal with a 1000-dimensional reaction-diffusion system evolving over a complex network model, called the Dorogovtsev model [54]. This complex network model is a directed graph given as a generalization of the Barabashi-Albert model (bidirected graph) having small-world and scale-free properties. The reaction-diffusion system on this directed graph is shown in FIGURE. 3.2.

The parameters of the reaction-diffusion system are given as follows: The dynamics of this system is described as (3.32) where the diffusion terms $a_{i,j}$ are randomly chosen from

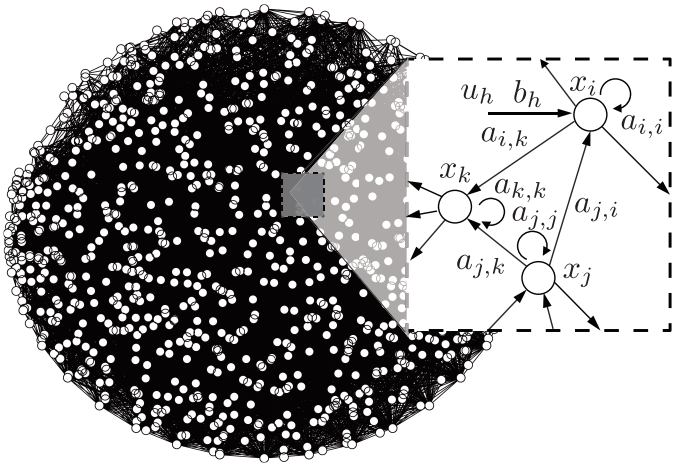


FIGURE 3.2: Dynamical system evolving over the Dorogovtsev model

$[-1, 1]$ if the ith node is directed from the jth node, otherwise 0, and the reaction terms $a_{i,i}$ are randomly chosen from $[-1, 0]$. Using these variables, we take $A \in \mathbb{R}^{1000\times1000}$ in (3.1) as

$$A_{ij} = \begin{cases} -a_{i,j} & i \neq j \\ a_{i,i} + \sum_{k\neq i} a_{i,k} & i = j \end{cases}$$

Let the input signal $u \in \mathbb{R}^5$ be applied to the states of some five nodes, and the measurement output signal $y \in \mathbb{R}^{10}$ is obtained as the states of some 10 nodes.

We show a design result of average state observers. Taking $\alpha = 0$, $\delta = 5.0$ and $\rho = 4.3$, an average state observer is constructed by the procedure shown in Section 3.3.3. In addition, S is constructed by (3.35), which implies that z estimates an average state in the sense of (3.36). The resultant dimension of the obtained observer is $\hat{n} = 5$, which implies that the resultant number of clusters is $L = 5$. We plot all trajectories of x and $\hat{z}$ in FIGURE. 3.3 where we take an input signal u randomly. The indications are as follows: the trajectory of $\hat{z} = [\hat{z}_1, \ldots, \hat{z}_5]^\mathsf{T}$ is depicted by the dotted lines with circles where $\hat{z}_l$ is color-coded for each $l \in \{1, \ldots, 5\}$. In addition, x_i for each $i \in \mathcal{I}_{[l]}$ is also color-coded according to its cluster index $l \in \{1, \ldots, 5\}$. From this figure, we can see that each trajectory of five elements of $\hat{z}$ is around the center of colored trajectory sets of x. Furthermore, the resultant estimation error turns out to be $\|z - \hat{z}\|^2_{\mathcal{L}_2} = 1.3\times10^{-2}$, which implies that $\hat{z}$ estimates an average state in the sense of (3.36) efficiently. As demonstrated in this numerical example, the proposed average state observer can efficiently find and estimate the average behavior of network systems.

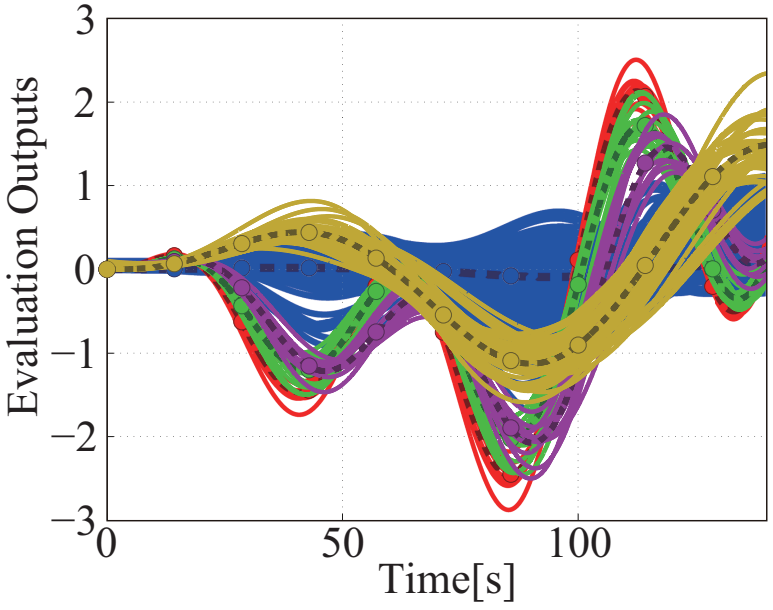


FIGURE 3.3: Trajectories of $\hat{z}$ and x.

3.5 Chapter Summary

In this chapter, we have proposed a novel framework of low-dimensional functional state observers called average state observers, which estimates average behavior of systems from a macroscopic point of view. By the proposed method, we can construct a performance guaranteed average state observer with systematic determination of a set of states capturing average behavior of systems. Towards systematic design of average state observers as well as determination of a set of states capturing average behavior, we have first derived a tractable representation of the error system. Then, we have clarified differences between the design of classical functional state observers and that of average state observers and shown a theoretical $\mathcal{L}_2$-error bound of estimation error. The efficiency of the proposed average state observers has been shown through a numerical example of a reaction-diffusion system evolving over a directed complex network.

Chapter 4

Hierarchical Distributed Control

4.1 Introduction

In this chapter, we propose a design method of hierarchical distributed controllers for general linear network systems. The hierarchical distributed controller has an advantage that an $\mathcal{L}_2$-performance of the closed-loop system is guaranteed for all sets of locally stabilizing controllers. Towards systematic distributed design, we first introduce state-space expansion, similar to one in [55], to independently deal with the state variables associated with disjoint subsystems and those associated with the interference among hierarchically clustered subsystems. This state-space expansion enables us to construct a hierarchically structured controller that attenuates the negative interference not only among hierarchically clustered subsystems but also among locally stabilizing controllers.

By the hierarchical distributed controller, whose compositional units can be designed individually, it is shown that an $\mathcal{L}_2$-performance of closed-loop systems improves as just improving an $\mathcal{L}_2$-performance of local controllers. Moreover, by the integration of a hierarchical distributed observer proposed in [56] having good compatibility with the hierarchical distributed controller, we build a framework to implement an observer-based hierarchical distributed control. The efficiency of the proposed control is shown through a numerical example for power networks.

The organization of this chapter is as follows: In Section 4.2, providing a mathematical formulation of hierarchically clustered network systems, we first formulate a design problem of hierarchical distributed controllers. In Section 4.3.1, on the basis of state-space expansion which enables us to deal with network systems in a hierarchical fashion, we give a solution to the hierarchical distributed control design problem. Furthermore, in Section 4.3.2, we build a framework to implement an observer-based hierarchical distributed control. In Section 4.4 we demonstrate the efficiency of the proposed control

structure through a numerical example of power network systems. Finally, concluding remarks are given in Section 4.5.

4.2 Problem Formulation

4.2.1 Review of Decentralized Control

In this chapter, we deal with linear network systems composed of N subsystems. For $i \in \mathcal{N} := \{1, \ldots, N\}$, we give the dynamics of the ith subsystem as follows:

$$\Sigma_i : \begin{cases} \dot{x}_i = A_i x_i + \sum_{j \neq i}^{N} A_{i,j} x_j + B_i u \\ y_i = C_i x_i \end{cases} \tag{4.1}$$

where $A_i \in \mathbb{R}^{n_i \times n_i}$, $A_{i,j} \in \mathbb{R}^{n_i \times n_j}$, $B_i \in \mathbb{R}^{n_i \times m_i}$, and $C_i \in \mathbb{R}^{p_i \times n_i}$. In this notation, we consider a family of local controllers that stabilizes each Σ_i by an input signal u_i and an measurement output y_i. The local controller associated with Σ_i is described by

$$\kappa_i : \begin{cases} \dot{\xi}_i = K_i \xi_i + H_i y_i \\ u_i = M_i \xi_i \end{cases} \tag{4.2}$$

where $K_i \in \mathbb{R}^{r_i \times r_i}$, $H_i \in \mathbb{R}^{r_i \times p_i}$, and $M_i \in \mathbb{R}^{m_i \times r_i}$.

Let us consider the disjoint subsystem with the local controller described by

$$\begin{bmatrix} \dot{x}_i \\ \dot{\xi}_i \end{bmatrix} = \begin{bmatrix} A_i & B_i M_i \\ H_i C_i & K_i \end{bmatrix} \begin{bmatrix} x_i \\ \xi_i \end{bmatrix}, \quad i \in \mathcal{N}. \tag{4.3}$$

For a given constant $\theta_i \in \mathbb{R}_+$, we suppose that (4.3) satisfies

$$\|x_i(t)\|_{\mathcal{L}_2} \leq \theta_i$$

for all $x_i(0) \in \mathbb{R}^{n_i}$ such that $\|x(0)\| = 1$ where $x := [x_1^\mathsf{T}, \ldots, x_N^\mathsf{T}]^\mathsf{T}$. Clearly, if all subsystems are disjoint, then the closed-loop system $(\{\Sigma_i\}_{i \in \mathcal{N}}, \{\kappa_i\}_{i \in \mathcal{N}})$ achieves the $\mathcal{L}_2$-performance such that

$$\|x(t)\|_{\mathcal{L}_2} \leq \|\theta\| \tag{4.4}$$

where $\theta := [\theta_1, \ldots, \theta_N]^\mathsf{T}$. In what follows, we denote a set of local controllers achieving the $\mathcal{L}_2$-performance in (4.4) for disjoint subsystems, i.e., Σ_i in (4.1) with $A_{i,j} = 0$ for all $j \in \mathcal{N} \backslash \{i\}$, as

$$\mathcal{K}_\theta := \{\{\kappa_i\}_{i \in \mathcal{N}} : \text{satisfying (4.4)}\} \tag{4.5}$$

However, if $A_{i,j} \neq 0$, i.e., if subsystems are connected, the $\mathcal{L}_2$-performance of disjoint closed-loop systems does not provide any guarantee for the overall closed-loop system in general. To make matters worse, the overall closed-loop system possibly becomes unstable even though disjoint closed-loop systems are stable. In this chapter, we consider constructing a hierarchical distributed controller that attenuates negative interference among subsystems.

4.2.2 Hierarchical Distributed Control Problem

In what follows, we use the notation of

$$n := \sum_{i=1}^{N} n_i, \quad m := \sum_{i=1}^{N} m_i, \quad p := \sum_{i=1}^{N} p_i \quad r := \sum_{i=1}^{N} r_i$$

and

$$A := \begin{bmatrix} A_1 & \cdots & A_{1,N} \\ \vdots & \ddots & \vdots \\ A_{N,1} & \cdots & A_N \end{bmatrix} \in \mathbb{R}^{n \times n}. \tag{4.6}$$

We consider introducing a hierarchical structure into network systems. Let $\mathcal{L} := \{1, \ldots, L\}$ with an integer L that represents the number of system layers. We define a family of index sets $\{\mathcal{N}^{(l)}\}_{l \in \mathcal{L}}$ satisfying

$$N \geq |\mathcal{N}^{(1)}| \geq \cdots \geq |\mathcal{N}^{(L)}| = 1, \quad \mathcal{N}^{(l)} = \{1, \ldots, |\mathcal{N}^{(l)}|\}. \tag{4.7}$$

Moreover, for each $l \in \{0, \ldots, L-1\}$, we define a set of cluster sets $\{\mathcal{C}_i^{(l)}\}_{i \in \mathcal{N}^{(l+1)}}$ satisfying

$$\bigcup_{i \in \mathcal{N}^{(l+1)}} \mathcal{C}_i^{(l)} = \mathcal{N}^{(l)}, \quad \mathcal{C}_i^{(l)} \cap \mathcal{C}_j^{(l)} = \emptyset, \quad i \neq j, \tag{4.8}$$

where $\mathcal{N}^{(0)}$ is regarded as $\mathcal{N}$.

Let $A_i^{(l)} \in \mathbb{R}^{n_i^{(l)} \times n_i^{(l)}}$ be the principal submatrix of A compatible with $\mathcal{C}_i^{(l-1)}$. Note that we have

$$\sum_{i \in \mathcal{N}^{(l)}} n_i^{(l)} = n, \quad l \in \mathcal{L},$$

and $A^{(L)} = A$. In what follows, we regard $A_i^{(0)}$ as A_i for all $i \in \mathcal{N}$.

We give the dynamics of the overall network system as

$$\Sigma : \begin{cases} \dot{x} = Ax + \mathrm{dg}(B_i)u + \sum_{l=1}^{L} \mathrm{dg}(B_i^{(l)})u^{(l)} \\ y = \mathrm{dg}(C_i)x \end{cases} \tag{4.9}$$

where the input signal $u := [u_1^\mathsf{T}, \ldots, u_N^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^m$ and the measurement output signal $y := [y_1^\mathsf{T}, \ldots, y_N^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^p$ are used for the interconnection of local controllers, and $u^{(l)}$ represents an additional input signal from a hierarchical distributed controller described below. In what follows, the pair $(A_i^{(l)}, B_i^{(l)})$, which is defined as being compatible with the hierarchical structure of network systems, is supposed to be stabilizable for any $i \in \mathcal{N}^{(l)}$ and $l \in \mathcal{L}$. Similarly, for $\xi := [\xi_1^\mathsf{T}, \ldots, \xi_N^\mathsf{T}]^\mathsf{T} \in \mathbb{R}^r$, we give the dynamics of local controllers by

$$\{\kappa_i\}_{i\in\mathcal{N}} : \begin{cases} \dot{\xi} = \mathrm{dg}(K_i)\xi + \mathrm{dg}(H_i)(y+z) \\ u = \mathrm{dg}(M_i)\xi \end{cases} \tag{4.10}$$

where the term of z expresses an additional input signal as well.

To construct appropriate additional input signals $\{u^{(l)}\}_{l\in\mathcal{L}}$ and z, we consider designing a hierarchical distributed controller given by

$$\Phi^{(l)} : \begin{cases} \dot{\phi}^{(l)} = \mathrm{dg}(\mathbf{E}_i^{(l)})\phi^{(l)} + \mathbf{G}^{(l)}x + \sum_{k=l}^{L} \mathrm{dg}(\mathbf{B}_i^{(k)})u^{(k)} \\ u^{(l)} = \mathrm{dg}(\mathbf{F}_i^{(l)})(\phi^{(l)} - \phi^{(l+1)}) \end{cases} \tag{4.11}$$

where $\phi^{(L+1)}$ is regarded as zero, and

$$\mathbf{E}_i^{(l)} \in \mathbb{R}^{n_i^{(l)}\times n_i^{(l)}}, \quad \mathbf{F}_i^{(l)} \in \mathbb{R}^{m_i^{(l)}\times n_i^{(l)}}, \quad \mathbf{B}_i^{(l)} \in \mathbb{R}^{n_i^{(l)}\times m_i^{(l)}}, \quad \mathbf{G}^{(l)} \in \mathbb{R}^{n\times n}$$

are design parameters. Furthermore, the additional input to local controllers is generated as

$$z = \mathrm{dg}(\mathbf{H}_i)\phi^{(1)}$$

where $\mathbf{H}_i \in \mathbb{R}^{p_i\times n_i}$ is another design parameter. For simplicity, we assume that $\xi(0) = 0$ and $\phi^{(l)}(0) = 0$ for all $l \in \mathcal{L}$. Moreover, $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ denotes the hierarchical distributed controller. In this setting, we address the following control problem for the closed-loop system $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$:

Problem 4.1. *Given $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8), consider Σ in (4.9) with $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10). Then, for a given constant $\epsilon > 0$, find $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ in (4.11) satisfying*

$$\|x(t)\|_{\mathcal{L}_2} \le \|\theta\| + \epsilon \tag{4.12}$$

for all $x(0) \in \mathbb{R}^n$ such that $\|x(0)\| = 1$ and for all $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.

In Problem 4.1, we formulate a problem to find a hierarchical distributed controller achieving an $\mathcal{L}_2$-performance of the closed-loop system, which necessarily implies the stability of the closed-loop system, is guaranteed for all sets of local controllers in a class $\mathcal{K}_\theta$ in (4.5). In FIGURE. 4.1, we provide an example of hierarchical distributed control

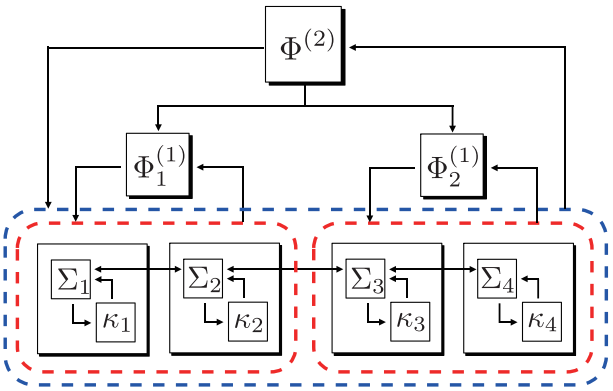


FIGURE 4.1: Hierarchical distributed control systems with $L = 2$ and $N = 4$.

systems $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ with $L = 2$ and $N = 4$. In this example, the number of controllers in each layer is given by

$$|\mathcal{N}^{(1)}| = 2, \quad |\mathcal{N}^{(2)}| = 1$$

and the family of cluster sets is given by

$$\mathcal{C}_1^{(0)} = \{1, 2\}, \quad \mathcal{C}_2^{(0)} = \{3, 4\}, \quad \mathcal{C}_1^{(1)} = \{1, 2\},$$

which satisfy $\mathcal{C}_1^{(0)} \cup \mathcal{C}_2^{(0)} = \mathcal{N}$ and $\mathcal{C}_1^{(1)} = \mathcal{N}^{(1)}$.

Remark 4.1. The development (or growth) of local subsystems can be interpreted as the variation of local controllers. For example, let us consider the case where new equipment is installed into a networked system. Regarding the additional equipment as the change of a local controller, we see that the hierarchical distributed controller solving Problem 4.1 can guarantee the stability of the whole networked systems, as long as the additional installation does not violate the stability of the local closed-loop system. In this sense, the hierarchical distributed controller to be designed is robust against the variation of local subsystems.

4.3 Hierarchical Distributed Control Systems

4.3.1 Design of Hierarchical Distributed Controllers

For systematic design of hierarchical distributed controllers, we consider transforming the realization of Σ into a tractable one based on the following state-space expansion:

Lemma 4.2. *Given* $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ *and* $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ *such that (4.7) and (4.8), consider* Σ *in (4.9). For* $l \in \mathcal{L}$*, define*

$$\begin{cases} \dot{\tilde{x}}^{(l)} = \mathrm{dg}(A_i^{(l)})\tilde{x}^{(l)} + \mathrm{dg}(B_i^{(l)})u^{(l)} + \Gamma^{(l)}\sum_{k=0}^{l-1}\tilde{x}^{(k)} \\ \dot{\tilde{x}}^{(0)} = \mathrm{dg}(A_i)\tilde{x}^{(0)} + \mathrm{dg}(B_i)u \end{cases} \tag{4.13}$$

where

$$\Gamma^{(l)} := \mathrm{dg}(A_i^{(l)})_{i\in\mathcal{N}^{(l)}} - \mathrm{dg}(A_i^{(l-1)})_{i\in\mathcal{N}^{(l-1)}}. \tag{4.14}$$

If $x(0) = \sum_{l=0}^{L}\tilde{x}_l(0)$*, then*

$$x(t) = \sum_{l=0}^{L}\tilde{x}_l(t), \quad t \geq 0 \tag{4.15}$$

for any u *and* $\{u^{(l)}\}_{l\in\mathcal{L}}$*.*

Proof. Let $\tilde{x} = [(\tilde{x}^{(L)})^\mathsf{T}, \ldots, (\tilde{x}^{(1)})^\mathsf{T}, (\tilde{x}^{(0)})^\mathsf{T}]^\mathsf{T}$. The state trajectory of (4.13) is given by

$$\tilde{x}(t) = e^{\tilde{A}t}\tilde{x}(0) + \int_0^t e^{\tilde{A}(t-\tau)}\tilde{B}\tilde{u}(\tau)d\tau$$

where $\tilde{A}$ and $\tilde{B}$ are defined as

$$\begin{aligned} \tilde{A} &:= \begin{bmatrix} A^{(L)} & \Gamma^{(L)} & \cdots & \Gamma^{(L)} & \Gamma^{(L)} \\ & \mathrm{dg}(A_i^{(L-1)}) & \cdots & \Gamma^{(L-1)} & \Gamma^{(L-1)} \\ & & \ddots & \vdots & \vdots \\ & & & \mathrm{dg}(A_i^{(1)}) & \Gamma^{(1)} \\ & & & & \mathrm{dg}(A_i) \end{bmatrix} \\ \tilde{B} &:= \begin{bmatrix} B^{(L)} & & & & \\ & \mathrm{dg}(B_i^{(L-1)}) & & & \\ & & \ddots & & \\ & & & \mathrm{dg}(B_i^{(1)}) & \\ & & & & \mathrm{dg}(B_i) \end{bmatrix} \end{aligned} \tag{4.16}$$

Noting that

$$T\tilde{A} = AT, \quad T\tilde{B} = \left[B^{(L)}, \ldots, \mathrm{dg}(B_i^{(1)}), \mathrm{dg}(B_i)\right]$$

for $T := [I_n, \ldots, I_n] \in \mathbb{R}^{n\times(L+1)n}$, we have

$$T\tilde{x}(t) = e^{At}T\tilde{x}(0) + \int_0^t e^{A(t-\tau)}\left[B^{(L)}, \ldots, \mathrm{dg}(B_i^{(1)}), \mathrm{dg}(B_i)\right]\tilde{u}(\tau)d\tau.$$

Hence, the claim follows. □

Lemma 4.2 shows that the summation of all trajectories of state variables of the expanded system in (4.13), which has a cascade structure shown in (4.16), coincides with the original trajectory $x(t)$ for any input signals. The cascade structure of (4.16) gives a clear insight into controlling the original system Σ in (4.9) by using input signals u and $\{u^{(l)}\}_{l\in\mathcal{L}}$. On the basis of this lemma, we have the following result:

Theorem 4.3. *Given $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8), consider Σ in (4.9) with $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10). Define $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ in (4.11) with*

$$\mathbf{E}_i^{(l)} = \mathrm{dg}(A_j^{(l-1)})_{j\in\mathcal{C}_i^{(l-1)}}, \quad \mathbf{F}_i^{(l)} = F_i^{(l)}, \quad \mathbf{B}_i^{(l)} = B_i^{(l)}, \quad \mathbf{G}^{(l)} = \textstyle\sum_{k=l}^{L}\Gamma^{(k)}, \quad \mathbf{H}_i = -C_i \tag{4.17}$$

where $F_i^{(l)}$ satisfies that $A_i^{(l)} + B_i^{(l)}F_i^{(l)}$ is stable, and $\Gamma^{(l)}$ in (4.14). In addition, define

$$\gamma^{(l)} := \left\| \left(sI_n - \mathrm{dg}(A_i^{(l)} + B_i^{(l)}F_i^{(l)}) \right)^{-1} \Gamma^{(l)} \right\|_{\mathcal{H}_\infty} \tag{4.18}$$

for each $l \in \mathcal{L}$. Then

$$\|x(t)\|_{\mathcal{L}_2} \le \|\theta\| \prod_{l=1}^{L} \left(1 + \gamma^{(l)}\right) \tag{4.19}$$

for all $x(0) \in \mathbb{R}^n$ such that $\|x(0)\| = 1$ and for all $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.

Proof. From Lemma 4.2, we consider the state feedback of

$$u^{(l)} = \mathrm{dg}(F_i^{(l)})\tilde{x}^{(l)}, \quad l \in \mathcal{L},$$

and the output feedback described by

$$\begin{cases} \dot{\xi} = \mathrm{dg}(K_i)\xi + \mathrm{dg}(H_iC_i)\tilde{x}^{(0)} \\ u = \mathrm{dg}(M_i)\xi \end{cases}$$

for the expanded system in (4.13). Note that this feedback system is stable because $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$, and $A_i^{(l)} + B_i^{(l)}F_i^{(l)}$ is stable for all $i \in \mathcal{N}$ and $l \in \mathcal{L}$. Taking the coordinate transformation as

$$\phi^{(l)} = \sum_{k=l}^{L} \tilde{x}^{(k)}, \quad l \in \mathcal{L} \tag{4.20}$$

with (4.15), we have the autonomous system described by

$$\begin{bmatrix} \dot{\phi}^{(L)} \\ \dot{\phi}^{(L-1)} \\ \vdots \\ \dot{\phi}^{(1)} \\ \dot{x} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} \Lambda^{(L)} & & & & \Gamma^{(L)} & \\ \Theta^{(L)} & \Lambda^{(L-1)} & & & \Gamma^{(L-1)}+\Gamma^{(L)} & \\ \vdots & \vdots & \ddots & & \vdots & \\ \Theta^{(L)} & \Theta^{(L-1)} & \dots & \Lambda^{(1)} & \Gamma^{(1)}+\cdots+\Gamma^{(L)} & \\ \Theta^{(L)} & \Theta^{(L-1)} & \dots & \Theta^{(1)} & A & \mathrm{dg}(B_iM_i) \\ & & & -\mathrm{dg}(H_iC_i) & \mathrm{dg}(H_iC_i) & \mathrm{dg}(K_i) \end{bmatrix} \begin{bmatrix} \phi^{(L)} \\ \phi^{(L-1)} \\ \vdots \\ \phi^{(1)} \\ x \\ \xi \end{bmatrix} \tag{4.21}$$

where

$$\Lambda^{(l)} := \mathrm{dg}\left(\mathrm{dg}(A_j^{(l-1)})_{j\in\mathcal{C}_i^{(l-1)}} + B_i^{(l)}F_i^{(l)}\right)_{i\in\mathcal{N}^{(l)}}$$
$$\Theta^{(l)} := \mathrm{dg}\left(B_i^{(l)}F_i^{(l)} - \mathrm{dg}(B_j^{(l-1)}F_j^{(l-1)})_{j\in\mathcal{C}_i^{(l-1)}}\right)_{i\in\mathcal{N}^{(l)}}$$

with $B_i^{(0)}F_i^{(0)} = 0$. Taking $\tilde{x}^{(0)}(0) = x(0)$, we have

$$\|\tilde{x}^{(0)}(t)\|_{\mathcal{L}_2} \le \|\theta\|.$$

Thus, (4.19) is proven by (4.15) in conjunction with the triangle inequality of the $\mathcal{L}_2$-norm and the cascade structure in (4.13). □

Theorem 4.3 shows that the hierarchical distributed controller $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ given by (4.17), whose compositional units can be designed independently of designing local controllers, achieves the $\mathcal{L}_2$-performance in (4.19). Thus, we solve Problem 4.1 by constructing the feedback gains $F_i^{(l)}$ that make the values of $\gamma^{(l)}$ in (4.18) sufficiently small. We can see from (4.19) that the $\mathcal{L}_2$-performance of the overall closed-loop system improves as just improving the $\mathcal{L}_2$-performance of local controllers in (4.4).

We see from the structure of the transfer matrix in (4.18) that the function of the controller $\Phi^{(l)}$ is to attenuate negative interference among clustered subsystems, and the magnitude of interference attenuation is measured by $\gamma^{(l)}$. Furthermore, $\mathbf{G}^{(l)}$ in (4.17) shows that the lth layer controller $\Phi^{(l)}$ uses

$$w^{(l)} := \sum_{k=l}^{L} \Gamma^{(k)} x \tag{4.22}$$

as its input signal. Note that $\Gamma^{(l)}$ in (4.14) represents the interconnection among clusters in the $(l-1)$th layer. Thus, the signal $w^{(l)}$ have the information on the interaction among clustered subsystems. In the following Sections 4.3.2 and 4.3.3, we consider the availability of $\{w^{(l)}\}_{l\in\mathcal{L}}$.

4.3.2 Integration with Hierarchical Distributed Observers

In Section 4.3.1, we have proposed a hierarchical distributed controller that can guarantee an $\mathcal{L}_2$-performance of closed-loop systems for all sets of locally stabilizing controllers. The proposed hierarchical distributed controller has an advantage that each compositional unit can be individually designed.

The hierarchical distributed controller $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ given by (4.17) requires us to measure $\{w^{(l)}\}_{l\in\mathcal{L}}$ in (4.22). In view of this, a number of sensors are possibly required to implement this hierarchical distributed controller for large-scale network systems. For example, to implement $\Phi^{(1)}$ in the first layer, we need to measure $w^{(1)}$ that contains the information on the interaction among *all* subsystems Σ_i.

In view of this, we consider estimating $w^{(l)}$ for lower layer controllers from other sensor signals by utilizing a hierarchical distributed observer [56] having good compatibility with the hierarchical structure of control systems. To this end, for $\hat{\mathcal{L}} := \{1, \ldots, \hat{L}\}$ with an integer $\hat{L} < L$, we assume that

$$\begin{aligned} y^{(l)} &:= \mathrm{dg}(C_i^{(l)})x, \quad && l \in \hat{\mathcal{L}} \\ v^{(l)} &:= \Gamma^{(l)}x, && l \in \mathcal{L}\backslash\hat{\mathcal{L}} \end{aligned} \tag{4.23}$$

are available as sensor signals from clustered subsystems. In addition, the pair $(A_i^{(l)}, C_i^{(l)})$ is supposed to be detectable for any $i \in \mathcal{N}^{(l)}$ and $l \in \hat{\mathcal{L}}$. Note that the availability of $\{v^{(l)}\}_{l\in\mathcal{L}\backslash\hat{\mathcal{L}}}$ is equal to that of $\{w^{(l)}\}_{l\in\mathcal{L}\backslash\hat{\mathcal{L}}}$. In this setting, we show that the following observer-based hierarchical distributed control is available:

Theorem 4.4. *Given $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8), consider Σ in (4.9) with $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10). For $y^{(l)}$ in (4.23) with $H_i^{(l)}$ such that $A_i^{(l)} - H_i^{(l)}C_i^{(l)}$ is stable, define $\{o^{(l)}\}_{l\in\hat{\mathcal{L}}}$ by*

$$o^{(l)}: \begin{cases} \dot{\hat{x}}^{(l)} = \mathrm{dg}(A_i^{(l)} - H_i^{(l)}C_i^{(l)})\hat{x}^{(l)} + \mathrm{dg}(B_i)u + \sum_{k=1}^{L}\mathrm{dg}(B_i^{(k)})u^{(k)} + \mathrm{dg}(H_i^{(l)})y^{(l)} + \hat{w}^{(l+1)} \\ \hat{v}^{(l)} = \Gamma^{(l)}\hat{x}^{(l)} \end{cases} \tag{4.24}$$

with

$$\hat{w}^{(l)} := \begin{cases} \sum_{k=l}^{\hat{L}}\hat{v}^{(k)} + w^{(\hat{L}+1)}, & l \in \hat{\mathcal{L}}, \\ w^{(l)}, & l \in \mathcal{L}\backslash\hat{\mathcal{L}}, \end{cases} \tag{4.25}$$

where $w^{(l)}$ is defined as in (4.22). Furthermore, by replacing $w^{(l)}$ with $\hat{w}^{(l)}$, define $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ in (4.11) with (4.17). Then $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ with $\{o^{(l)}\}_{l\in\hat{\mathcal{L}}}$ is stable for all $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.

Proof. Define $e^{(l)} := x - \hat{x}^{(l)}$ for $l \in \hat{\mathcal{L}}$. Since $\hat{w}^{(l)}$ for $l \in \hat{\mathcal{L}}$ can be rewritten as

$$\hat{w}^{(l)} = w^{(l)} - \sum_{k=l}^{\hat{L}} \Gamma^{(k)} e^{(k)},$$

we see that the closed-loop system is stable as long as $e^{(l)} = 0$ for all $l \in \hat{\mathcal{L}}$. Thus, what remains to be shown is $\lim_{t\to\infty} e^{(l)}(t) = 0$. Noting that

$$A = \mathrm{dg}(A_i^{(l)})_{i\in\mathcal{N}^{(l)}} + \sum_{k=l+1}^{L} \Gamma^{(k)},$$

we can express the dynamics of $e^{(l)}$ by

$$\dot{e}^{(l)} = \mathrm{dg}(A_i^{(l)} - H_i^{(l)} C_i^{(l)})_{i\in\mathcal{N}^{(l)}} e^{(l)} + \sum_{k=l+1}^{\hat{L}} \Gamma^{(k)} e^{(k)},$$

where the last term is replaced with zero if $l = \hat{L}$. Hence, the claim follows from the stability of $A_i^{(l)} - H_i^{(l)} C_i^{(l)}$. □

The hierarchical distributed observer $\{o^{(l)}\}_{l\in\hat{\mathcal{L}}}$ in (4.24) gives the estimate $\{\hat{w}^{(l)}\}_{l\in\hat{\mathcal{L}}}$, which represent the interference among clustered subsystems, by the available measurement sensor signals in (4.23). Even though the hierarchical distributed observer needs to use the input signals u and $\{u^{(l)}\}_{l\in\mathcal{L}}$, each local observer only needs to obtain a part of input signals produced by its supervisor and subordinate controllers, owing to the hierarchical structure of control systems.

Remark 4.5. In Theorem 4.4, to simplify the arguments, we only provide a result on the stability of the observer-based hierarchical distributed control systems. A result on an $\mathcal{L}_2$-performance, similar to Theorem 4.3, is available also for the observer-based control, based on the separation principle for controller and observer design, which has been used to prove Theorem 4.4.

4.3.3 Discussion on Hierarchical Clustering and Sensor and Actuator Allocation Towards Scalable Implementation

In Sections 4.3.1 and 4.3.2, we have addressed a design problem of hierarchical distributed control systems assuming that the hierarchical clusters of networked systems, i.e., $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$, are given in advance. In what follows, we discuss how hierarchical clustering should be determined to implement the hierarchical distributed control systems in a desirable manner.

We notice that the low-rankness of $\Gamma^{(l)}$ for $l \in \mathcal{L}\backslash\hat{\mathcal{L}}$ has a direct relationship with the number of sensors to obtain $\{w^{(l)}\}_{l\in\mathcal{L}\backslash\hat{\mathcal{L}}}$. Note that $\Gamma^{(l)}$ becomes a lower-rank matrix if the interconnection among the corresponding clusters is sparser. Thus, a sparser interconnection structure among upper layer clusters is more desirable to reduce the number of required sensors.

Furthermore, recall that the function of $\Phi^{(l)}$ is to attenuate negative interference among clustered subsystems, and the degree of interference attenuation is measured by $\gamma^{(l)}$ in (4.18). Inspecting the structure of the transfer matrix in (4.18), we see that the magnitude of $\gamma^{(l)}$ can be efficiently reduced if

(i) the rank of $\Gamma^{(l)}$ is low enough, and

(ii) the input signal given by $B_i^{(l)}$ can effectively attenuate the interference signal injected through $\Gamma^{(l)}$.

Thus, we can expect that a suitable determination of hierarchical clustering as well as the allocation of actuators effectively contribute to improve the $\mathcal{L}_2$-performance of hierarchical distributed control systems.

Such a suitable determination of hierarchical clustering and actuator allocation can also contribute to make *practical dimension* of upper layer controllers lower. The meaning of practical dimension is explained as follows: Let us consider the topmost layer controller $\Phi^{(L)}$, for instance. By definition, the state-space of $\Phi^{(L)}$ necessarily has the dimension comparable with the whole networked system. On the other hand, if the items (i) and (ii) above are satisfied, the decay rate of the Hankel singular values of $\Phi^{(L)}$ tends to be fast [57]. This implies that the Hankel matrix associated with $\Phi^{(L)}$, whose rank is equal to its McMillan degree, turns out to be low-rank or near low-rank. Thus, the topmost layer controller can potentially be implemented as a lower-dimensional model as long as the items (i) and (ii) are satisfied; see Section 4.4.2 below for a numerical demonstration.

The rapid decay of the Hankel singular values has a deep connection with the possibility to finely approximate a dynamical system by a low-dimensional model. Indeed, the magnitude of them is closely related to the approximation error via the Hankel norm approximation as well as the balanced truncation/residualization [17]. In conclusion, we see that sparse interconnection among upper layer clusters is desirable to reduce the number of required sensors as well as that of actuators to attenuate negative interference among clustered subsystems.

4.4 Numerical Example

4.4.1 Power Network Model

In this section, we demonstrate the proposed hierarchical distributed control through an example of electric power network systems. We deal with a power network model [11] composed of N subnetworks (subsystems), where the ith subsystem includes n_i^{G} generators and n_i^{L} loads.

For $k \in \mathcal{N}_i^{\mathrm{G}} := \{1, \ldots, n_i^{\mathrm{G}}\}$, the model of the kth generator is given by

$$\Sigma_{[i]k}^{\mathrm{G}} : \begin{cases} \dot{\zeta}_{[i]k} = A_{[i]k}^{\mathrm{G}} \zeta_{[i]k} + \frac{1}{M_{[i]k}^{\mathrm{G}}} b^{\mathrm{G}} \theta_{[i]k}^{\mathrm{G}} + \frac{1}{T_{[i]k}^{\mathrm{G}}} b u_{[i]k} \\ \delta_{[i]k}^{\mathrm{G}} = c^{\mathrm{G}} \zeta_{[i]k} \end{cases} \tag{4.26}$$

where each element of $\zeta_{[i]k} \in \mathbb{R}^3$ represents a phase angle difference, an angular velocity difference and a mechanical input difference, and $\theta_{[i]k}^{\mathrm{G}} \in \mathbb{R}$, $\delta_{[i]k}^{\mathrm{G}} \in \mathbb{R}$, and $u_{[i]k} \in \mathbb{R}$ represent an electric output difference, a phase angle difference, and a valve position difference, respectively. In addition, we give the system matrices in (4.26) by

$$A_{[i]k}^{\mathrm{G}} := \begin{bmatrix} 0 & 1 & 0 \\ 0 & -D_{[i]k}^{\mathrm{G}}/M_{[i]k}^{\mathrm{G}} & -1/M_{[i]k}^{\mathrm{G}} \\ 0 & 0 & -1/T_{[i]k}^{\mathrm{G}} \end{bmatrix}, \quad b^{\mathrm{G}} := e_2^3, \quad c^{\mathrm{G}} := (e_1^3)^{\mathsf{T}}, \quad b := e_3^3$$

where $M_{[i]k}^{\mathrm{G}}$, $D_{[i]k}^{\mathrm{G}}$ and $T_{[i]k}^{\mathrm{G}}$ represent a mechanical inertia, a damping coefficient and a turbine time constant, respectively.

In a similar fashion, for $i \in \mathcal{N}_{[i]}^{\mathrm{L}} := \{1, \ldots, n_{[i]}^{\mathrm{L}}\}$, we give the model of the ith load by

$$\Sigma_{[i]k}^{\mathrm{L}} : \begin{cases} \dot{\psi}_{[i]k} = A_{[i]k}^{\mathrm{L}} \psi_{[i]k} + \frac{1}{M_{[i]k}^{\mathrm{L}}} b^{\mathrm{L}} \theta_{[i]k}^{\mathrm{L}} \\ \delta_{[i]k}^{\mathrm{L}} = c^{\mathrm{L}} \psi_{[i]k} \end{cases} \tag{4.27}$$

where each state of $\psi_{[i]k} \in \mathbb{R}^2$ represents a phase angle difference and an angular velocity difference, and $\theta_{[i]k}^{\mathrm{L}} \in \mathbb{R}$ and $\delta_{[i]k}^{\mathrm{L}} \in \mathbb{R}$ represent an electric output difference and a phase angle difference, respectively. Then, we give the system matrices in (4.27) by

$$A_{[i]k}^{\mathrm{L}} := \begin{bmatrix} 0 & 1 \\ 0 & -D_{[i]k}^{\mathrm{L}}/M_{[i]k}^{\mathrm{L}} \end{bmatrix}, \quad b^{\mathrm{L}} := e_2^2, \quad c^{\mathrm{L}} := (e_1^2)^{\mathsf{T}}$$

where $M_{[i]k}^{\mathrm{L}}$ and $D_{[i]k}^{\mathrm{L}}$ represent an inertia constant and a damping coefficient, respectively.

The interconnection structure among generators and loads are given by

$$\theta = -Y\delta, \quad \theta := \begin{bmatrix} \theta_1^{\mathrm{G}} \\ \theta_1^{\mathrm{L}} \\ \vdots \\ \theta_N^{\mathrm{G}} \\ \theta_N^{\mathrm{L}} \end{bmatrix}, \quad \delta := \begin{bmatrix} \delta_1^{\mathrm{G}} \\ \delta_1^{\mathrm{L}} \\ \vdots \\ \delta_N^{\mathrm{G}} \\ \delta_N^{\mathrm{L}} \end{bmatrix} \tag{4.28}$$

where $Y \in \mathbb{R}^{N_Y \times N_Y}$ represents an admittance matrix satisfying

$$Y = Y^{\mathsf{T}}, \quad Y\mathbf{1}_{N_Y} = 0, \quad N_Y := \sum_{i=1}^{N} n_i^{\mathrm{G}} + n_i^{\mathrm{L}},$$

and

$$\theta_i^{\star} := \begin{bmatrix} \theta_{[i]1}^{\star} \\ \vdots \\ \theta_{[i]n_i^{\star}}^{\star} \end{bmatrix}, \quad \delta_i^{\star} := \begin{bmatrix} \delta_{[i]1}^{\star} \\ \vdots \\ \delta_{[i]n_i^{\star}}^{\star} \end{bmatrix}, \quad \star \in \{\mathrm{G}, \mathrm{L}\}.$$

We define a state variable as $x := [\zeta_1^{\mathsf{T}}, \psi_1^{\mathsf{T}}, \ldots, \zeta_N^{\mathsf{T}}, \psi_N^{\mathsf{T}}]^{\mathsf{T}}$ where

$$\zeta_i := \begin{bmatrix} \zeta_{[i]1} \\ \vdots \\ \zeta_{[i]n_i^{\mathrm{G}}} \end{bmatrix}, \quad \psi_i := \begin{bmatrix} \psi_{[i]1} \\ \vdots \\ \psi_{[i]n_i^{\mathrm{L}}} \end{bmatrix}.$$

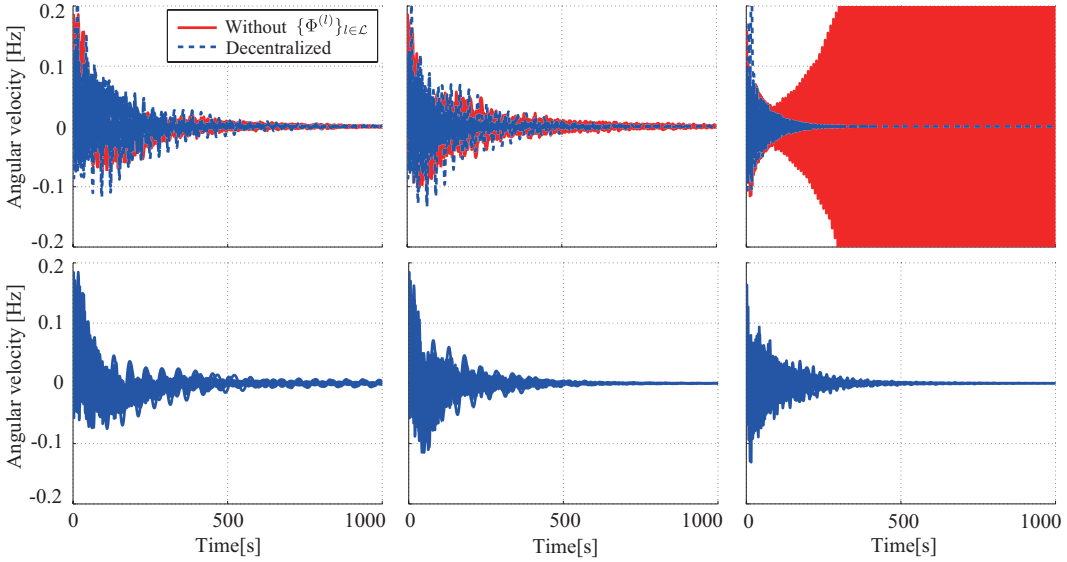
Furthermore, we define the input u in (4.9) by

$$u := \begin{bmatrix} u_{[1]} \\ \vdots \\ u_{[N]} \end{bmatrix}, \quad u_{[i]} := \begin{bmatrix} u_{[i]1} \\ \vdots \\ u_{[i]n_i^{\mathrm{G}}} \end{bmatrix} \tag{4.29}$$

and the output y by

$$y := \begin{bmatrix} \zeta_{[1]}^{1:2} \\ \vdots \\ \zeta_{[N]}^{1:2} \end{bmatrix}, \quad \zeta_{[i]}^{1:2} := \begin{bmatrix} \zeta_{[i]1}^{1:2} \\ \vdots \\ \zeta_{[i]n_i^{\mathrm{G}}}^{1:2} \end{bmatrix} \tag{4.30}$$

where $\zeta_{[i]k}^{1:2} \in \mathbb{R}^2$ denotes the first and second elements of $\zeta_{[i]k} \in \mathbb{R}^3$. In this notation, for Σ in (4.9), the system matrices of the whole power network is given by



(a) Low-gain local controllers (b) Middle-gain local controllers (c) High-gain local controllers

FIGURE 4.2: Initial value responses of power network model.

$$
\begin{aligned}
A &= \mathrm{dg}\left(\mathrm{dg}(A^{\mathrm{G}}_{[i]k}), \mathrm{dg}(A^{\mathrm{L}}_{[i]k})\right)_{i\in\mathbb{N}} \\
&\quad -\mathrm{dg}\left(\mathrm{dg}(\tfrac{1}{M^{\mathrm{G}}_{[i]k}} b^{\mathrm{G}}), \mathrm{dg}(\tfrac{1}{M^{\mathrm{L}}_{[i]k}} b^{\mathrm{L}})\right)_{i\in\mathbb{N}} Y \mathrm{dg}\left(\mathrm{dg}(c^{\mathrm{G}})_{k\in\mathbb{N}^{\mathrm{G}}_i}, \mathrm{dg}(c^{\mathrm{L}})_{k\in\mathbb{N}^{\mathrm{L}}_i}\right)_{i\in\mathbb{N}} \\
B_i &= \begin{bmatrix} \mathrm{dg}(\frac{1}{T^{\mathrm{G}}_{[i]k}} b) \\ 0_{2n^{\mathrm{L}}_i \times n^{\mathrm{G}}_i} \end{bmatrix}, \quad C_i = \begin{bmatrix} I_{n^{\mathrm{G}}_i} \otimes [I_2 \; 0_{2\times 1}] & 0_{2n^{\mathrm{G}}_i \times 2n^{\mathrm{L}}_i} \end{bmatrix}
\end{aligned} \tag{4.31}
$$

where $\otimes$ denote the Kronecker product.

Finally, we consider giving additional input ports used by a hierarchical distributed controller. Assuming that several generators have the ports for controllers in the lth layer, we give $B^{(l)}$ in (4.9) as a matrix composed of a part of columns of $\mathrm{dg}(B_i)_{i\in\mathcal{N}}$. Similarly to this, we give the additional output $y^{(l)}$ in (4.23) for observers in the lth layer as a part of y in (4.30).

4.4.2 Hierarchical Distributed Control of Power Networks

We design a hierarchical distributed controller for a power network system composed of five subsystems, i.e., $N = 5$. The interconnection structure among generators and loads is shown in FIGURE. 4.3 where generators and loads are denoted by circles and diamonds, respectively. The blue and red circles represent the generators having additional input ports for a hierarchical distributed controller $\Phi^{(1)}$ and $\Phi^{(2)}$, respectively. In addition,

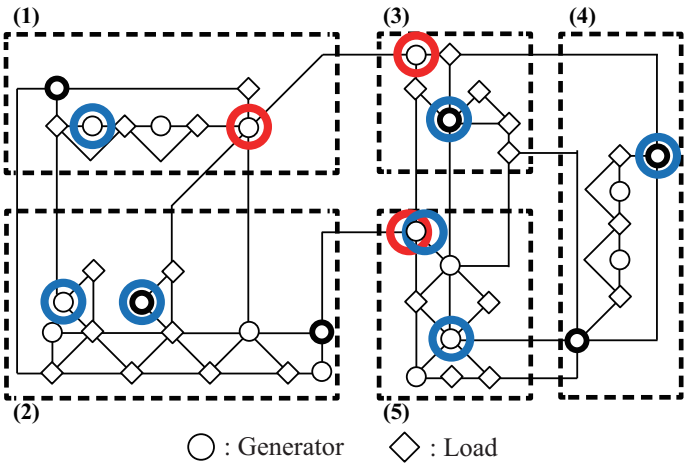


FIGURE 4.3: Interconnection structure among generators and loads.

thick circles denotes the generators equipped with sensors to be used for a hierarchical distributed observer.

This power network system involves 20 generators and 24 loads, which yields that the overall network system is 108-dimensional, i.e., $n = 108$. For generators and loads, the parameters $M^{\mathrm{G}}_{[i]k}$, $D^{\mathrm{G}}_{[i]k}$, $T^{\mathrm{G}}_{[i]k}$, $M^{\mathrm{G}}_{[i]k}$ and $D^{\mathrm{G}}_{[i]k}$ are randomly chosen from $\{10, 90\}$, $\{0.1, 0.4\}$, $\{3.0, 10\}$, $\{5, 10, 30\}$ and $\{0.1, 0.3, 0.5\}$, respectively. We give the elements of Y in (4.28) compatible with interconnection among subsystems as 1. In addition, those compatible with interconnection inside the subsystems are randomly chosen from $[0.1, 1.0]$. In what follows, to simulate a situation where the frequency of the power system suddenly varies, we give nonzero initial values for the angular velocity of generators, i.e., $\dot{\delta}^{\mathrm{G}}_{[i]k}(0) \neq 0$.

First, we design a set of locally stabilizing controllers $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10) by the LQR design techniques. Changing the weighting parameters for the LQR design, we obtain three sets of $\{\kappa_i\}_{i\in\mathcal{N}}$. The resultant values of $\|\theta\|$ in (4.19) are 1388, 423 and 126, respectively. To see the behavior of the closed-loop system, we show its initial value responses of the angular velocities of all generators and loads in the upper half of FIGURE. 5.1 (a)-(c). In this figure, we plot the trajectories of disjoint subsystems, i.e., x_i in (4.3), by the red lines, and those of subsystems with interconnection by the blue lines. Even though the convergence rate of the system without subsystem interconnection becomes higher as improving the $\mathcal{L}_2$-performance of local controllers, we see from this figure that the instability of the closed-loop system is induced by negative interference among subsystems.

Next, using a hierarchical distributed controller and observer, we consider improving the $\mathcal{L}_2$-performance of the overall closed-loop system. Let $L = 2$ and $\hat{L} = 1$ in (4.23). We

give a family of cluster sets by

$$\mathcal{C}_1^{(0)} = \{1, 2\}, \quad \mathcal{C}_2^{(0)} = \{3, 4, 5\}, \quad \mathcal{C}_1^{(1)} = \{1, 2\}, \tag{4.32}$$

where each of $\mathcal{C}_1^{(0)}$ and $\mathcal{C}_2^{(0)}$ includes 10 generators and 12 loads.

We design each controller $\Phi^{(l)}$ by minimizing $\gamma^{(l)}$ in (4.18). From the structure of hierarchical distributed observers in (4.24), we design the observer gain $H_i^{(l)}$ by minimizing the $\mathcal{H}_\infty$-norm of the transfer matrix compatible with the pair $(\mathrm{dg}(A_i^{(l)} - H_i^{(l)} C_i^{(l)}), \Gamma^{(l)})$. Then, we implement the hierarchical distributed controller and observer. In the lower half of FIGURE. 5.1 (a)-(c), we show the initial value responses of the hierarchical distributed control system. For each case of (a)-(c), the resultant value of $\|x\|_{\mathcal{L}_2}$ turns out to be 1189, 491 and 234, respectively. We see from this result that the $\mathcal{L}_2$-performance of the closed-loop system improves as improving the performance of local controllers, owing to the attenuation of interference among subsystems. This application demonstrates that the hierarchical distributed controller can be a new frequency controller (post-LFC) of power systems.

4.5 Chapter Summary

In this chapter, we have proposed a design method of hierarchical distributed controllers for linear network systems. For systematic design, we have used state-space expansion that enables us to construct a hierarchically structured controller that attenuates negative interference among hierarchically clustered subsystems as well as among locally stabilizing controllers. On the basis of this state-space expansion, we have devised a design method to construct a hierarchical distributed controller whose compositional units can be designed individually. Furthermore, the hierarchical distributed controller has an advantage that an $\mathcal{L}_2$-performance of the closed-loop system improves as just improving an $\mathcal{L}_2$-performance of local controllers that stabilizes disjoint subsystems individually. Moreover, we have built a framework to implement an observer-based hierarchical distributed control by integrating a hierarchical distributed observer having good compatibility with the hierarchical distributed controller. Finally, the efficiency of the proposed method has been shown through an illustrative example of power networks. The proposed method has a potential to robustly control large-scale power systems against modification of local controllers.

Chapter 5

Low-dimensional Hierarchical Distributed Controller Design

5.1 Introduction

In Chapter 4, we have proposed a distributed design method of hierarchical distributed controllers for general linear systems. The hierarchical distributed controller has an advantage that an $\mathcal{L}_2$-performance of the closed-loop system is guaranteed for any sets of locally stabilizing controllers. However, the dimension of designed hierarchical distributed controller $\Phi^{(l)}$ in (5.5) coincides with n, which is the dimension of the system to be controlled. Thus, the designed hierarchical distributed controllers do not fully comply with practical application for large-scale systems from a viewpoint of computational costs for implementation.

In this chapter, we propose a design method of low-dimensional hierarchical distributed controllers having an $\mathcal{L}_2$-performance of the closed-loop system for any sets of locally stabilizing controllers. We take a controller reduction approach to theoretically evaluate a performance of low-dimensional hierarchical distributed controllers for the closed-loop system. More specifically, supposing that a hierarchical distributed controller is given by the method in the previous chapter, we find a low-dimensional controller such that

- the trajectory of the system state controlled by the low-dimensional controller needs to be close to that controlled by the original controller *for any locally stabilizing controllers*, and
- the low-dimensional controller needs to have the same hierarchical distributed structure as that of the original controller.

Since the existing controller reduction techniques, e.g., [24, 25], cannot explicitly take into account the interconnection structure among controllers, this controller reduction problem cannot be solved by the straightforward use of those techniques. Thus, we explicitly utilize the hierarchical distributed structure of the closed-loop system, which has compositional units of the controller in upper layers and those in lower layers. More specifically, taking into account the inherent hierarchy of information transmission which can be represented as the block-triangular structure of a coordinate transformed closed-loop system, it is shown that the approximation error of compositional units in upper layers does not affect those in lower layers. Next, using biorthogonal projection [17], we clarify the relation between approximation errors of the compositional units and the performance degradation of the closed-loop system. Finally, we demonstrate the proposed method is demonstrated through a numerical example of power networks.

The organization of this chapter is as follows: In Section 5.2, we formulate a controller reduction problem to design low-dimensional hierarchical distributed controllers. In Section 5.3, we first show that the approximation error of the closed-loop system can be independently evaluated by using that of a system associated with the controllers in the upper layer. On the basis of this result, we provide an approximation error bound of the closed-loop system with a construction procedure of low-dimensional hierarchical distributed controllers. In Section 5.4, we demonstrate the efficiency of the proposed method through an example of power networks.

5.2 Problem Formulation

In this chapter, we deal with linear network systems similar to that in the previous chapter. In the first half of this section, we summarize some different settings.

For each $i \in \mathcal{N} := \{1, \ldots, N\}$, we give the dynamics of the ith subsystem as follows:

$$\Sigma_i : \begin{cases} \dot{x}_i = A_i x_i + B_i u_i + \sum_{j \in \mathcal{J}_i} J_{i,j} w_j \\ y_i = C_i x_i \\ w_i = S_i x_i \end{cases} \tag{5.1}$$

where $A_i \in \mathbb{R}^{n_i \times n_i}$, $B_i \in \mathbb{R}^{n_i \times m_i}$, $J_{i,j} \in \mathbb{R}^{n_i \times q_j}$, $C_i \in \mathbb{R}^{p_i \times n_i}$ and $S_i \in \mathbb{R}^{q_i \times n_i}$, and $\mathcal{J}_i \subseteq \mathcal{N}$ denotes a set of indices of the subsystems connected to the ith subsystem. In addition, let us consider a set of local controllers $\{\kappa_i\}_{i \in \mathcal{N}}$ in (4.10). In what follows, we

use the notation of $q := \sum_{i=1}^{N} q_i$ and

$$J := \begin{bmatrix} 0 & J_{1,2} & \cdots & J_{1,N} \\ J_{2,1} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & J_{N-1,N} \\ J_{N,1} & \cdots & J_{N,N-1} & 0 \end{bmatrix} \in \mathbb{R}^{n \times p} \tag{5.2}$$

where $J_{i,j} = 0$ if the ith and jth subsystems are disjoint. In addition, we define

$$A := \mathrm{dg}(A_i) + J\mathrm{dg}(S_i) \in \mathbb{R}^{n \times n}. \tag{5.3}$$

Let $\mathcal{L} := \{1, \ldots, L\}$ with an integer L that represents the number of system layers. Define $\{\mathcal{N}^{(l)}\}_{l \in \mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i \in \mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8). Let $J_i^{(l)} \in \mathbb{R}^{n_i^{(l)} \times p_i^{(l)}}$ be the principal submatrices of J compatible with $\mathcal{C}_i^{(l-1)}$, respectively. By definition, it follows that

$$\sum_{i \in \mathcal{N}^{(l)}} q_i^{(l)} = q$$

for each $l \in \mathcal{L}$, and $J^{(L)} = J$.

We give the dynamics of the whole networked system as

$$\Sigma : \begin{cases} \dot{x} = Ax + \mathrm{dg}(B_i)u + \sum_{l=1}^{L} \mathrm{dg}(B_i^{(l)})u^{(l)} \\ y = \mathrm{dg}(C_i)x \\ w = \mathrm{dg}(S_i)x \end{cases} \tag{5.4}$$

where the input signal $u := [u_1^{\mathsf{T}}, \ldots, u_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^m$ and the measurement output signal $y := [y_1^{\mathsf{T}}, \ldots, y_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^p$ are used for the interconnection to local controllers, and $u^{(l)} \in \mathbb{R}^{m^{(l)}}$ expresses an additional input signal from a hierarchical distributed controller to be explained below. In addition, $w := [w_1^{\mathsf{T}}, \ldots, w_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{R}^q$ is used for not only the interconnection among the subsystems Σ_i, but also the interconnection to the hierarchical distributed controller. Furthermore, we consider designing a hierarchical distributed controller given by

$$\Phi^{(l)} : \begin{cases} \dot{\phi}^{(l)} = \mathrm{dg}(E_i^{(l)})\phi^{(l)} + \Gamma^{(l)}y + \sum_{k=l}^{L} \mathrm{dg}(\Lambda_i^{(k)})u^{(k)} \\ u^{(l)} = \mathrm{dg}(F_i^{(l)})\phi^{(l)} + G^{(l+1)}\phi^{(l+1)} \end{cases} \tag{5.5}$$

and

$$z = \mathrm{dg}(U_i)\phi^{(1)}$$

where $G^{(L+1)}$ and $\phi^{(L+1)}$ are regarded as zero, and

$$\begin{array}{lll} E_i^{(l)} \in \mathbb{R}^{n_i^{(l)} \times n_i^{(l)}}, & \Gamma^{(l)} \in \mathbb{R}^{n \times q}, & \Lambda_i^{(l)} \in \mathbb{R}^{n_i^{(l)} \times m_i^{(l)}} \\ F_i^{(l)} \in \mathbb{R}^{m_i^{(l)} \times n_i^{(l)}}, & G^{(l+1)} \in \mathbb{R}^{m^{(l)} \times n}, & U_i \in \mathbb{R}^{p_i \times n_i^{(1)}} \end{array} \tag{5.6}$$

are design parameters. It should be noted that $\Phi^{(l)}$ is composed of $|\mathcal{N}^{(l)}|$ units such that the ith unit is an $n_i^{(l)}$-dimensional system. In what follows, we take $\phi^{(l)}(0) = 0$ for simplicity. Furthermore, we denote the hierarchical distributed controller by $\{\Phi^{(l)}\}_{l \in \mathcal{L}}$. Regarding to this hierarchical distributed controller, we have the following proposition, which is Theorem 4.3 in Chapter 4:

Proposition 5.1. *Given $\{\mathcal{N}^{(l)}\}_{l \in \mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i \in \mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8), consider Σ in (5.4) with $\{\kappa_i\}_{i \in \mathcal{N}}$ in (4.10). Give $\{\Phi^{(l)}\}_{l \in \mathcal{L}}$ in (5.5) with*

$$E_i^{(l)} = \mathrm{dg}(A_j^{(l-1)})_{j \in \mathcal{C}_i^{(l-1)}},\ \Gamma^{(l)} = \textstyle\sum_{k=l}^{L} J^{(k)},\ \Lambda_i^{(l)} = B_i^{(l)},\ \ G^{(l+1)} = -\mathrm{dg}(F_i^{(l)}),\ U_i = -C_i \tag{5.7}$$

where $F_i^{(l)}$ satisfies that $A_i^{(l)} + B_i^{(l)} F_i^{(l)}$ is stable. For each $l \in \mathcal{L}$, define

$$\gamma^{(l)} := \left\| \left(sI - \mathrm{dg}(A_i^{(l)} + B_i^{(l)} F_i^{(l)}) \right)^{-1} \mathrm{dg}(J_i^{(l)}) \mathrm{dg}(S_i) \right\|_{\mathcal{H}_\infty}. \tag{5.8}$$

Then (4.19) follows for all $x(0) \in \mathbb{R}^n$ such that $\|x(0)\| = 1$ and any $\{\kappa_i\}_{i \in \mathcal{N}} \in \mathcal{K}_\theta$ where $\mathcal{K}_\theta$ is defined as in (4.5).

Proposition 5.1 shows that the hierarchical distributed controller $\{\Phi^{(l)}\}_{l \in \mathcal{L}}$ given by (5.5), whose compositional units can be designed independently of designing local controllers, achieves an $\mathcal{L}_2$-performance of the closed-loop system for any sets of locally stabilizing controllers belonging $\mathcal{K}_\theta$. However, the $n_i^{(l)}$-dimensional compositional unit becomes the larger in scale in the upper layer. To make matters worse, the topmost controller $\Phi^{(L)}$ is n-dimensional. Thus, this hierarchical distributed controller is not practical for large-scale network systems due to high computational costs of the compositional units in the upper layer.

In the rest of this subsection, we consider designing a low-dimensional hierarchical distributed controller that the $\mathcal{L}_2$-performance of the closed-loop system is robustly guaranteed for any sets of locally stabilizing controllers in $\mathcal{K}_\theta$. To this end, let us consider

$$\hat{\Phi}^{(l)} : \begin{cases} \dot{\hat{\phi}}^{(l)} = \mathrm{dg}(\hat{\mathbf{E}}_i^{(l)}) \hat{\phi}^{(l)} + \hat{\mathbf{\Gamma}}^{(l)} y + \sum_{k=l}^{L} \mathrm{dg}(\hat{\mathbf{\Lambda}}_i^{(l,k)})_{i \in \mathcal{N}^{(k)}} u^{(k)} \\ u^{(l)} = \mathrm{dg}(\hat{\mathbf{F}}_i^{(l)}) \hat{\phi}^{(l)} + \hat{\mathbf{G}}^{(l+1)} \hat{\phi}^{(l+1)} \end{cases} \tag{5.9}$$

with $\hat{\phi}^{(l)}(0) = 0$ where $\hat{\mathbf{G}}^{(L+1)}$ and $\hat{\phi}^{(L+1)}$ are regarded as zero, and $\hat{\mathbf{\Lambda}}_i^{(l,k)}$ such that $\mathrm{dg}(\hat{\mathbf{\Lambda}}_i^{(l,k)})_{i\in\mathcal{N}^{(k)}} \in \mathbb{R}^{\hat{n}^{(l)}\times m^{(k)}}$ and

$$\hat{\mathbf{E}}_i^{(l)} \in \mathbb{R}^{\hat{n}_i^{(l)}\times \hat{n}_i^{(l)}}, \quad \hat{\mathbf{F}}_i^{(l)} \in \mathbb{R}^{m_i^{(l)}\times \hat{n}_i^{(l)}}, \quad \hat{\mathbf{\Gamma}}^{(l)} \in \mathbb{R}^{\hat{n}^{(l)}\times p}, \quad \hat{\mathbf{G}}^{(l+1)} \in \mathbb{R}^{m^{(l)}\times \hat{n}^{(l+1)}}$$

are design parameters in conjunction with $\hat{n}_i^{(l)}$ and $\hat{n}^{(l)}$ satisfying

$$\hat{n}^{(l)} = \sum_{i\in\mathcal{N}^{(l)}} \hat{n}_i^{(l)}.$$

We deal with the case of $\hat{n}_i^{(l)} \leq n_i^{(l)}$ without loss of generality. Furthermore, the additional input to local controllers is given by

$$z = \mathrm{dg}(\hat{\mathbf{U}}_i)\hat{\phi}^{(1)}$$

where $\hat{\mathbf{U}}_i \in \mathbb{R}^{q_i\times \hat{n}_i^{(1)}}$ is another design parameter.

We consider designing $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ in (5.9) such that the closed-loop system with $\{\kappa_i\}_{i\in\mathcal{N}}$ and this low-dimensional hierarchical distributed controller, i.e., $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$, achieves a desirable $\mathcal{L}_2$-performance for any $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$ by taking a controller reduction approach. Suppose that $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ given by (5.5) and (5.7) guarantees a desirable $\mathcal{L}_2$-performance of the closed-loop system $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ for any $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$, e.g., for a given $\delta > 0$, $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ is given such that

$$\|x(t)\|_{\mathcal{L}_2} \leq \delta \tag{5.10}$$

without depending on $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$. Then, it suffices that $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ approximates the given $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ to make the trajectory of the state variables of Σ in the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ close to that in the closed-loop system $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ for any $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$. In view of this, we address the following controller reduction problem:

Problem 5.1. *Given $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ such that (4.7) and (4.8), consider Σ in (5.4) with $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10). Give $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ in (5.5) and (5.7) such that it achieves a desirable $\mathcal{L}_2$-performance of the closed-loop system $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ for any $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$. Denote $\hat{x} \in \mathbb{R}^n$ by the state variables of Σ in the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$. Then, for a given constant $\epsilon > 0$, find $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ in (5.9) satisfying*

$$\|x(t) - \hat{x}(t)\|_{\mathcal{L}_2} \leq \epsilon \tag{5.11}$$

where $x(0) = \hat{x}(0) = x_0$ for all $x_0 \in \mathbb{R}^n$ such that $\|x_0\| = 1$ and $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.

It should be noted that the difficulties of this problem are that

- the criterion (5.11) is satisfied for any sets of locally stabilizing controllers in $\mathcal{K}_\theta$, and
- the approximant $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ has the hierarchical distributed structure in (5.9).

Since existing controller reduction techniques, e.g., [24, 25], cannot explicitly take into account the interconnection structure among controllers, Problem 5.1 cannot be solved by straightforwardly using those techniques. In the next section, we give a solution to Problem 5.1 by explicitly utilizing a hierarchically distributed structure of the closed-loop system.

5.3 Main Results

5.3.1 Analysis of Controller Reduction

In the rest of this chapter, for the sake of simple explanation, we focus on the case of $L = 3$, which yields $\mathcal{L} = \{1, 2, 3\}$. Similar results are also availble for general cases. In addition, we omit the subscript 1 of the matrices associated with $\Phi^{(3)}$ and $\hat{\Phi}^{(3)}$, e.g., $E^{(3)}$ denotes $E_1^{(3)}$.

Note that each compositional unit of $\Phi^{(2)}$ and $\Phi^{(3)}$, which are an $n_i^{(2)}$-dimensional system and an n-dimensional system respectively, are higher-dimensional than that of $\Phi^{(1)}$. In view of this, we consider reducing $\Phi^{(2)}$ and $\Phi^{(3)}$. In other words, we take $\hat{\Phi}^{(1)}$ as $\Phi^{(1)}$, i.e., the parameters of $\hat{\Phi}^{(1)}$ in (5.9) are taken as

$$\hat{\mathbf{E}}_i^{(1)} = E_i^{(1)}, \quad \hat{\mathbf{F}}_i^{(1)} = F_i^{(1)}, \quad \hat{\mathbf{\Gamma}}^{(1)} = \Gamma^{(1)}, \quad \hat{\mathbf{U}}_i = U_i, \quad \hat{\mathbf{\Lambda}}_i^{(1,k)} = \mathbf{\Lambda}_i^{(1)} \tag{5.12}$$

for each $i \in \mathcal{N}^{(1)}$ and $k \in \mathcal{L}$. Note that $\hat{\mathbf{G}}^{(2)}$ depends on also $\hat{\Phi}^{(2)}$.

Let us consider the following two transfer function matrices associated with $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ and $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ as

$$g(s) := [0, I_n](sI - \mathcal{A})^{-1}\mathcal{B}, \quad \hat{g}(s) := [0, I_n](sI - \hat{\mathcal{A}})^{-1}\hat{\mathcal{B}} \tag{5.13}$$

where

$$\mathcal{A} := \begin{bmatrix} E^{(3)} + \Lambda^{(3)}F^{(3)} & 0 & \Gamma^{(3)}\mathrm{dg}(C_i) \\ \Lambda^{(3)}F^{(3)} + \mathrm{dg}(\Lambda_i^{(2)})G^{(3)} & \mathrm{dg}(E_i^{(2)} + \Lambda_i^{(2)}F_i^{(2)}) & \Gamma^{(2)}\mathrm{dg}(C_i) \\ \Lambda^{(3)}F^{(3)} + \mathrm{dg}(\Lambda_i^{(2)})G^{(3)} & \mathrm{dg}(\Lambda_i^{(2)}F_i^{(2)}) + \mathrm{dg}(\Lambda_i^{(1)})G^{(2)} & A + \mathrm{dg}(\Lambda_i^{(1)}F_i^{(1)}) \end{bmatrix}$$

$$\hat{\mathcal{A}} := \begin{bmatrix} \hat{\mathbf{E}}^{(3)} + \hat{\mathbf{\Lambda}}^{(3,3)}\hat{\mathbf{F}}^{(3)} & 0 & \hat{\mathbf{\Gamma}}^{(3)}\mathrm{dg}(C_i) \\ \hat{\mathbf{\Lambda}}^{(2,3)}\hat{\mathbf{F}}^{(3)} + \mathrm{dg}(\hat{\mathbf{\Lambda}}_i^{(2,2)})\hat{\mathbf{G}}^{(3)} & \mathrm{dg}(\hat{\mathbf{E}}_i^{(2)} + \hat{\mathbf{\Lambda}}_i^{(2,2)}\hat{\mathbf{F}}_i^{(2)}) & \hat{\mathbf{\Gamma}}^{(2)}\mathrm{dg}(C_i) \\ \Lambda^{(3)}\hat{\mathbf{F}}^{(3)} + \mathrm{dg}(\Lambda_i^{(2)})\hat{\mathbf{G}}^{(3)} & \mathrm{dg}(\Lambda_i^{(2)}\mathbf{F}_i^{(2)}) + \mathrm{dg}(\Lambda_i^{(1)})\hat{\mathbf{G}}^{(2)} & A + \mathrm{dg}(\Lambda_i^{(1)}F_i^{(1)}) \end{bmatrix}$$

$$\mathcal{B} := \begin{bmatrix} \Gamma^{(3)}\mathrm{dg}(C_i) \\ \Gamma^{(2)}\mathrm{dg}(C_i) \\ \Gamma^{(1)}\mathrm{dg}(C_i) \end{bmatrix}, \quad \hat{\mathcal{B}} := \begin{bmatrix} \mathbf{\Gamma}^{(3)}\mathrm{dg}(C_i) \\ \mathbf{\Gamma}^{(2)}\mathrm{dg}(C_i) \\ \Gamma^{(1)}\mathrm{dg}(C_i) \end{bmatrix}. \tag{5.14}$$

We can see that $g(s)$ and $\hat{g}(s)$ do not include any information on local controllers $\{\kappa_i\}_{i\in\mathcal{N}}$. To guarantee (5.11) for any sets of local controllers in $\mathcal{K}_\theta$, it suffices that the approximation error of the closed-loop systems can be independently evaluated by using these systems, which do not include any information on local controllers, and a system associated with local controllers. Thus, we give the following theorem:

Theorem 5.2. *Consider Problem 5.1 and $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ in (5.9) and (5.12). Define $g(s)$ and $\hat{g}(s)$ in (5.13). If $\hat{\mathcal{A}}$ is stable, then*

$$\|x(t) - \hat{x}(t)\|_{\mathcal{L}_2} \le \|\theta\| \, \|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty} \tag{5.15}$$

where $x(0) = \hat{x}(0) = x_0$ for all $x_0 \in \mathbb{R}^n$ such that $\|x_0\| = 1$ and $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.

Proof. Let $\hat{\phi} := [(\hat{\phi}^{(3)})^\mathsf{T}, (\hat{\phi}^{(2)})^\mathsf{T}, (\phi^{(1)})^\mathsf{T}]^\mathsf{T}$. By taking a coordinate transformation as $\hat{\chi} = \hat{x} - \phi^{(1)}$, the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ is transformed into

$$\begin{bmatrix} \dot{\hat{\phi}} \\ \dot{\hat{\chi}} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} \hat{\mathcal{A}} & \hat{\mathcal{B}} & 0 \\ 0 & \mathrm{dg}(A_i) & \mathrm{dg}(B_iM_i) \\ 0 & \mathrm{dg}(H_iC_i) & \mathrm{dg}(K_i) \end{bmatrix} \begin{bmatrix} \hat{\phi} \\ \hat{\chi} \\ \xi \end{bmatrix}. \tag{5.16}$$

Similarly, by taking $\phi := [(\phi^{(3)})^\mathsf{T}, (\phi^{(2)})^\mathsf{T}, (\phi^{(1)})^\mathsf{T}]^\mathsf{T}$ and $\chi = x - \phi^{(1)}$, we have

$$\begin{bmatrix} \dot{\phi} \\ \dot{\chi} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} & 0 \\ 0 & \mathrm{dg}(A_i) & \mathrm{dg}(B_iM_i) \\ 0 & \mathrm{dg}(H_iC_i) & \mathrm{dg}(K_i) \end{bmatrix} \begin{bmatrix} \phi \\ \chi \\ \xi \end{bmatrix}. \tag{5.17}$$

Note that $\hat{\chi}(t) \equiv \chi(t)$ for all $t \ge 0$ because $\chi(0) = \hat{\chi}(0) = x_0$. Thus, we have $x - \hat{x} = [0, I_n]\phi - [0, I_n]\hat{\phi}$. Owing to Lemma 5.1, $g(s)$ is stable. Thus, $g(s) - \hat{g}(s)$ is also stable

if $\hat{\mathcal{A}}$ is stable. Hence, we have

$$\|x(t) - \hat{x}(t)\|_{\mathcal{L}_2} \leq \|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty} \|\chi(t)\|_{\mathcal{L}_2}.$$

By definition, we have $\|\chi(t)\|_{\mathcal{L}_2} \leq \theta$. Thus, the claim follows. □

Theorem 5.2 shows that we solve Problem 5.1 by finding $\hat{\Phi}^{(2)}$ and $\hat{\Phi}^{(3)}$ such that $\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty} < \epsilon/\theta$. On the basis of this theorem, we next clarify the relation between the approximation error of $\Phi^{(l)}$ and $\hat{\Phi}^{(l)}$, and that of $g(s)$ and $\hat{g}(s)$. More specifically, we take the biorthogonal projection [17], i.e., the parameters in (5.9) are taken as

$$\begin{array}{lll} \hat{\mathbf{E}}_i^{(l)} = P_i^{(l)} E_i^{(l)} Q_i^{(l)}, & \hat{\mathbf{F}}_i^{(l)} = F_i^{(l)} Q_i^{(l)}, & \hat{\mathbf{\Gamma}}^{(l)} = \mathrm{dg}(P_i^{(l)}) \Gamma^{(l)} \\ \hat{\mathbf{\Lambda}}_i^{(l,l)} = P_i^{(l)} \Lambda_i^{(l)} m & \hat{\mathbf{G}}^{(l)} = -\mathrm{dg}(F_i^{(l-1)}) \mathrm{dg}(Q_i^{(l)}), & \hat{\mathbf{\Lambda}}^{(2,3)} = \mathrm{dg}(P_i^{(2)}) \Lambda^{(3)} \end{array} \tag{5.18}$$

for $l \in \{2, 3\}$ where

$$P_i^{(l)} \in \mathbb{R}^{\hat{n}_i^{(l)} \times n_i^{(l)}}, \quad Q_i^{(l)} \in \mathbb{R}^{n_i^{(l)} \times \hat{n}_i^{(l)}}$$

satisfy $P_i^{(l)} Q_i^{(l)} = I_{\hat{n}_i^{(l)}}$. In this formulation, finding $\hat{\Phi}^{(l)}$ coincides with finding the biorthogonal projection described by $P_i^{(l)}$ and $Q_i^{(l)}$ in (5.18). Then, we give the following theorem:

Theorem 5.3. *Consider Problem 5.1 and $\{\hat{\Phi}^{(l)}\}_{l \in \mathcal{L}}$ in (5.9), (5.12) and (5.18). Define*

$$\sigma^{(l,k)} := \|(I - \mathrm{dg}(Q_i^{(l)} P_i^{(l)}))(sI - \mathrm{dg}(A_i^{(k)} + B_i^{(k)} F_i^{(k)}))^{-1} \mathrm{dg}(J_i^{(k)}) \mathrm{dg}(C_i)\|_{\mathcal{H}_\infty} \tag{5.19}$$

for each $k \geq l$ and $l \in \{2, 3\}$. Let $\mathcal{A}$ be given by (5.14) and define

$$\mathcal{P} := \mathrm{dg}(P^{(3)}, \mathrm{dg}(P_i^{(2)}), I_n), \quad \mathcal{Q} := \mathrm{dg}(Q^{(3)}, \mathrm{dg}(Q_i^{(2)}), I_n). \tag{5.20}$$

If $P_i^{(l)}$ and $Q_i^{(l)}$ satisfy that $\mathcal{PAQ}$ is stable, then

$$\|g(s) - \hat{g}(s)\|_{\mathcal{L}_2} \leq \sigma\mu \tag{5.21}$$

where $g(s)$ and $\hat{g}(s)$ are defined in (5.13) and

$$\begin{array}{l} \sigma := ((\sigma^{(2,3)} + \sigma^{(3,3)})(1 + \gamma^{(2)}) + \sigma^{(2,2)})(1 + \gamma^{(1)}) \\ \mu := \|[0, I_n](sI - \mathcal{PAQ})^{-1} \mathcal{PA}[I_{2n}, 0]^{\mathsf{T}}\|_{\mathcal{H}_\infty} \end{array} \tag{5.22}$$

with $\gamma^{(l)}$ in (5.8).

Proof. In what follows, we use the notation of

$$\Xi^{(l)} := \mathrm{dg}(\Xi_i^{(l)}), \quad \Xi_i^{(l)} := A_i^{(l)} + B_i^{(l)} F_i^{(l)}, \quad P^{(l)} := \mathrm{dg}(P_i^{(l)}), \quad Q^{(l)} := \mathrm{dg}(Q_i^{(l)}). \tag{5.23}$$

Consider the similarity transformation of the error system $g(s) - \hat{g}(s)$ by

$$\mathcal{T} = \begin{bmatrix} I_{\hat{n}^{(3)}+\hat{n}^{(2)}+n} & -\mathcal{P} \\ 0 & I_{3n} \end{bmatrix}, \ \mathcal{T}^{-1} = \begin{bmatrix} I_{\hat{n}^{(3)}+\hat{n}^{(2)}+n} & \mathcal{P} \\ 0 & I_{3n} \end{bmatrix}.$$

Then, we have

$$\begin{aligned} &\mathcal{T}\mathrm{dg}(\mathcal{P}\mathcal{A}\mathcal{Q}, \mathcal{A})\mathcal{T}^{-1} = \begin{bmatrix} \mathcal{P}\mathcal{A}\mathcal{Q} & -\mathcal{P}\mathcal{A}\overline{\mathcal{Q}\mathcal{P}} \\ 0 & \mathcal{A} \end{bmatrix}, \quad \mathcal{T}\begin{bmatrix} \mathcal{P}\mathcal{B} \\ \mathcal{B} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{B} \end{bmatrix} \\ &[-[0, I_n]\mathcal{Q}, \ [0, I_n]]\,\mathcal{T}^{-1} = [0, -I_n, 0_{n\times 3n}] \end{aligned}$$

where $0_{n\times m}$ denotes the n-by-m zero matrix and

$$\begin{aligned} \overline{\mathcal{P}} &:= \left[\mathrm{dg}(\overline{P}^{(3)}, \overline{P}^{(2)}), 0\right] \in \mathbb{R}^{(2n-\hat{n}^{(3)}-\hat{n}^{(2)})\times 3n} \\ \overline{\mathcal{Q}} &:= \left[\mathrm{dg}((\overline{Q}^{(3)})^{\mathsf{T}}, (\overline{Q}^{(2)})^{\mathsf{T}}), 0\right]^{\mathsf{T}} \in \mathbb{R}^{3n\times(2n-\hat{n}^{(3)}-\hat{n}^{(2)})} \end{aligned} \tag{5.24}$$

for $\overline{P}^{(l)} \in \mathbb{R}^{(n-\hat{n}^{(l)})\times n}$ and $\overline{Q}^{(l)} \in \mathbb{R}^{n\times(n-\hat{n}^{(l)})}$ such that $\overline{P}^{(l)}\overline{Q}^{(l)} = I_{n-\hat{n}^{(l)}}$. Thus, it follows that

$$g(s) - \hat{g}(s) = [0, -I_n]\mathcal{Q}(sI - \mathcal{P}\mathcal{A}\mathcal{Q})^{-1}\mathcal{P}\mathcal{A}\overline{\mathcal{Q}\mathcal{P}}(sI - \mathcal{A})^{-1}\mathcal{B}.$$

Noting that $\overline{\mathcal{Q}\mathcal{P}} = [I_{2n}, 0]^{\mathsf{T}}\overline{\mathcal{Q}\mathcal{P}}$, we have

$$\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty} \le \mu\|\overline{\mathcal{Q}\mathcal{P}}(sI - \mathcal{A})^{-1}\mathcal{B}\|_{\mathcal{H}_\infty}.$$

Furthermore, giving

$$T := \begin{bmatrix} I_n & 0 & 0 \\ I_n & I_n & 0 \\ I_n & I_n & I_n \end{bmatrix} \in \mathbb{R}^{3n\times 3n},$$

we have

$$\begin{aligned} &\overline{\mathcal{Q}\mathcal{P}}(sI - \mathcal{A})^{-1}\mathcal{B} = \overline{\mathcal{Q}\mathcal{P}}T(sI - T^{-1}\mathcal{A}T)^{-1}T^{-1}\mathcal{B} \\ &= \begin{bmatrix} \overline{Q}^{(3)}\overline{P}^{(3)} & 0 & 0 \\ \overline{Q}^{(2)}\overline{P}^{(2)} & \overline{Q}^{(2)}\overline{P}^{(2)} & 0 \\ 0 & 0 & 0 \end{bmatrix} \left(sI - \begin{bmatrix} \Xi^{(3)} & J^{(3)}S & J^{(3)}S \\ 0 & \mathrm{dg}(\Xi^{(2)}) & J^{(2)}S \\ 0 & 0 & \mathrm{dg}(\Xi^{(1)}) \end{bmatrix} \right)^{-1} \begin{bmatrix} J^{(3)}S \\ J^{(2)}S \\ J^{(1)}S \end{bmatrix} \end{aligned} \tag{5.25}$$

where

$$J^{(l)} := \mathrm{dg}(J_i^{(l)}), \quad S := \mathrm{dg}(S_i).$$

Hence, the claim follows. □

Theorem 5.3 shows that the approximation error $\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty}$ is bounded by $\sigma^{(l,k)}$ associated with the approximation error of $\Phi^{(l)}$ and $\hat{\Phi}^{(l)}$. Therefore, combining Theorem 5.2 and 5.3, we can construct a low-dimensional hierarchical distributed controller $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ to achieve an $\mathcal{L}_2$-performance of the closed-loop system for any sets of locally stabilizing controllers in $\mathcal{K}_\theta$. In Theorem 5.3, we assume that $P_i^{(l)}$ and $Q_i^{(l)}$ guarantee the stability of $\mathcal{PAQ}$. One approach to construct $P_i^{(l)}$ and $Q_i^{(l)}$ guaranteeing the stability is shown in Section 5.3.2 below.

5.3.2 Stability Preservation of Low-dimensional Hierarchical Distributed Controller

To construct $P^{(l)}$ and $Q^{(l)}$ in (5.18) preserving the stability of the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$, we first introduce the following lemma:

Lemma 5.4. *Consider Problem 5.1 and the autonomous system (5.28) with* $\tilde{x}^{(0)}(0) = x_0$ *and* $\tilde{x}^{(l)}(0) = 0$ *for* $l \in \mathcal{L}$*. If*

$$\mathrm{im}(J^{(3)}) \subseteq \mathrm{im}(Q^{(3)}), \quad \mathrm{im}([J^{(2)}, Q^{(3)}]) \subseteq \mathrm{im}(Q^{(2)}) \tag{5.26}$$

then it follows that

$$\hat{x}(t) \equiv \sum_{l=2}^{3} Q^{(l)}\tilde{x}^{(l)}(t) + \sum_{l=0}^{1} \tilde{x}^{(l)}(t). \tag{5.27}$$

Proof. Note that (5.26) implies

$$Q^{(3)}P^{(3)}J^{(3)} = J^{(3)}, \quad Q^{(2)}P^{(2)}[J^{(2)}, Q^{(3)}] = [J^{(2)}, Q^{(3)}].$$

and it follows that

$$\mathrm{dg}(A_i^{(l+1)}) = \mathrm{dg}(A_j^{(l)}) + J^{(l+1)}C$$

by definition of $J^{(l)}$. By taking a coordinate transformation given by

$$\tilde{x}^{(l)} = \begin{cases} -P^{(l)}Q^{(l+1)}\hat{\phi}^{(l+1)} + \hat{\phi}^{(l)}, & l \in \mathcal{L} \\ x - \hat{\phi}^{(1)}, & l = 0 \end{cases}$$

where $\hat{\phi}^{(4)}$ is regarded as zero, the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ is transformed into

$$\begin{bmatrix} \dot{\tilde{x}}^{(3)} \\ \dot{\tilde{x}}^{(2)} \\ \dot{\tilde{x}}^{(1)} \\ \dot{\tilde{x}}^{(0)} \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} P^{(3)}\Xi^{(3)}Q^{(3)} & P^{(3)}J^{(3)}SQ^{(2)} & P^{(3)}J^{(3)}S & P^{(3)}J^{(3)}S & 0 \\ P^{(2)}\overline{Q}^{(3)}\overline{P}^{(3)}\Xi^{(3)}Q^{(3)} & P^{(2)}\Xi^{(2)}Q^{(2)} & P^{(2)}J^{(2)}S & P^{(2)}J^{(2)}S & 0 \\ \overline{Q}^{(2)}\overline{P}^{(2)}\Xi^{(3)}Q^{(3)} & \overline{Q}^{(2)}\overline{P}^{(2)}\Xi^{(2)}Q^{(2)} & \Xi^{(1)} & J^{(1)}S & 0 \\ 0 & 0 & 0 & \mathrm{dg}(A_i) & \mathrm{dg}(B_iM_i) \\ 0 & 0 & 0 & \mathrm{dg}(H_iC_i) & \mathrm{dg}(K_i) \end{bmatrix} \begin{bmatrix} \tilde{x}^{(3)} \\ \tilde{x}^{(2)} \\ \tilde{x}^{(1)} \\ \tilde{x}^{(0)} \\ \xi \end{bmatrix}. \tag{5.28}$$

In addition, $\hat{x}$ can be described as (5.27). □

In Lemma 5.4, we can see from (5.28) that the state variables $\tilde{x}^{(1)}$ depend on the state variables in the higher layers, i.e., $\tilde{x}^{(2)}$ and $\tilde{x}^{(3)}$. If we do not reduce $\Phi^{(l)}$, i.e., $P^{(l)} = Q^{(l)} = I_n$, then there are no feedback from the variables in the higher layer to that in the lower layer. Thus, on the basis of small gain theorem [58], we consider constructing $P^{(l)}$ and $Q^{(l)}$ such that they make the magnitude of the feedback small. In view of this, we provide the following theorem:

Theorem 5.5. *Consider Problem 5.1. Let $\rho^{(3)} > 0$ and $\mathcal{V}^{(3)} \succ \mathcal{O}_n$ be given such that*

$$\mathcal{S}_{\rho^{(3)}}(\mathcal{V}^{(3)}; \Xi^{(3)}, J^{(3)}, \Xi^{(3)}) \prec 0_n \tag{5.29}$$

where

$$\mathcal{S}_\rho(\mathcal{V}; A, B, C) := A^\mathsf{T}\mathcal{V} + \mathcal{V}A + \rho^{-1}(\mathcal{V}BB^\mathsf{T}\mathcal{V} + C^\mathsf{T}C). \tag{5.30}$$

Let $V^{(3)}$ be a Cholesky factor of $\mathcal{V}^{(3)}$, i.e., $V^{(3)} \succ 0_n$ such that $\mathcal{V}^{(3)} = (V^{(3)})^\mathsf{T}V^{(3)}$. Define

$$P^{(3)} = W^{(3)}V^{(3)}, \quad Q^{(3)} = (V^{(3)})^{-1}(W^{(3)})^\mathsf{T} \tag{5.31}$$

for $W^{(3)} \in \mathbb{R}^{\hat{n}^{(3)}\times n}$ such that $W^{(3)}(W^{(3)})^\mathsf{T} = I_{\hat{n}^{(3)}}$ and the first condition in (5.26) holds. Furthermore, suppose that there exist $\rho^{(2)} \in \mathbb{R}^{|\mathcal{N}^{(2)}|}$ and $\mathcal{V}_i^{(2)} \succ 0_{n_i^{(2)}}$ satisfying

$$\mathcal{S}_{\rho_i^{(2)}}(\mathcal{V}_i^{(2)}; \Xi_i^{(2)}, E_i(I - Q^{(3)}P^{(3)}), S_i^{(2)}) \prec 0_{n_i^{(2)}}, \tag{5.32}$$

for $i \in \mathcal{N}^{(2)}$ and

$$\|\rho^{(2)}\| < (\rho^{(3)})^{-1} \tag{5.33}$$

where $\rho_i^{(2)}$ denotes the ith element of $\rho^{(2)}$, $E_i \in \mathbb{R}^{n_i^{(2)}\times n}$ such that $[E_1^\mathsf{T}, \ldots, E_{|\mathcal{N}^{(2)}|}^\mathsf{T}]^\mathsf{T} = I_n$ and $S_i^{(2)} := \mathrm{dg}(\mathrm{dg}(S_k)_{k\in\mathcal{C}_j^{(0)}})_{j\in\mathcal{C}_i^{(1)}}$. Furthermore, define

$$P_i^{(2)} = W_i^{(2)}V_i^{(2)}, \quad Q_i^{(2)} = (V_i^{(2)})^{-1}(W_i^{(2)})^\mathsf{T} \tag{5.34}$$

where $V_i^{(2)}$ *is a Cholesky factor of* $\mathcal{V}_i^{(2)}$ *and* $W_i^{(2)} \in \mathbb{R}^{\hat{n}_i^{(2)} \times n_i^{(2)}}$ *satisfies* $W_i^{(2)}(W_i^{(2)})^{\mathsf{T}} = I_{\hat{n}_i^{(2)}}$. *If* $P^{(2)}$ *and* $Q^{(2)}$ *satisfy the second condition in (5.26) and*

$$\|C(sI - \Xi^{(1)})^{-1}(I - Q^{(2)}P^{(2)})\|_{\mathcal{H}_\infty} < \|[\Xi^{(3)}Q^{(3)}, \Xi^{(2)}Q^{(2)}](sI - \mathcal{X})^{-1}\begin{bmatrix} P^{(3)}J^{(3)} \\ P^{(2)}J^{(2)} \end{bmatrix}\|_{\mathcal{H}_\infty}^{-1} \tag{5.35}$$

where $\mathcal{X}$ *denotes the principal submatrix of the system matrix in (5.28) corresponding to the first* $\hat{n}^{(3)} + \hat{n}^{(2)}$ *rows and columns, then the closed-loop system* $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ *is stable.*

Proof. From Lemma 5.4, it suffices that we show the stability of the principal submatrix of the system matrix in (5.28) corresponding to the first $\hat{n}^3 + \hat{n}^2 + n$ rows and columns. Let $\overline{P}^{(3)}$ and $\overline{Q}^{(3)}$ be given by

$$\overline{P}^{(3)} = \overline{W}^{(3)}V^{(3)}, \quad \overline{Q}^{(3)} = (V^{(3)})^{-1}(\overline{W}^{(3)})^{\mathsf{T}}$$

where $\overline{W}^{(3)} \in \mathbb{R}^{(n-\hat{n}^{(3)})\times n}$ such that $[(W^{(3)})^{\mathsf{T}}, (\overline{W}^{(3)})^{\mathsf{T}}]^{\mathsf{T}}$ is unitary. First, we show the stability of $\mathcal{X}$ in Theorem 5.3. From small gain theorem [58], it suffices to show

$$\|SQ^{(2)}(sI - P^{(2)}\Xi^{(2)}Q^{(2)})^{-1}P^{(2)}\overline{Q}^{(3)}\overline{P}^{(3)}\Xi^{(3)}Q^{(3)}(sI - P^{(3)}\Xi^{(3)}Q^{(3)})^{-1}P^{(3)}J^{(3)}\|_{\mathcal{H}_\infty} < 1. \tag{5.36}$$

Note that there exist $\rho^{(3)} > 0$ and $V^{(3)} \succ 0_n$ satisfying (5.29) because $\Xi^{(3)}$ is stable. It follows from Bounded Real Lemma [48] that (5.29) is equivalent to

$$\begin{bmatrix} \mathcal{S}_{\rho^{(3)}}(\mathcal{V}^{(3)}; \Xi^{(3)}, 0, \Xi^{(3)}) & J^{(3)} \\ * & -(\rho^{(3)})^{-1}I_p \end{bmatrix} \prec 0_{n+p}. \tag{5.37}$$

Pre- and Post- multiplication of (5.37) by $\mathrm{dg}((V^{(3)})^{-1}(W^{(3)})^{\mathsf{T}}, I_p)$ and $\mathrm{dg}(W^{(3)}(V^{(3)})^{-1}, I_p)$, we have

$$\begin{bmatrix} \mathcal{S}_{\rho^{(3)}}(I; P^{(3)}\Xi^{(3)}Q^{(3)}, 0, \Xi^{(3)}Q^{(3)}) & P^{(3)}J^{(3)} \\ * & -(\rho^{(3)})^{-1}I_p \end{bmatrix} \prec 0_{\hat{n}^{(3)}+p}.$$

Hence, $P^{(3)}\Xi^{(3)}Q^{(3)}$ is stable and

$$\|\Xi^{(3)}Q^{(3)}(sI - P^{(3)}\Xi^{(3)}Q^{(3)})^{-1}P^{(3)}J^{(3)}\|_{\mathcal{H}_\infty} < \rho^{(3)} \tag{5.38}$$

holds for all $W^{(3)}$ satisfying $W^{(3)}(W^{(3)})^{\mathsf{T}} = I_{\hat{n}^{(3)}}$. Similarly to this, (5.32) shows that $P_i^{(2)}\Xi_i^{(2)}Q_i^{(2)}$ is stable and

$$\|S_i^{(2)}Q_i^{(2)}(sI - P_i^{(2)}\Xi_i^{(2)}Q_i^{(2)})^{-1}P_i^{(2)}E_i^{\mathsf{T}}\overline{Q}^{(3)}\overline{P}^{(3)}\|_{\mathcal{H}_\infty} < \rho_i^{(2)} \tag{5.39}$$

holds for all $i \in \mathcal{N}^{(2)}$ and all sets of $\{W_i^{(2)}\}_{i\in\mathcal{N}^{(2)}}$ satisfying $W_i^{(2)}(W_i^{(2)})^{\mathsf{T}} = I_{\hat{n}_i^{(2)}}$. Note that

$$\left\| \begin{bmatrix} G_1(s) \\ G_2(s) \end{bmatrix} \right\|_{\mathcal{H}_\infty}^2 \leq \|G_1(s)\|_{\mathcal{H}_\infty}^2 + \|G_2(s)\|_{\mathcal{H}_\infty}^2$$

holds for stable proper transfer matrices $G_1(s)$ and $G_2(s)$. Thus, (5.39) yields

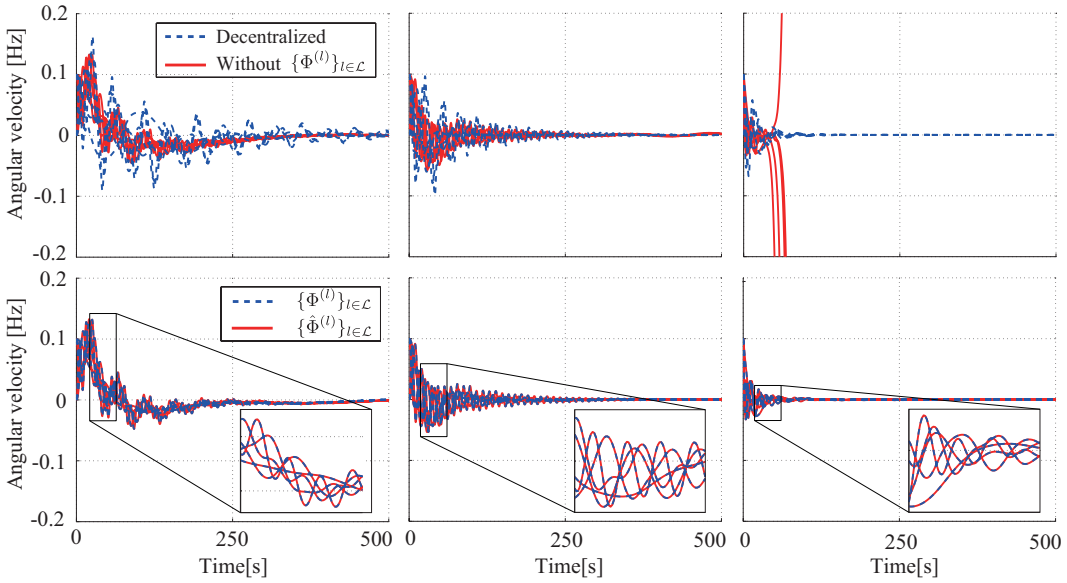
$$\|SQ^{(2)}(sI - P^{(2)}\Xi^{(2)}Q^{(2)})^{-1}P^{(2)}\overline{Q}^{(3)}\overline{P}^{(3)}\|_{\mathcal{H}_\infty}^2 \leq \|\rho^{(2)}\|^2$$

Hence, (5.36) follows from (5.33) and (5.38). Thus, it follows from small gain theorem that $\mathcal{X}$ is stable for all sets of $\{W_i^{(2)}\}_{i\in\mathcal{N}^{(2)}}$ satisfying $W_i^{(2)}(W_i^{(2)})^{\mathsf{T}} = I_{\hat{n}_i^{(2)}}$. Therefore, small gain theorem shows that (5.35) implies the stability of the closed-loop system $(\Sigma, \{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$. □

5.3.3 Design Algorithm of Low-dimensional Hierarchical Distributed Controller

In this subsection, we provide a design procedure of low-dimensional hierarchical distributed controllers. In general, it is difficult to find a biorthogonal projection to satisfy a criterion evaluated by the $\mathcal{H}_\infty$-norm, such as (5.19). Intuitively, the $\mathcal{H}_\infty$-norm of a transfer matrix is expected to be small if so is the $\mathcal{H}_2$-norm. In view of this, we consider constructing $P_i^{(l)}$ and $Q_i^{(l)}$ for a given $\hat{n}_i^{(l)}$ to make the $\mathcal{H}_2$-norm of the system in the left side of (5.19) small. The specific procedure is provided in Remark 2.4 in Chapter 3. For simplicity, we omit a procedure to preserve the stability of the closed-loop system. Note that, if an approximation error is sufficiently small, then $\mathcal{PAQ}$ is expected to be stable. Thus, we summarize a design procedure of $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ being a solution of Problem 5.1 as follows:

1. Give positive values of ϵ and θ.
2. For a given system Σ in (5.4) and a hierarchical structure $\{\mathcal{N}^{(l)}\}_{l\in\mathcal{L}}$ and $\{\mathcal{C}_i^{(l)}\}_{i\in\mathcal{N}^{(l+1)}}$ for $l \in \{0, 1, 2\}$, construct $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ in (5.5) and (5.7) such that it achieves a desirable $\mathcal{L}_2$-performance of the closed-loop system $(\Sigma, \{\Phi^{(l)}\}_{l\in\mathcal{L}}, \{\kappa_i\}_{i\in\mathcal{N}})$ for any $\{\kappa_i\}_{i\in\mathcal{N}} \in \mathcal{K}_\theta$.
3. Find $P_i^{(l)}$ and $Q_i^{(l)}$ satisfying $P_i^{(l)}Q_i^{(l)} = I_{\hat{n}_i^{(l)}}$ and (5.19) for a given $\hat{n}_i^{(l)}$.
4. If $\hat{g}(s)$ is unstable or $\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty} > \epsilon/\theta$, then take a larger value of $\hat{n}_i^{(l)}$ and go to the step in 3).
5. Construct $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ in (5.9), (5.12) and (5.18).



(a) Low-gain local controllers (b) Middle-gain local controllers (c) High-gain local controllers

FIGURE 5.1: Initial value responses of a power network model.

5.4 Numerical Example

In this section, we demonstrate the efficiency of the proposed low-dimensional hierarchical distributed control through an example of power network systems. We deal with a power network model provided in Section 4.4.1 in the previous chapter. We summarize the different settings below:

We deal with a power network system composed of $N = 50$ subsystems and each subsystem consists of three generators and two loads. Each generator and load are three- and two-dimensional system, respectively. Thus, each subsystem is 13-dimensional, i.e., $n_i = 13$ for $i \in \mathcal{N} := \{1, \ldots, 50\}$, and the overall power network system is 650-dimensional, i.e., $n = 650$. The interconnection structure among generators and loads in each subsystem is given as a graph Laplacian of complex network model, called the Wattz-Strogatz(WS) model [59]. For the interconnection among subsystems, the first generators in individual subsystems are interconnected and the graph Laplacian is given as WS model.

Furthermore, the elements of the admittance matrix compatible with interconnection among subsystems (resp. inside the subsystems) are randomly chosen from $[0.1, 0.5]$ (resp. $[0.1, 1.0]$).

First, we design a set of locally stabilizing controllers $\{\kappa_i\}_{i\in\mathcal{N}}$ in (4.10) by the LQR design techniques. We obtain three sets of locally stabilizing controllers where the resultant values of $\|\theta\|$ in (4.5) are 4100, 686 and 415, respectively. In the upper half of FIGURE. 5.1, we show the initial responses of angular velocities of all generators and

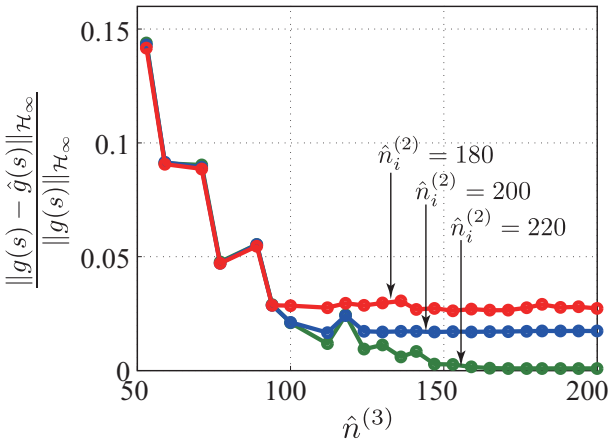


FIGURE 5.2: Resultant approximation error $\|g-\hat{g}\|_{\mathcal{H}_\infty}/\|g\|_{\mathcal{H}_\infty}$ versus $\hat{n}_i^{(l)}$.

loads in the first subsystem. In this figure, we depict the trajectories of disjoint subsystems and those of interconnected subsystems by blue and red lines, respectively. We can see from this figure that the instability of the closed-loop system is induced as we get higher-gain local controllers due to negative interference among subsystems.

Next, using a hierarchical distributed controller $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$, we consider improving the $\mathcal{L}_2$-performance of the whole closed-loop system. For the hierarchical structure, we take $\mathcal{N}^{(l)}$ such that

$$|\mathcal{N}^{(1)}| = 10, \quad |\mathcal{N}^{(2)}| = 2, \quad |\mathcal{N}^{(3)}| = 1$$

and take cluster sets $\mathcal{C}_i^{(l)}$ having the same number of subsystems in each layer, i.e.,

$$|\mathcal{C}_i^{(0)}| = 5, \quad |\mathcal{C}_i^{(1)}| = 5, \quad |\mathcal{C}_i^{(2)}| = 2$$

for $i \in \mathcal{N}^{(l)}$. Thus, we have $n_i^{(1)} = 65$, $n_i^{(2)} = 325$ and $n^{(3)} = 650$ for $i \in \mathcal{N}^{(l)}$. In addition, let the number of generators having additional input ports compatible with $B^{(l)}$ be $m^{(1)} = 25$, $m^{(2)} = 17$ and $m^{(3)} = 13$ and take those generators randomly from the first generators among 50 subsystems. We design individual controllers $\Phi^{(l)}$ by minimizing $\gamma^{(l)}$ in (5.8).

In the lower half of (a)-(c) in FIGURE. 5.1, we depict the initial value responses of the closed-loop system with a hierarchical distributed controller $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ by the blue dotted lines. From this figure, we see that the $\mathcal{L}_2$-performance of the closed-loop system improves as improving the performance of local controllers, owing to the attenuation of interference among subsystems. However, each compositional unit of the designed controller $\Phi^{(2)}$ and $\Phi^{(3)}$ are 325- and 650-dimensional system, respectively. Next, we aim at reducing these two controllers while preserving a similar quality.

We design $\hat{\Phi}^{(l)}$ along the procedure shown in Section 5.3.3 for several given values of $\hat{n}_i^{(l)}$. Let $\hat{n}_i^{(2)} = 180$, 200 and 220 for $i \in \mathcal{N}^{(2)}$. In FIGURE. 5.2, we plot the resultant approximation error of $\{\Phi^{(l)}\}_{l\in\mathcal{L}}$ and $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$, i.e., $\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty}/\|g(s)\|_{\mathcal{H}_\infty}$ where $g(s)$ and $\hat{g}(s)$ are defined in (5.13), with respect to several $\hat{n}^{(3)}$ for each case of $\hat{n}_i^{(2)}$ by the red, blue and green lines, respectively. From this figure, we can see that the performance of $\hat{\Phi}^{(l)}$ appropriately improves as increasing the dimension of $\hat{\Phi}^{(2)}$ and $\hat{\Phi}^{(3)}$. If we choose the dimension of $\hat{\Phi}^{(l)}$ as $\hat{n}_i^{(2)} = 200$ and $\hat{n}_i^{(3)} = 137$, the resultant approximation error turns out to be $\|g(s) - \hat{g}(s)\|_{\mathcal{H}_\infty}/\|g(s)\|_{\mathcal{H}_\infty} = 0.017$.

Finally, in the lower half of (a)-(c) in FIGURE. 5.1, we plot the initial value responses of the closed-loop system with $\{\hat{\Phi}^{(l)}\}_{l\in\mathcal{L}}$ by the red solid lines. We can see from this figure that the trajectories compatible with $\hat{\Phi}^{(l)}$ is close to that compatible with $\Phi^{(l)}$. Furthermore, for each case of (a)-(c), the resultant value of $\sup_{x_0}(\|x\|_{\mathcal{L}_2}/\|x_0\|_2)$ is 3568, 477 and 229, respectively. This result implies that the $\mathcal{L}_2$-performance of the closed-loop system with $\{\hat{\Phi}_l\}_{l\in\mathcal{L}}$ improves as improving the performance of local controllers.

5.5 Chapter Summary

In this chapter, we have proposed a design method of low-dimensional hierarchical distributed controllers for large-scale network systems. The problem of designing low-dimensional hierarchical distributed controllers has been formulated as a structured controller reduction problem for any sets of locally stabilizing controllers. To solve this problem, explicitly utilizing a hierarchically distributed structure of the closed-loop system, we have shown that the approximation error of the closed-loop system can be evaluated by that of the system associated with the hierarchical distributed controller without local controllers. Furthermore, we have derived the relation between the approximation errors of the individual hierarchical controllers and the performance degradation of the closed-loop system. Finally, we have shown the efficiency of the proposed method through a numerical example of power network systems.

Chapter 6

Low-dimensional Nonlinear Modelling of Plasticization Cylinders

6.1 Introduction

In this chapter, as a first step towards low-dimensional observer/controller design for nonlinear large-scale network systems, we show the importance and necessity of nonlinear low-dimensional modelling through an example of real industrial applications. More specifically, we construct a low-dimensional nonlinear model of the plasticization cylinder that is a key component of plastic injection molding machines. FIGURE. 6.1 shows schematic depiction of the plasticization cylinder that plays an important role in plasticization process where resin is melted by heat exchange with the internal surface of a barrel heated by band heaters. The quality of plastic products highly depends on temperatures in the plastic injection molding process. Thus, towards quality management and improvement of plastic products, modelling temperature dynamics of the plasticization cylinders in FIGURE. 6.1 has tremendous potential.

In the first half of this chapter, we construct the nonlinear model of plasticization cylinders including thermal properties of heaters, radiation to a water-cooling cylinder and outer air while taking into account the temperature-dependent properties of heaters. Furthermore, we show that the temperature dependency is not negligible and show the validity of the resultant model by experiment. Due to the high dimensionality of the resultant model (PDEs or a discretized model), simulation and controller design based on the model do not become computationally friendly. Thus, in the second half of this chapter, we consider to reduce the dimension of nonlinear models. More specifically,

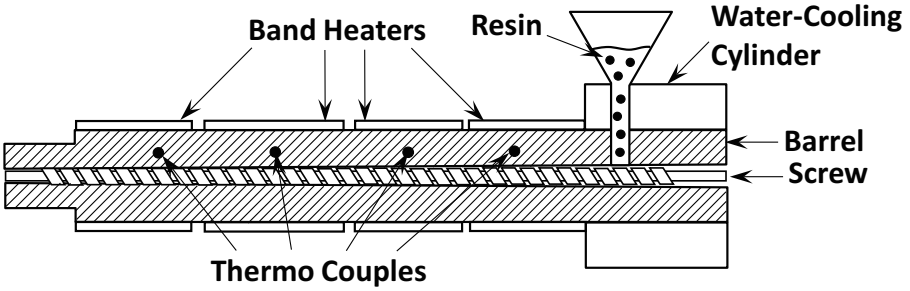


FIGURE 6.1: Plasticization cylinders.

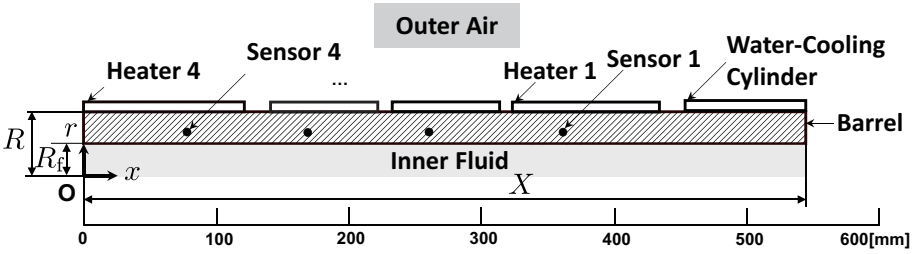


FIGURE 6.2: Schematic of plasticization cylinder models.

utilizing a particular structure of the nonlinear model arising from radiation to the air, we show that the reduced nonlinear model preserves the stability with an a priori approximation error bound.

This chapter is organized as follows: We provide a nonlinear model of plasticization cylinders in Section 6.2. In Section 6.3, we validate the resultant nonlinear model by experiment. In Section 6.4.1, we provide the theoretical approximation error bound of reduced order nonlinear models with the provision of systematic reduction procedure. In Section 6.4.2, we show the validity of the resultant 28-dimensional model (via an 808-dimensional spatially discretized model) by numerical simulation. Finally, in Section 6.5, we show concluding remarks.

6.2 Nonlinear Modelling of Plasticization Cylinders

In FIGURE. 6.2, we show the schematic depiction of plasticization model composed of a barrel, outer air, inner fluid, heaters and a water-cooling cylinder. The inner fluid represents melted resin inside the barrel in real injection machines. The water-cooling cylinder has internally a pipe line over the barrel, which cools the barrel by flowing water from IN to OUT shown in FIGURE. 6.3. In what follows, we describe the detail of the plasticization cylinder model.

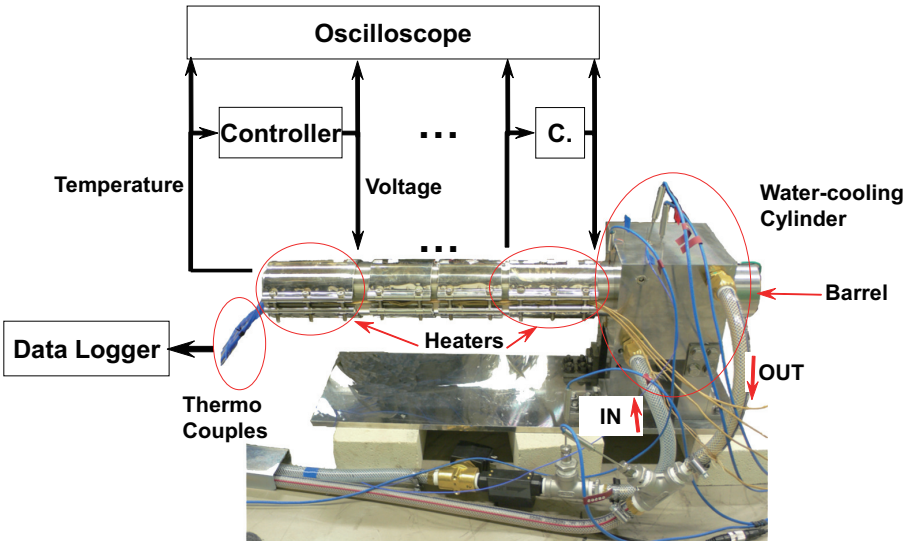


FIGURE 6.3: Prototype systems.

6.2.1 Barrel, Outer air, Inner fluid and Water-cooling cylinder

In this subsection, we model a barrel, outer air, inner fluid and a water-cooling cylinder. In what follows, we denote the time variable by t and spatial variables along the longer and radial direction by x and r such that

$$(x, r) \in \mathcal{D} := [0, X] \times [R_{\mathrm{f}}, R]$$

with the origin shown in FIGURE. 6.2 where X [m] is the length of the barrel, R_{f} and R [m] are the internal radius and the external radius of the barrel, respectively. In this notation, state variables are

- Temperature of barrel [deg]: $T(t, x, r)$
- Temperature of inner fluid [deg]: $\tilde{T}(t, x)$
- Temperature of the kth heater [deg]: $H_k(t), \ \ k \in \{1, \ldots, N\}$

where N denotes the number of heaters. Note that $\tilde{T}$ (resp. H_k) does not depend on r (resp. x and r). In addition, preliminary experiment shows that the temperature of outer air and that of a water-cooling cylinder can be regarded constant. Thus, we denote the temperature of outer air and that of a water-cooling cylinder by T_{o} [deg] and T_{c} [deg], respectively.

First, the heat transfer property of barrels is described by a cylindrical coordinate diffusion equation[1] as

$$\frac{\partial T}{\partial t} = \alpha \left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial x^2} \right), \quad (x, r) \in \mathrm{int}(\mathcal{D}), \tag{6.1}$$

where α [m^2/s] denotes a diffusion coefficient. In addition, the heat budget on the barrel surface is given by Neumann type boundary conditions

$$-\beta \frac{\partial T}{\partial \mathbf{n}} = h_\bullet (T_\bullet - T), \quad (x, r) \in \mathcal{S}_\bullet, \ \bullet = \{\mathrm{o}, \mathrm{c}\} \tag{6.2}$$

where β [W/(mK)] is the coefficient of thermal conductivity, $h_\bullet$ [W/(m^{2}K)] is the coefficient of heat transfer and $\mathcal{S}_\mathrm{o}, \mathcal{S}_\mathrm{c} \subset \partial\mathcal{D}$ are sets of contacts to the outer air and the water-cooling cylinder, respectively. In addition, the boundary condition on the inner fluid interface over $r = R_\mathrm{f}$ is given by

$$-\beta \frac{\partial T}{\partial r} = h_\mathrm{f} (\tilde{T} - T), \quad r = R_\mathrm{f}, \ x \in (0, X), \tag{6.3}$$

where h_f [W/(m^{2}K)] denotes a coefficient of heat transfer. For $k \in \{1, \ldots, N\}$, we describe the heat budget to the kth heater as

$$-\beta \frac{\partial T}{\partial r} = h_k (H_k - T), \quad r = R, \ x \in \mathcal{X}_k \tag{6.4}$$

where $\mathcal{X}_k \subset [0, X]$ is a set of contacts to the kth heater over $r = R$ and h_k [W/(m^{2}K)] is a coefficient of heat transfer at the contact points. Moreover, assuming that the inner fluid remains stationary, we describe the heat transfer dynamics of the inner fluid as

$$\frac{\partial \tilde{T}}{\partial t} = \tilde{\alpha} \frac{\partial^2 \tilde{T}}{\partial x^2} + h_\mathrm{f} (T(x, R_\mathrm{f}) - \tilde{T}), \quad x \in (0, X) \tag{6.5}$$

where $\tilde{\alpha}$ [m^2/s] is a coefficient of thermal conductivity of the inner fluid, the second term in the right-hand side in (6.5), i.e., $h_\mathrm{f}(T(x, R_\mathrm{f}) - \tilde{T})$, represents the total thermal flow from the barrel to the inner fluid. Moreover, we describe heat exchange at $x \in \{0, X\}$ of the outer air as

$$-\beta_\mathrm{f} \frac{\partial \tilde{T}}{\partial x} = h_\mathrm{o} (T_\mathrm{o} - \tilde{T}), \quad x \in \{0, X\}. \tag{6.6}$$

where β_f [W/(mK)] is a coefficient of thermal conductivity of the inner fluid. Finally, for output equations, $Y_\mathrm{d} \in \mathbb{R}^N$ [deg] and $Y \in \mathbb{R}^l$ [deg] denote temperatures measured by N sensors shown in FIGURE. 6.2 (called controlling thermo couples) and by other sensors (for measurement). The each element of Y_d and Y is given by appropriately spatially-weighted integration of T.

[1] $\partial\mathcal{D}$ is the border of $\mathcal{D}$, $\mathrm{int}(\mathcal{D})$ is the inside of $\mathcal{D}$, $\mathbf{n}$ denotes a normal unit vector to $\partial\mathcal{D}$.

6.2.2 Modelling heaters

In this subsection, we describe a model of heaters. In general, the coefficient of heat transfer between heaters and outer air is not static. This is because the rate of heat loss to the outer air depends on the temperature differences between heaters and the outer air under natural convection [8]. Toward quality management of plastic products, this temperature-dependency is not negligible and incorporated explicitly as follows:

Assumption 6.1. *For $k \in \{1, \ldots, N\}$, the heat transfer coefficient between the kth heater and outer air is given by*

$$\bar{h}_k(H_k - T_\mathrm{o}) \tag{6.7}$$

where the smooth function $\bar{h}_k : \mathbb{R} \mapsto \mathbb{R}_+$ is non-decreasing in $\mathbb{R}_+$ and non-increasing in $\mathbb{R}_-$.

For $k \in \{1, \ldots, N\}$, we describe the model of the kth heater as

$$\dot{H}_k = \frac{1}{c_k}\left(\frac{V_k^2(t)}{r_k} - \bar{h}_k(H_k - T_\mathrm{o}) \cdot a_k(H_k - T_\mathrm{o}) - 2\pi R \int_{\mathcal{X}_k} h_k(H_k - T(x, R))dx\right) \tag{6.8}$$

where c_k [J/K], r_k [ohm] and a_k [m^2] denote the heat capacity, the impedance and the outer area of the kth heater, respectively. Then, $2\pi R|\mathcal{X}_k|$ [m^2] coincides with the inner area of the kth heater and $V_k(t)$ [V] is the signal of input voltage. In the right side of (6.8), the second and third term represent the total thermal flow rate to outer air and the barrel, respectively. Note that the other coefficients of heat transfer $h_\mathrm{o}, h_\mathrm{c}, h_\mathrm{f}$ and h_k are constant because the corresponding temperatures are low or the corresponding area is sufficiently small.

6.2.3 Network System Given by Spatial discretization

In this subsection, spatialy discretizing the overall model described by partial differential equations(PDEs), we give a finite-dimensional nonlinear error system around an equilibrium state. More specifically, we discretize (6.1)-(6.6) with steps Δx and Δr for x and r axes by means of the finite element method [7]. Combining the discretized model

with (6.7) and (6.8), we have the following nonlinear thermal diffusion network system:

$$\begin{cases} \dot{\bar{x}} &= A\bar{x} + B_v\bar{v} + \bar{b}_1 \\ \bar{y}_{\mathrm{d}} &= C_{\mathrm{d}}\bar{x} \\ \bar{y} &= C_y\bar{x} \\ \bar{w} &= C_w\bar{x} + D_w\bar{u} \end{cases} \tag{6.9}$$

$$\begin{cases} \dot{\bar{z}} &= \begin{bmatrix} -\frac{1}{c_1}\bar{z}_1\psi_1(\bar{z}_1) \\ \vdots \\ -\frac{1}{c_N}\bar{z}_N\psi_N(\bar{z}_N) \end{bmatrix} + \bar{b}_2 + \bar{w} \\ \bar{v} &= \bar{z} + T_{\mathrm{o}}\mathbf{1}_N \end{cases} \tag{6.10}$$

$$\psi_k(\bar{z}_k) := a_k\bar{h}_k(\bar{z}_k) + 2\pi R|\mathcal{X}_k|h_k. \tag{6.11}$$

The system in (6.9) represents the temperature dynamics of a barrel and inner fluid while that in (6.10) and (6.11) represents the temperature dynamics of heaters. The physical meanings of variables are as follows: $\bar{x} \in \mathbb{R}^n$ is a vector of spatial discretized temperature T and $\tilde{T}$, $\bar{y}_{\mathrm{d}} \in \mathbb{R}^N$ denotes partial temperatures of T measured by sensors, $\bar{y} \in \mathbb{R}^l$ denotes partial temperatures of T for evaluation and

$$\bar{z} := [H_1 - T_{\mathrm{o}}, \ldots, H_N - T_{\mathrm{o}}]^{\mathsf{T}} \in \mathbb{R}^N, \quad \bar{v} := [H_1, \ldots, H_N]^{\mathsf{T}} \in \mathbb{R}^N.$$

Furthermore, $\bar{w} \in \mathbb{R}^N$ denotes a heat quantity to heaters and $\bar{u}(t)$ is input signal such that

$$\bar{u}(t) = [\bar{u}_1(t), \ldots, \bar{u}_N(t)]^{\mathsf{T}} \in \mathbb{R}^N, \quad \bar{u}_k(t) := V_k^2(t).$$

Finally, the first and second term in (6.11) represent radiation to the air and a barrel, respectively. Thus, ψ_k represents the coefficient of whole radiation of the kth heater.

The configuration of controller of injection molding machines is as follows: Input to the kth heater is determined by the compatible with measured temperature y_{d}, i.e., the kth element of y_{d}. Thus, operators give the target value of controllers as a desired output temperature y_{d}^* of y_{d}, which depends on molding products. In general, nonlinear systems do not necessarily have a unique desired state corresponding to a desired output. However, the following proposition guarantees that the nonlinear model in (6.9)-(6.11) has a unique equilibrium points.

Proposition 6.1. *Consider a nonlinear system (6.9)-(6.11). If* $\begin{bmatrix} A & B_v \\ C_{\mathrm{d}} & 0 \end{bmatrix}$ *is non-singular, then, for any* $y_{\mathrm{d}}^* \in \mathbb{R}^N$ *there exist unique equilibrium states and inputs* $\bar{x}^*, \bar{v}^*, \bar{u}^*$ *satisfying* $\bar{y}_{\mathrm{d}}(t) \equiv y_{\mathrm{d}}^*$.

Proof. Suppose $\dot{\bar{x}} \equiv 0$, $\bar{y}_{\rm d} \equiv y_d^*$. It follows from the first and second equations in (6.9) that

$$\begin{bmatrix} -\bar{b}_1 \\ y_{\rm d}^* \end{bmatrix} = \begin{bmatrix} A & B_v \\ C_{\rm d} & 0 \end{bmatrix} \begin{bmatrix} \bar{x}^* \\ \bar{v}^* \end{bmatrix}. \tag{6.12}$$

Note that $\begin{bmatrix} A & B_v \\ C_{\rm d} & 0 \end{bmatrix}$ is non-singular. Hence, there exist unique $\bar{x}^*$ and $\bar{v}^*$ satisfying (6.12). Let $z^* := \bar{v}^* - T_o \mathbf{1}_N$ and $D_w := \mathrm{dg}\,(1/(c_1 r_1), \ldots, 1/(c_N r_N))$. Then, $\bar{u}^*$ is given by

$$\bar{u}^* = -D_w^{-1} \left(\begin{bmatrix} -z_1^* \psi_k(z_1^*) \\ \vdots \\ -z_N^* \psi_k(z_N^*) \end{bmatrix} + \bar{b}_2 + C_w \bar{x}^* \right).$$

□

In the remainder of this section, we rewrite (6.9)-(6.11) as an error system from the desired value. Define errors with respect to each variable, e.g., $z := \bar{z} - z^*$, $y := \bar{y} - C_y \bar{x}^*$. From simple calculation, we have

$$\bar{z} = \begin{bmatrix} -(z_1 + z_1^*)\psi_k(z_1 + z_1^*) \\ \vdots \\ -(z_N + z_N^*)\psi_k(z_N + z_N^*) \end{bmatrix} + C_w x + D_w u + \bar{b}_2 + C_w \bar{x}^* + D_w \bar{u}^*.$$

Thus, we have a nonlinear network system

$$\begin{cases} \dot{x} &= Ax + B_v z \\ y_{\rm d} &= C_{\rm d} x \\ y &= C_y x \\ w &= C_w x + D_w u \end{cases} \tag{6.13}$$

$$\Sigma_{nl} : \dot{z} = \tilde{\Psi}(z) + w, \quad v = z \tag{6.14}$$

where

$$\tilde{\Psi}(z) := \left[\tilde{\psi}_k(z_1), \ldots, \tilde{\psi}_k(z_N) \right]^{\mathsf{T}} \in \mathbb{R}^N, \quad \tilde{\Psi}(0) = 0 \tag{6.15}$$

$$\tilde{\psi}_k(z_k) := -\left((z_k + z_k^*)\psi_k(z_k + z_k^*) - z_k^* \psi_k(z_k^*) \right) \tag{6.16}$$

and z_k denotes the kth element of z for any $k \in \{1, \ldots, N\}$.

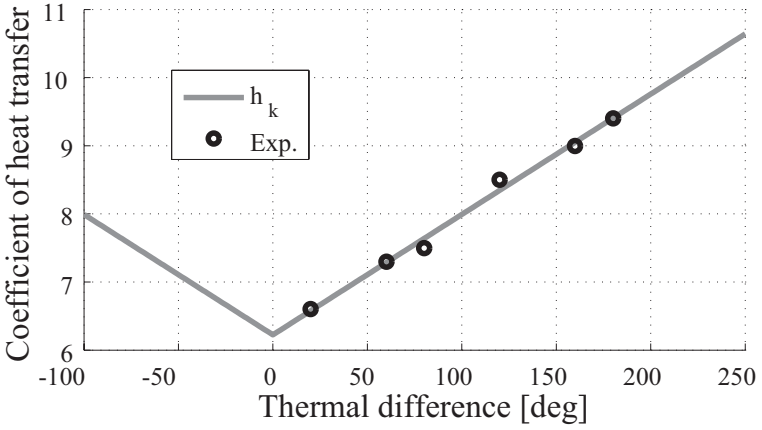


FIGURE 6.4: Coefficients of heat transfer $\bar{h}_k$.

6.3 Experimental Results

6.3.1 Configuration of Real Systems

In this section, we show the validity of the nonlinear network model in (6.13)-(6.16) by experiment. We first describe the configurations of the prototype system shown in FIGURE. 6.3 as follows: The prototype system consists of a barrel, $N = 4$ heaters and a water-cooling cylinder. The cylindrical barrel has several thermo couples on the interior wall of the barrel as well as four implanted several thermo couples (controlling thermo couples) inside the barrel. In addition, sequential experimental data is sampled by a data logger during 30 [min] with a sufficiently short sampling interval. The velocity of water flowing through a water-cooling cylinder is constant because preliminary experiments show that the velocity does not have influences on variance of barrel temperatures.

Furthermore, the nonlinear function of heat transfer $\bar{h}_k(\bar{z}_k)$ in (6.7) is determined as follows: In FIGURE. 6.4, we depict several constant coefficients of the heat transfer identified by experiment for the first heater as circles. In addition, $\bar{h}_1(\bar{z}_1)$ for $\bar{z}_1 \in \mathbb{R}_+$ is determined by the least square method in $\mathbb{R}_+$ for the resultant experimental data (black circles in FIGURE. 6.4). Furthermore, we take $\bar{h}_1(-\bar{z}_1) = \bar{h}_1(\bar{z}_1)$ and the resultant $\bar{h}_1(\bar{z}_1)$ is plotted by the glay line in FIGURE. 6.4. We take $\bar{h}_k = \bar{h}_1$ for $k \in \{2, 3, 4\}$ because similar experimental results are obtained. The other parameters, e.g., α in (6.1), $h_{\rm o}$ and $h_{\rm c}$ in (6.2), are referred to [8]. Taking $\Delta x = \Delta r = 5$ [mm], we have an 808-dimensional nonlinear thermal diffusion network system.

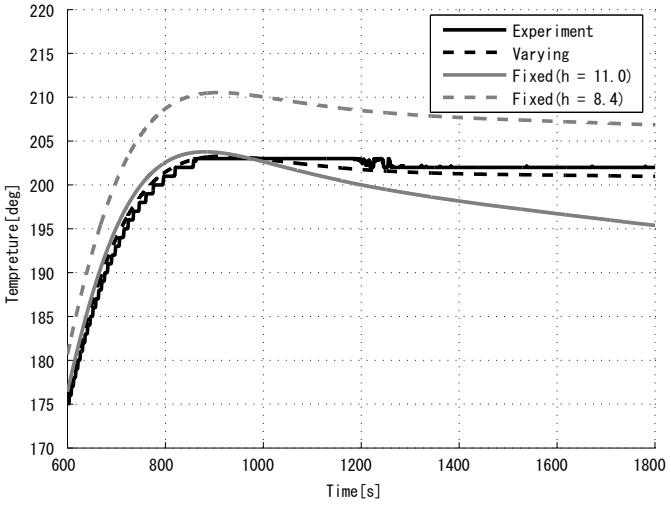


FIGURE 6.5: Temperature-dependency of heat transfer coefficient.

6.3.2 Model Validation by Experiment

In what follows, we validate the nonlinear network model in (6.13)-(6.16) by comparing experimental results and simulation results where measured experimental input signals are used. In FIGURE. 6.5, the black solid line depicts the transient temperature measured by a controlling thermo couple. In addition, glay solid and dotted line depict the temperature by simulation of the obtained model (6.13) and (6.14) with $\bar{h} \equiv 11.0$ and $\bar{h} \equiv 8.4$. This result show that glay solid (resp. dotted) line is below (resp. above) the experimental result plotted by the black solid line around $t = 1800$ (resp. $t = 1000$). This is because the low (resp. high) constant value of $\bar{h}$ expresses small (resp. high) heat radiation to the outer air. Furthermore, in FIGURE. 6.5, the black dotted line depicts the temperature by simulation with the nonlinear function of $\bar{h}$ given in the Section 6.3.1. This figure shows that the temperature-dependency of $\bar{h}$ is not negligible and the nonlinear network model in (6.13) and (6.14) has potential to approximate experimental results.

Next, we validate steady-state characteristics of the nonlinear network model in (6.13)-(6.16). In FIGURE. 6.6, we plot the steady temperature distribution inside the barrel obtained by the experiment as circles for the desired temperatures 200 [deg] and 100 [deg], respectively where the two results are plotted in each case. Moreover, the solid lines are simulation results of $\bar{x}^*$ in Proposition 6.1. In FIGURE. 6.6, the vertical axis indicates temperatures and the horizontal axis indicates positions inside the barrel in FIGURE. 6.2. From this figure, we can see that the obtained nonlinear network model accurately explains experimental results.

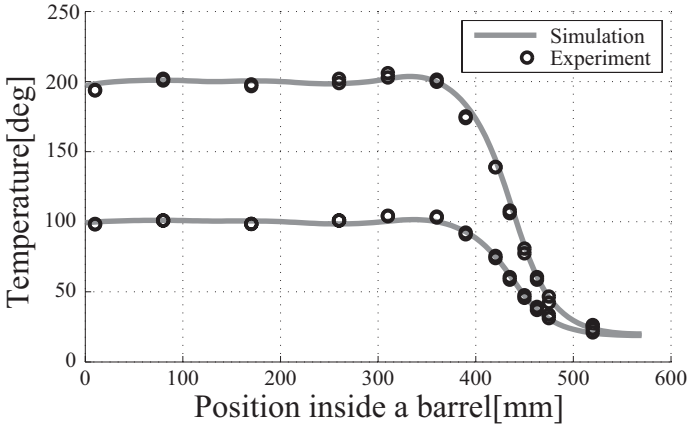


FIGURE 6.6: Steady temperature distribution inside barrel.

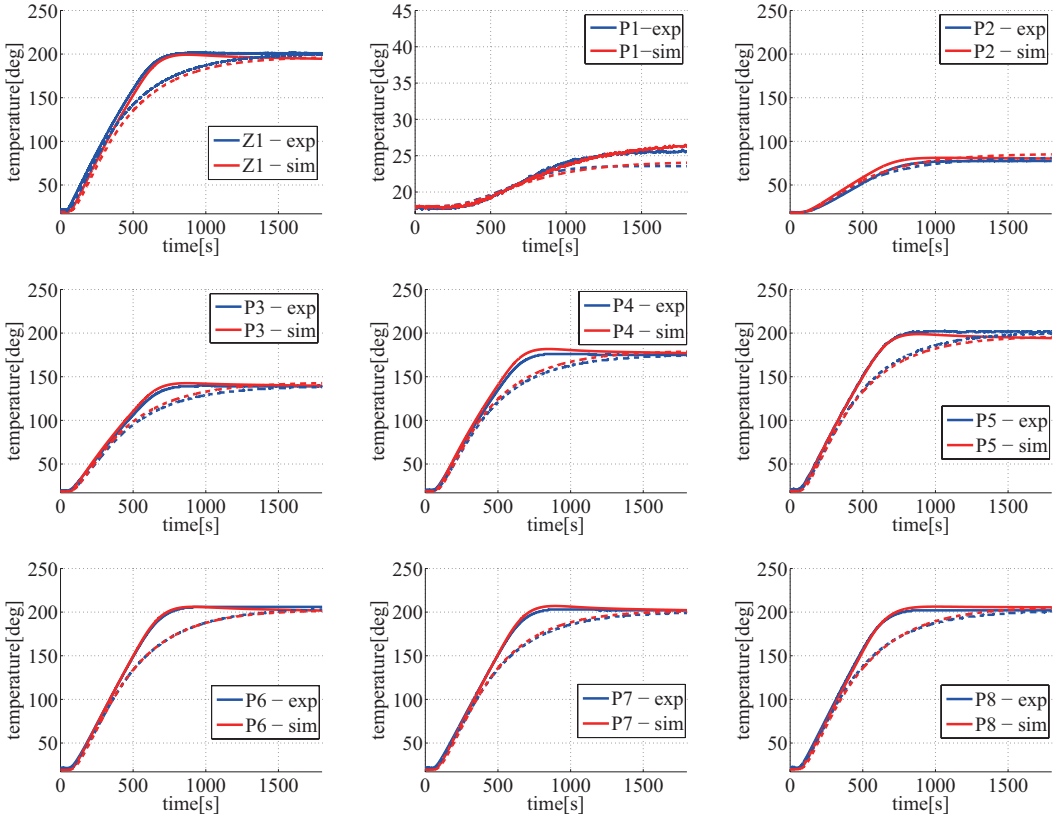


FIGURE 6.7: Transient responses at several measurement points shown in FIGURE. 6.8.

Finally, we show transient responses obtained by the experiment and simulation as the blue and red lines in FIGURE. 6.7. In addition, the solid and dotted lines depict those for different controlling configurations. Measurement points are shown in FIGURE. 6.8 as $Z1$ and $P1$-$P8$. These figures show that the derived model accurately explains experimental results. Therefore, we conclude that an accurate nonlinear model is obtained.

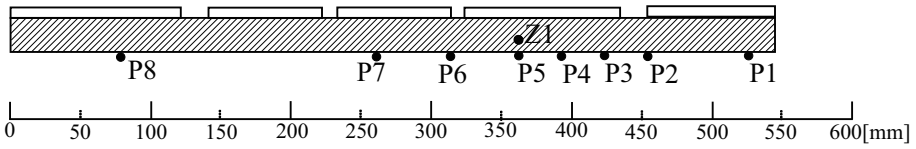


FIGURE 6.8: Measurement points.

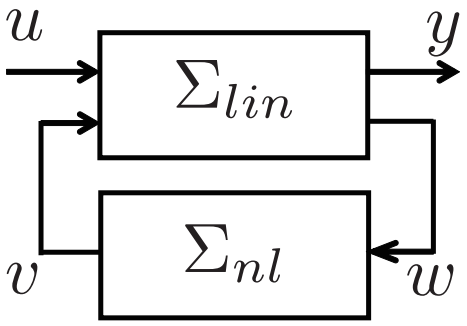


FIGURE 6.9: Overall nonlinear system.

6.4 Model Order Reduction of Nonlinear Network System

In the previous section, we have shown the validity of the obtained nonlinear thermal diffusion network model in (6.13)-(6.16). However, the resultant dimension of the model is 808. Thus, designing controllers and observers based on this model is not useful. One naive approach to obtain low-dimensional model is to discretize the PDEs by coarser spatial step sizes. However, a 409-dimensional model by spatial discretization with $\Delta x = 10$ [mm] does not simulate accurately. Thus, in this section, we reduce the dimension of the whole nonlinear system while preserving its input-output performance. Since model reduction of general nonlinear systems is challenging, we utilize a particular structure: the dimension of the nonlinear system, i.e., $N = 4$, is lower than that of the linear system, i.e., 804. In view of this, we reduce the linear system only and connect the resultant low-dimensional linear system with the original nonlinear system as FIGURE. 6.9. We describe the linear part of the whole system as

$$\Sigma_{lin} : \begin{cases} \dot{x} &= Ax + B_v v \\ y &= C_y x \\ w &= C_w x + D_w u \end{cases} \tag{6.17}$$

where y denotes an evaluating output in (6.13) and the overall structure of the nonlinear system $\Sigma := (\Sigma_{lin}, \Sigma_{nl})$ is shown in FIGURE. 6.9.

6.4.1 Error analysis

We first assume that matrix A in (6.17) is stable. This is the case for our model, since Σ_{lin} is a diffusive system. Thus, a reduced nonlinear model must keep stability with a small approximation error.

As preliminary of main results, we introduce some notations as follows: Let a reduced linear system be $\hat{\Sigma}_{lin}$ and an inter-connected reduced nonlinear model be $\hat{\Sigma} := (\hat{\Sigma}_{lin}, \Sigma_{nl})$. We denote the $\mathcal{H}_\infty$-norm of G_{ij} by γ_{ij} where G_{ij} is a transfer function from $i \in \{y, w\}$ to $j \in \{u, v\}$ of Σ_{lin}. Similarly, $\hat{\gamma}_{ij}$ is the $\mathcal{H}_\infty$-norm of $\hat{G}_{ij}$, that is a transfer function from i to j when $\hat{\Sigma}_{lin}$ is stable. We denote the $\mathcal{H}_\infty$-norm of an error system $G_{ij} - \hat{G}_{ij}$ by ϵ_{ij}. In addition, we assume that $\hat{\Sigma}_{lin}$ is minimal realization without loss of generality.

Using this notation, we have the following theorem:

Theorem 6.2. *Consider Σ_{nl} in (6.14) and Σ_{lin} in (6.17). Define*

$$\mu := \max_{k=1,\dots,N} (\mu_k) \tag{6.18}$$

$$\mu_k := \frac{c_k}{\nu_k} \tag{6.19}$$

$$\nu_k := \min_{z \in \mathbb{R}^N} \psi_k(z) (= \psi_k(0)). \tag{6.20}$$

Let $\hat{\Sigma}_{lin}$ be a stable linear system satisfing

$$\hat{\gamma}_{wv} \mu < 1 \tag{6.21}$$

then the overall reduced order system $\hat{\Sigma} = (\hat{\Sigma}_{lin}, \Sigma_{nl})$ is asymptotically stable. Moreover, for any $\mathcal{L}_2$ bounded input signals u, we have

$$\|y - \hat{y}\|_{\mathcal{L}_2} \leq \epsilon \|u\|_{\mathcal{L}_2} \tag{6.22}$$

where

$$\epsilon = \frac{\epsilon_{yv} \mu \gamma_{wu}}{1 - \gamma_{wv} \mu} + \frac{\hat{\gamma}_{yv} \mu}{1 - \hat{\gamma}_{wv} \mu} \frac{\epsilon_{wv} \mu \gamma_{wu}}{1 - \gamma_{wv} \mu}. \tag{6.23}$$

Proof. To prove this theorem, we first introduce the following notion of *incremental gain*:

Definition 6.3. Consider a nonlinear system

$$\dot{x} = f(x, u), \quad y = g(x, u) \tag{6.24}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^l$. Suppose that $f(0,0) = 0, g(0,0) = 0$ and $x = 0$ are stable equilibriums. If there exists a bounded function $\beta(p,s) : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ such that $\beta(p,s) \geq 0, \beta(0,0) = 0$ and

$$\|y_2 - y_1\|_{\mathcal{L}_2}^2 \leq \mu^2 \|u_2 - u_1\|_{\mathcal{L}_2}^2 + \beta(x_{1,0}, x_{2,0}) \tag{6.25}$$

for any u_1, u_2 in the class of m-dimensional $\mathcal{L}_2$ bounded signals, then the nonlinear system (6.24) has a $\mathcal{L}_2$-bounded incremental gain μ.

Similarly, an incremental gain for linear system is defined, which results in the $\mathcal{H}_\infty$-norm. The following lemma clarifies the relation between the incremental gain of (6.24) and the dissipativity of systems:

Lemma 6.4. *Consider the augmented system of a nonlinear system (6.24) as*

$$\Sigma_{\text{aux}} : \begin{cases} \dot{x_1} = f(x_1, u_1), & y_1 = g(x_1, u_1) \\ \dot{x_2} = f(x_2, u_2), & y_2 = g(x_2, u_2) \end{cases} . \tag{6.26}$$

The nonlinear system (6.24) has incremental gain μ if and only if Σ_{aux} is dissipative with a supply rate

$$s(y_1, u_1, y_2, u_2) = \mu^2 |u_2 - u_1|^2 - |y_2 - y_1|^2. \tag{6.27}$$

Proof. See [60]. □

In this setting, we have the following proposition for the nonlinear system having the structure as FIGURE. 6.9:

Proposition 6.5. *Consider a given Σ_{lin} in (6.17). Let Σ_{lin} be stable and $\hat{\Sigma}_{lin}$ be a stable reduced linear system. In addition, suppose that a given Σ_{nl} in (6.14) is zero-state detectable while $\hat{\Sigma}_{lin}$ and Σ_{nl} have incremental gain $\hat{\gamma}_{ij}$ and μ. If we have*

$$\gamma_{wv}\mu < 1 \tag{6.28}$$

then, $\Sigma = (\hat{\Sigma}_{lin}, \Sigma_{nl})$ has an incremental gain and is asymptotically stable with zero input. Furthermore, an output error bound ϵ in (6.22) is given by

$$\epsilon = \epsilon_{yu} + \frac{\epsilon_{yv}\mu\gamma_{wu}}{1 - \gamma_{wv}\mu} + \frac{\hat{\gamma}_{yv}\mu}{1 - \hat{\gamma}_{wv}\mu}\left(\epsilon_{wu} + \frac{\epsilon_{wv}\mu\gamma_{wu}}{1 - \gamma_{wv}\mu}\right) \tag{6.29}$$

Proof. See [61]. □

Using this proposition, we show that a nonlinear scalar system $\dot{z}_{k,1} = \tilde{\psi}_k(z_{k,1}) + w_{k,1},\ v_{k,1} = z_{k,1}$ has an incremental gain μ_k. To this end, we define a nonlinear system

$$\Sigma_{nl}^{(k)} : \dot{z}_k = \begin{bmatrix} \tilde{\psi}_k(z_{k,1}) \\ \tilde{\psi}_k(z_{k,2}) \end{bmatrix} + w_k, \quad v_k = z_k$$

where $z_k := [z_{k,1}, z_{k,2}]^\mathsf{T}$ and $w_k := [w_{k,1}, w_{k,2}]^\mathsf{T}$. For this system, we have the following lemma:

Lemma 6.6. *Define $\nu_k > 0$ in (6.20). If assumption 6.1 holds and $z_{k,1} \geq z_{k,2}$ is satisfied for all $z_{k,1}, z_{k,2} \in \mathbb{R}$, then it follows that*

$$z_{k,1}\psi_k(z_{k,1}) - z_{k,2}\psi_k(z_{k,2}) \geq \nu_k \left(z_{k,1} - z_{k,2}\right). \tag{6.30}$$

Proof. It is obviously proven in the case of $\psi_k(z_{k,1}) = \nu_k$. In what follows, we consider the case of $\psi_k(z_{k,1}) \neq \nu_k$. Then, $z_{k,1}(\psi_k(z_{k,1}) - \nu_k) \geq z_{k,2}(\psi_k(z_{k,2}) - \nu_k)$ is equivalent to

$$z_{k,1} \geq \frac{\psi_k(z_{k,2}) - \nu_k}{\psi_k(z_{k,1}) - \nu_k} z_{k,2}.$$

Assumption 6.1 yields that $\psi_k(z_{k,1}) \geq \psi_k(z_{k,2})$ when $z_{k,1} \geq z_{k,2} \geq 0$. Thus, (6.30) follows. Similarly to this, we have (6.30) when $0 > z_{k,1} \geq z_{k,2}$. It is obviously proven in the case of $z_{k,1} \geq 0 \geq z_{k,2}$. Hence, the claim follows. □

From Lemma 6.4, we show that $\Sigma_{nl}^{(k)}$ has a storage function $S_k(z_k) := \mu_k(z_{k,1} - z_{k,2})^2$ with a supply rate $s_k(\cdot) = \mu_k^2|w_{k,2} - w_{k,1}|^2 - |v_{k,2} - v_{k,1}|^2$, i.e., it suffices to show that

$$\dot{S}_k(z_k) \leq \mu_k^2|w_{k,2} - w_{k,1}|^2 - |v_{k,2} - v_{k,1}|^2. \tag{6.31}$$

Lemma 6.6 for $z_{k,1} \geq z_{k,2}$ yields that

$$\begin{aligned}
\dot{S}_k(z_k) &= 2\mu_k(z_{k,1} - z_{k,2}) \left(-\frac{1}{c_k} \left((z_{k,1} + z_k^*)\ \psi_k(z_{k,1} + z_k^*) - z_k^*\psi_k(z_k^*) - (z_{k,2} + z_k^*) \right.\right. \\
&\quad \left.\left. \psi_k(z_{k,2} + z_k^*) + z_k^*\psi_k(z_k^*) \right) + w_{k,1} - w_{k,2} \right) \\
&\leq -\frac{2}{c_k}\mu_k(z_{k,1} - z_{k,2})^2\nu_k + 2\mu_k(z_{k,1} - z_{k,2})(w_{k,1} - w_{k,2}).
\end{aligned}$$

Similarly to the case of $z_{k,1} < z_{k,2}$, we see that the sufficient condition for (6.31) is

$$-\frac{2\mu_k\nu_k}{c_k}p^2 + 2\mu_k pq - \mu_k^2 q^2 + p^2 \leq 0 \tag{6.32}$$

for any $p, q \in \mathbb{R}$. Eq. (6.32) is equivalently written as

$$\left(1 - \frac{2\mu_k\nu_k}{c_k}\right)\left(p + \frac{\mu_k c_k}{c_k - 2\mu_k\nu_k}q\right)^2 - \mu_k^2\left(1 + \frac{c_k}{c_k - 2\mu_k\nu_k}\right)q^2 \leq 0. \tag{6.33}$$

On the other hand, it follows from (6.19) that

$$c_k - 2\mu_k\nu_k \leq 0, \quad 1 + \frac{c_k}{c_k - 2\mu_k\nu_k} \geq 0.$$

Thus, (6.32) follows. Hence, the nonlinear scalar system

$$\dot{z}_{k,1} = \tilde{\psi}_k(z_{k,1}) + w_{k,1}, \ v_{k,1} = z_{k,1}$$

has the incremental gain μ_k. Second, we show that Σ_{nl} has the incremental gain $\mu := \max_k(\mu_k)$. Define

$$S(z_1, \ldots, z_N) := \sum_{k=1}^{N} S_k(z_k). \tag{6.34}$$

Then, (6.31) yields that $S(\cdot)$ satisfies

$$\dot{S}(z_1, \ldots, z_N) \leq \mu^2 |w_{:,2} - w_{:,1}|^2 - |v_{:,2} - v_{:,1}|^2$$

where $w_{:,1} := [w_{1,1}, \ldots, w_{N,1}]^\mathsf{T}$. Thus Σ_{nl} has the incremental gain $\mu := \max_k(\mu_k)$. Finally, we show the stability of $\hat{\Sigma}$ and the error bound. Eq. (6.14) implies that Σ_{nl} is zero-state detectable. Thus, Proposition 6.5 shows that the overall nonlinear system $\hat{\Sigma}$ is asymptotically stable with zero input. Furthermore, since $\epsilon_{yu} = \epsilon_{wu} = 0$ holds, the error bound is given by (6.23). □

The parameter μ in (6.18) and μ_k in (6.19) act as an incremental gain of nonlinear systems; see the details in the proof. The parameter μ_k is characterized by the lower value of the coefficient of radiation ψ_k. Since ψ_k dominates the decay rate of energy of heaters, it implies that (6.23) evaluates an approximation error with the minimum decay rate. It is reasonable that the error ϵ becomes larger as ψ_k gets smaller.

Note that $\epsilon_{ij} \leq \|G - \hat{G}\|_{\mathcal{H}_\infty}$ holds where G and $\hat{G}$ are transfer functions of Σ_{lin} and $\hat{\Sigma}_{lin}$, respectively. In addition, we define δ satisfying $\epsilon_{ij} \leq \delta$, then $\hat{\gamma}_{ij} \leq \gamma_{ij} + \delta$ holds. In this setting, the reduction procedure using this theorem is summarized as follows:

1. For given Σ_{lin} and Σ_{nl}, compute γ_{ij} and μ in (6.18).
2. Find maximum $\delta > 0$ such that the error bound ϵ in (6.23) is less than a desired value and (6.21) holds. Note that we can replace ϵ_{ij} and $\hat{\gamma}_{ij}$ by δ and $\gamma_{ij} + \delta$ as a priori bounds.

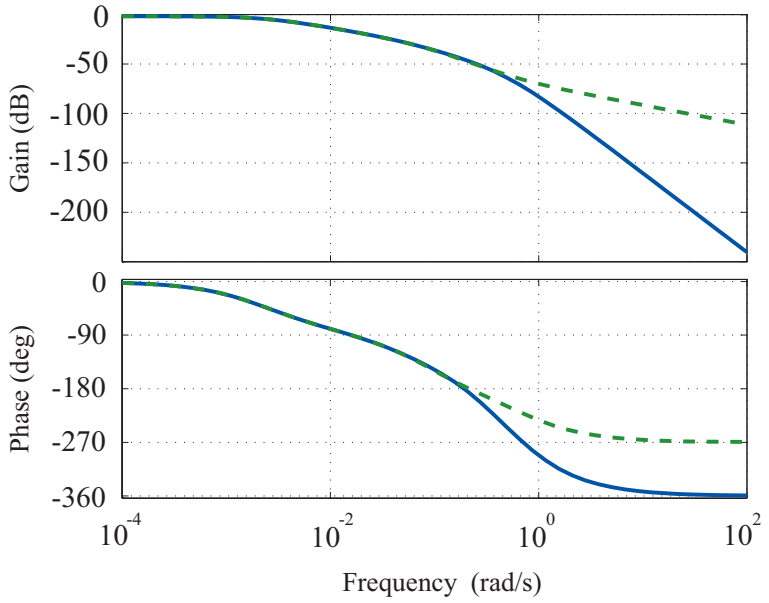


FIGURE 6.10: Model reduction of the linear part: Bode diagram of the transfer function from v_1 to y_1 for the original (blue) and the reduced order models (green).

3. Find a reduced order model $\hat{\Sigma}_{lin}$ satisfying $\|G - \hat{G}\|_{\mathcal{H}_\infty} \leq \delta$ by means of a model reduction method with preserving the stability of $\hat{G}$, e.g., balanced truncation [17, 26].

4. Obtain $\hat{\Sigma}$ by inter-connecting $\hat{\Sigma}_{lin}$ and Σ_{nl} as shown in FIGURE. 6.9.

6.4.2 Simulation results

In this subsection, we demonstrate the efficiency of the model reduction method described in the previous subsection.

For given Σ_{lin} and Σ_{nl}, we have $\mu = 30.0$, $\gamma_{wu} = 1.3 \times 10^{-4}$, $\gamma_{yv} = 1.7$ and $\gamma_{wv} = 3.2 \times 10^{-2}$. In the second step of the procedure given above, taking $\delta = 3 \times 10^{-4}$, we have $\epsilon < 1.5 \times 10^{-1}$. Reducing Σ_{lin} by means of balanced truncation, we have a further reduced 28-dimensional nonlinear model via the 808-dimensional original nonlinear model. The obtained model satisfies (6.21) with $\mu\hat{\gamma}_{wv} = 0.96$, which implies that the model is stable. This model is the smallest one satisfying the condition in (6.21). Moreover, we have $\epsilon = 2.7 \times 10^{-2}$. Taking into account the fact that $y(t)$ and $u(t)$ are $\mathcal{O}(10^3)$ and $\mathcal{O}(10^2)$, we can see that the resultant low-dimensional model approximates the original 808-dimensional model accurately.

In FIGURE. 6.10, we show the bode diagram of the original transfer function G from v_1 to y_1 and that of $\hat{G}$ by blue and green lines, respectively. This figure shows that $\hat{G}$ appropriately approximates G.

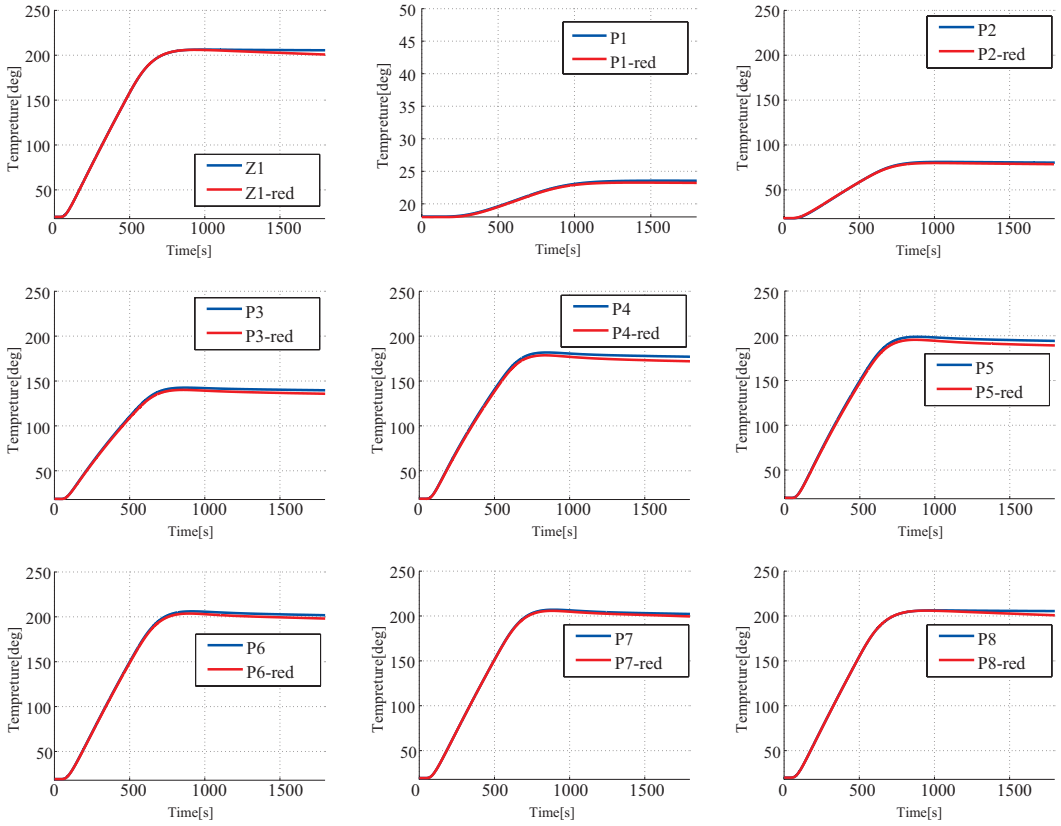


FIGURE 6.11: Transient trajectory difference between the original (black) and reduced order (glay) models.

Finally, in FIGURE. 6.11, the blue and red lines depict transient responses of the overall low-dimensional model $(\hat{\Sigma}_{lin}, \Sigma_{nl})$ and the original model $(\Sigma_{lin}, \Sigma_{nl})$ where the input signal is the same as that used in Section 6.3. The points $Z1$, $P1$-$P8$ are shown in FIGURE. 6.8. These figures show that we have obtained the accurate low-dimensional network system for the spatially distributed nonlinear dynamics.

6.5 Chapter summary

In the first half of this chapter, we have provided the nonlinear model of plasticization cylinders including nonlinear radiation to outer air. Furthermore, we have shown the accuracy of the spatially discretized nonlinear network system by experiment. In the latter half of this chapter, we have reduced the order of the nonlinear network system with the theoretical guarantee about stability as well as the approximation error. The accuracy of the obtained low-dimensional nonlinear network system has been evaluated by numerical simulation with real data. These results have shown that we have obtained the appropriate low-dimensional nonlinear network system explaining experimental results.

Chapter 7

Conclusion

7.1 Summary of results

This thesis is summarized as a line of work towards development of systematic control theory for large-scale dynamical network systems. Summaries in individual chapters are as follows:

In Chapter 2, we have proposed a method of designing low-dimensional functional observers to estimate a given set of states via an observer reduction approach. This type of observer is useful for estimation of a limited number of states such as load power in a particular area of large-scale electric power network systems. We have clarified that we have to take into account not only an initial state estimation error but also external input signals. Analyzing these estimation error factors based on model reduction techniques, we have derived an a priori $\mathcal{L}_2$-error bound on the performance degradation with the provision of systematic design.

In Chapter 3, towards estimation of overall dynamical behavior of large-scale network systems with small computational costs, we have proposed an average state observer that estimates average behavior of the network system from a macroscopic point of view. In general, we do not know a set of states capturing average behavior of the network system in advance. By the proposed method, we can construct a performance guaranteed average state observer with systematic determination of a set of states capturing average behavior of systems. Towards systematic design of average state observers while determining a set of states capturing average behavior, we have first derived a tractable representation of the estimation error system by an average state observer. On the basis of the representation of error systems, we have provided a systematic design procedure with an a priori $\mathcal{L}_2$-error bound of the estimation error.

In Chapter 4, we have proposed hierarchical distributed control for general linear network systems. Towards systematic design, we have introduced state-space expansion that enables us to independently deal with the state variables associated with disjoint subsystems and those associated with the interference among hierarchically clustered subsystems in a tractable manner. On the basis of this state-space expansion, we have devised a design method to construct a hierarchical distributed controller, whose compositional units can be designed individually having an advantage that an $\mathcal{L}_2$-performance of the closed-loop system improves as just improving an $\mathcal{L}_2$-performance of local controllers that stabilizes disjoint subsystems individually.

However, the hierarchical distributed controller does not fully comply with practical application for large-scale network systems from a viewpoint of computational costs for implementation because the dimension of compositional units in upper layers is comparable with that of the system to be controlled. In view of this, in Chapter 5, we have provided a method of reducing the dimension of hierarchical distributed controllers that approximates the original hierarchical distributed controller for any sets of locally stabilizing controllers. To solve this problem, explicitly utilizing a hierarchically distributed structure of the closed-loop system, we have shown that the approximation error of the closed-loop system can be evaluated by that of the system associated with the hierarchical distributed controller without local controllers. Furthermore, we have derived the relation between the approximation errors of the individual hierarchical controllers and the performance degradation of the closed-loop system.

As a first step towards development of control theory for nonlinear large-scale network systems, in Chapter 6, we have shown the importance and necessity of nonlinear low-dimensional modelling through an example of plasticization cylinders including nonlinear radiation to outer air. Utilizing a particular structure of the dynamics of plasticization cylinders, we have provided a nonlinear model reduction method that preserves stability as well as input-output performances. The accuracy of the obtained low-dimensional nonlinear network system, which is 28-dimensional model whereas the original model is 808-dimensional nonlinear system, has been evaluated by experiment.

7.2 Future works

The average state observers proposed in Chapter 3 has potential to be useful for practical applications, e.g., weather prediction and data assimilation [1, 6]. In fact, we have extended the observer to the case of time-invariant Kalman filter dealing with system noises and measurement noises in [62]. Further extension to nonlinear and time-variant

filters is one of the most important study to develop practical tools for large-scale network systems.

In addition, in Chapters 4, 5, we have proposed hierarchical distributed control with a given sets of clusters. As discussed in Section 4.3.3, we have to devise a systematic method of hierarchical clustering and sensor/actuator allocation towards scalable implementation. Furthermore, the proposed controller has the spatial hierarchy such as local controllers in particular areas and a hierarchical distributed controller having jurisdiction over the network system of interest. Similarly to this, we can consider hierarchical controllers having temporal hierarchy based on spectral decomposition for example. This extension is expected to establish a novel control theory from a spatiotemporal point of view.

Bibliography

[1] G. Evensen, *Data assimilation.* Springer, 2007.

[2] ——, "Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics," *Journal of Geophysical Research-all Series*, vol. 99, pp. 10–10, 1994.

[3] E. Kalnay, *Atmospheric modeling, data assimilation and predictability.* Cambridge university press, 2002.

[4] A. Lawless, N. Nichols, C. Boess, and A. Bunse-Gerstner, "Using model reduction methods within incremental four-dimensional variational data assimilation," *Monthly Weather Review*, vol. 136, no. 4, pp. 1511–1522, 2008.

[5] P. Vermeulen and A. Heemink, "Model-reduced variational data assimilation," *Monthly Weather Review*, vol. 134, no. 10, pp. 2888–2899, 2006.

[6] M. Vermeulen, "Efficient kalman filtering algorithms for hydrodynamic models," *Ph.D. Thesis at Technische Universiteit, Delft, The Netherlands*, 1998.

[7] J. Ferziger and M. Perić, *Computational methods for fluid dynamics.* Springer Berlin etc, 1999.

[8] J. Holman, *Heat Transfer tenth edition.* McGraw Hill Higher Education, 2009.

[9] G. K. Batchelor, *An introduction to fluid dynamics.* Cambridge university press, 2000.

[10] P. Kundur, *Power system stability and control.* Tata McGraw-Hill Education, 1994.

[11] M. D. Ilic and S. Liu, *Hierarchical power systems control: its value in a changing industry.* Springer Heidelberg, 1996.

[12] A. Chakrabortty, "Wide-area damping control of large power systems using a model reference approach," in *Proc. of Decision and Control and European Control Conference.* IEEE, 2011, pp. 2189–2194.

[13] Y. Riffonneau, S. Bacha, F. Barruel, and S. Ploix, "Optimal power flow management for grid connected pv systems with batteries," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 309–320, 2011.

[14] T. Masuta and A. Yokoyama, "Supplementary load frequency control by use of a number of both electric vehicles and heat pump water heaters," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1253–1262, 2012.

[15] M. Mesbahi and M. Egerstedt, *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[16] S. H. Strogatz, "Exploring complex networks," *Nature*, vol. 410, no. 6825, pp. 268–276, 2001.

[17] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005.

[18] ——, "An overview of model reduction methods and a new result," in *Proc. of Decision and Control and Chinese Control Conference*. IEEE, 2009, pp. 5357–5361.

[19] A. K. Singh and J. Hahn, "State estimation for high-dimensional chemical processes," *Computers & chemical engineering*, vol. 29, no. 11, pp. 2326–2334, 2005.

[20] W. H. Schilders, H. A. Van der Vorst, and J. Rommes, *Model order reduction: theory, research aspects and applications*. Springer, 2008, vol. 13.

[21] T. Iwasaki and R. E. Skelton, "All controllers for the general $\mathcal{H}_\infty$ control problem: Lmi existence conditions and state space formulas," *Automatica*, vol. 30, no. 8, pp. 1307–1317, 1994.

[22] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to $\mathcal{H}_\infty$ control," *International journal of robust and nonlinear control*, vol. 4, no. 4, pp. 421–448, 1994.

[23] X. Xin, L. Guo, and C. Feng, "Reduced-order controllers for continuous and discrete-time singular $\mathcal{H}_\infty$ control problems based on LMI," *Automatica*, vol. 32, no. 11, pp. 1581–1585, 1996.

[24] P. J. Goddard and K. Glover, "Controller approximation: approaches for preserving $\mathcal{H}_\infty$ performance," *Automatic Control, IEEE Transactions on*, vol. 43, no. 7, pp. 858–871, 1998.

[25] B. D. Anderson and Y. Liu, "Controller reduction: concepts and approaches," *Automatic Control, IEEE Transactions on*, vol. 34, no. 8, pp. 802–812, 1989.

[26] G. Obinata and B. Anderson, *Model reduction for control system design*, ser. Communications and control engineering. Springer, 2000.

[27] A. Varga and B. Anderson, "Accuracy-enhancing methods for balancing-related frequency-weighted model and controller reduction," *Automatica*, vol. 39, no. 5, pp. 919–927, 2003.

[28] D. F. Enn, "Model reduction for control systems design," *Ph.D. dissertation, Stanford University*, 1984.

[29] G. A. Latham and B. Anderson, "Frequency-weighted optimal hankel-norm approximation of stable transfer functions," *Systems & control letters*, vol. 5, no. 4, pp. 229–236, 1985.

[30] M. K. Sundareshan and R. M. Elbanna, "Design of decentralized observation schemes for large-scale interconnected systems: some new results," *Automatica*, vol. 26, no. 4, pp. 789–796, 1990.

[31] F. Garin and L. Schenato, "A survey on distributed estimation and control applications using linear consensus algorithms," in *Networked Control Systems*. Springer, 2010, pp. 75–107.

[32] C. Langbort, R. S. Chandra, and R. D'Andrea, "Distributed control design for systems interconnected over an arbitrary graph," *Automatic Control, IEEE Transactions on*, vol. 49, no. 9, pp. 1502–1519, 2004.

[33] J. C. Willems, "Dissipative dynamical systems part i: General theory," *Archive for rational mechanics and analysis*, vol. 45, no. 5, pp. 321–351, 1972.

[34] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *Automatic Control, IEEE Transactions on*, vol. 51, no. 2, pp. 274–286, 2006.

[35] E. Camponogara, D. Jia, B. H. Krogh, and S. Talukdar, "Distributed model predictive control," *Control Systems, IEEE*, vol. 22, no. 1, pp. 44–52, 2002.

[36] L. Bakule, "Decentralized control: An overview," *Annual reviews in control*, vol. 32, no. 1, pp. 87–98, 2008.

[37] T. Ishizaki, H. Sandberg, K. H. Johansson, K. Kashima, J.-i. Imura, and K. Aihara, "Structured model reduction of interconnected linear systems based on singular perturbation," in *Proc. of American Control Conference*, 2013, pp. 5544–5549.

[38] T. Ishizaki, K. Kashima, J.-i. Imura, and K. Aihara, "Model reduction and clusterization of large-scale bidirectional networks," *Automatic Control, IEEE Transactions on*, vol. 59, no. 1, pp. 48–63, 2014.

[39] T. Ishizaki, K. Kashima, A. Girard, J. Imura, L. Chen, and K. Aihara, "Clustering-based $\mathcal{H}_2$-state aggregation of positive networks and its application to reduction of chemical master equations," in *Proc. of Decision and Control*. IEEE, 2012, pp. 4175–4180.

[40] C. Langbort and J. Delvenne, "Distributed design methods for linear quadratic control and their limitations," *Automatic Control, IEEE Transactions on*, vol. 55, no. 9, pp. 2085–2093, 2010.

[41] Y. Ebihara, D. Peaucelle, and D. Arzelier, "Decentralized control of interconnected positive systems using L_1-induced norm characterization," in *Proc. Conference on Decision and Control*. IEEE, 2012, pp. 6653–6658.

[42] D. Luenberger, "An introduction to observers," *Automatic Control, IEEE Transactions on*, vol. 16, no. 6, pp. 596–602, 1971.

[43] M. Darouach, "Existence and design of functional observers for linear systems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 940–943, 2000.

[44] H. Trinh and T. Fernando, *Functional observers for dynamical systems*. Springer Verlag, 2011, vol. 420.

[45] C.-C. Tsui, "What is the minimum function observer order?" *Journal of the Franklin Institute*, vol. 335, no. 4, pp. 623–628, 1998.

[46] ——, "A new algorithm for the design of multifunctional observers," *Automatic Control, IEEE Transactions on*, vol. 30, no. 1, pp. 89–93, 1985.

[47] C. Tsui, "On the order reduction of linear function observers," *Automatic Control, IEEE Transactions on*, vol. 31, no. 5, pp. 447–449, 1986.

[48] S. Boyd, L. El Ghaoul, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*. Society for Industrial and Applied Mathematics, 1987, vol. 15.

[49] S. Sastry, *Nonlinear systems: analysis, stability, and control*. Springer New York, 1999, vol. 10.

[50] A. Monticelli, *State estimation in electric power systems: a generalized approach*. Springer, 1999, vol. 507.

[51] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics reports*, vol. 424, no. 4, pp. 175–308, 2006.

[52] S. P. Boyd, *Linear matrix inequalities in system and control theory*. Siam, 1994, vol. 15.

[53] T. Ishizaki and J. Imura, "Clustered model reduction of interconnected second-order systems," *Nonlinear Theory and Its Applications, Institute of Electronics, Information and Communication Engineers, Special Section of Complex Systems Modelling and its Transdisciplinary Applications*, vol. E6-N, no. 1, 2015 (in Press).

[54] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking," *Physical Review Letters*, vol. 85, no. 21, pp. 4633–4636, 2000.

[55] M. Ikeda and D. D. Šiljak, "Overlapping decentralized control with input, state, and output inclusion," *Control Theory and Advanced Technology*, vol. 2, no. 2, pp. 155–172, 1986.

[56] T. Ishizaki, M. Koike, T. Sadamoto, and J. Imura, "Hierarchical decentralized observers for networked linear systems," in *Proc. American Control Conference*. IEEE, 2014.

[57] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, "On the decay rate of Hankel singular values and related issues," *Systems & control letters*, vol. 46, no. 5, pp. 323–342, 2002.

[58] K. Zhou, J. C. Doyle, K. Glover, *et al.*, *Robust and optimal control*. Prentice Hall New Jersey, 1996, vol. 40.

[59] A. Barrat, M. Barthlemy, and A. Vespignani, *Dynamical processes on complex networks*. Cambridge University Press, 2008.

[60] B. Romanchuk and M. James, "Characterization of the $\mathcal{L}_p$ incremental gain for nonlinear systems," in *Proc. of Conference on Decision and Control*, 1996, pp. 3270–3275.

[61] B. Besselink, N. Van De Wouw, and H. Nijmeijer, "Model reduction of nonlinear systems with bounded incremental $\mathcal{L}_2$ gain," in *Proc. of Conference on Decision and Control*, 2011, pp. 7170–7175.

[62] F. Watanabe, T. Sadamoto, T. Ishizaki, and J. Imura, "Average state kalman filters for large-scale stochastic networked linear systems," in *Proc. European Control Conference (submitted)*, 2015.