Subsystem decomposition and distributed state estimation of nonlinear process networks

by

Xunyuan Yin

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Process Control

Department of Chemical and Material Engineering

University of Alberta

Abstract

The distributed framework has been recognized as a promising framework for handling the large size and the strong interaction of modern integrated chemical processes. Compared to distributed process control, its dual problem - distributed state estimation which is equally important, has received relatively less attention in the process control community. This thesis deals with distributed state estimation for nonlinear process networks as well as subsystem structure decomposition and configuration for distributed state estimation. The contributions of this thesis include the development of distributed state estimation methods for nonlinear process networks, the development of systematic approaches to properly decompose general nonlinear processes into subsystems for distributed estimation, and the application of the developed methods in different processes and output-feedback fault detection and isolation.

First, two-time-scale nonlinear systems are taken into account. The system is decomposed into fast and slow subsystems based on the singular perturbation theory. Local observer-enhanced moving horizon state estimators are designed. A one-directional communication scheme is used. The convergence and boundedness of the estimation error is rigorously studied. A benchmark chemical process example is used to illustrate the proposed method. Then, attention is given to general nonlinear systems that can be divided into smaller subsystems. It is assumed that a decentralized state estimation scheme already exists for the system. And the aim is to form a distributed state estimation scheme based on the existing one without significant modifications. Compensators are designed for the subsystems and

augmented estimators that communicate with each other are obtained to form a distributed network. Stability analysis is carried out. The convergence and ultimate boundedness of the estimation error dynamics can be ensured subject to reasonable assumptions. The proposed method is applied to three application examples and good state estimates are obtained. This distributed state estimation method is also utilized in the development of a distributed output-feedback fault detection and isolation mechanism for cascade nonlinear processes. The state estimation scheme provides state estimates that are further used for generating residual signals for the subsystems. A distributed fault detection and isolation mechanism is proposed and applied to a froth flotation process.

With the development of distributed state estimation algorithm being completed, we further explore the decomposition and configuration of subsystem structure for distributed state estimation. A systematic procedure is proposed to address the considered subsystem decomposition problem. The procedure is based on the evaluation of physical closeness between state and output measurement variables. The proposed method is applied to the froth flotation process and subsystem models are configured. The decomposition result is consistent with physical topology and can be readily used for distributed state estimation design. Moreover, we consider systematic subsystem decomposition and distributed state estimation for a wastewater treatment plant. Based on an extension of the previously proposed method, the large-scale plant is decomposed into subsystems and local state estimators are developed. Good simulation results confirm the effectiveness of the systematic approach. Finally, to facilitate the synthesis of distributed state estimation and distributed control, a systematic approach on subsystem decomposition of process networks for simultaneous distributed estimation and control is presented.

Preface

The results presented in this thesis are part of the research that is under the supervision of Dr. Jinfeng Liu and is funded by Natural Sciences and Engineering Research Council (NSERC) of Canada and Alberta Innovates Technology Futures (AITF). Chapter 2 of this thesis is a revised version of X. Yin and J. Liu, Distributed moving horizon state estimation of two-time-scale nonlinear systems. *Automatica*, 79:152-161, 2017. Chapter 3 of this thesis is a revised version of X. Yin, J. Zeng, and J. Liu, Forming distributed state estimation network from decentralized estimators, *IEEE Transactions on Control Systems Technology*, 2018 (in press). Chapter 4 is a revised version of X. Yin and J. Liu, Distributed output-feedback fault detection and isolation of cascade process networks, *AIChE Journal*, 63:4329-4342, 2017. Chapter 5 is a revised version of X. Yin, K. Arulmaran, J. Liu, and J. Zeng, Subsystem decomposition and configuration for distributed state estimation. *AIChE Journal*, 62:1995-2003, 2017. Chapter 6 is a revised version of X. Yin, D.-N. Benjamin, and J. Liu, Subsystem decomposition and distributed moving horizon estimation of wastewater treatment plants. *Chemical Engineering Research and Design*, 134:405-419, 2018. Chapter 7 is a revised version of X. Yin and J. Liu, Subsystem decomposition of process networks for simultaneous distributed state estimation and control, which has been submitted to *AIChE Journal*.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Jinfeng Liu for his tremendous amount of guidance and generous support he has given to me during my graduate study. He is definitely the greatest advisor I could have had. His great patience, enthusiasm and determination have made him my role model. His constant encouragement and excellent guidance have greatly inspired me to move towards a success, significantly improved my confidence, and made profound impacts on my career goals. Dr. Liu has always been providing much freedom to me to do interesting research and the excellent opportunities to attend international conferences and to participate in teaching. Without Dr. Liu's guidance and support during my graduate study, this thesis would not have been possible.

I want to gratefully acknowledge Professor Sirish L. Shah, Professor Biao Huang, Professor Zukui Li and Professor Qing Zhao who have helped me a lot in terms of study and research during the past years. In addition, I sincerely thank Professor Biao Huang, Professor Vinay Prasad, Professor Zukui Li, Professor Joseph Sang-II Kwon, Professor Hyo-Jick Choi for serving on my doctoral examination.

Moreover, I would like to thank my colleagues whom I have worked with during my past four years in PSACE group, including Jing Zhang, Su Liu, Jannatun Nahar, Mohammad Rashedi, Kevin Arulmaran, Jayson McAllister, Benjamin Decardi-Nelson, Tianrui An, Guoyang Yan, An Zhang, Nirwair Bajwa, Yawen Mao, Soumya Sahoo, Song Bo, Rui Nian, Mengzhi Wang, as well as Mengqi Fang, Xiaodong Xu, Yuan Yuan, Zheyuan Liu, Lei Fan, Ming Ma and Yaojie Lu.

I would also like to gratefully acknowledge the financial support from Natural Sciences and Engineering Research Council of Canada (NSERC), Alberta Innovative Technology Futures (AITF), and the scholarships/awards offered by Faculty of Graduate Studies and Research, University of Alberta.

Last but not least, I am very grateful to my family for their constant encouragement and unconditional support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and research overview

Complex and tightly integrated process networks are common occurrences in manufacturing industries (e.g., chemicals, petrochemicals and mineral processes), and have attracted growing interest within the modern process community due to their economic efficiency [1, 2, 3]. A typical process network consists of several operating units (subsystems), which are connected with each other via material, energy and information flows. Due to the significant interaction between different subsystems, great challenges have been posed to the design of automatic control systems. Currently, most of these process networks are controlled in a decentralized manner such that a subsystem control system does not communicate with other subsystem control systems which further renders that interaction between subsystems is treated in conservative ways. During the past decade, considerable efforts have been given to the development of distributed control systems (in particular, distributed model predictive control [4, 5, 1, 6, 7, 8, 9, 10]). Within a distributed control framework, local controllers communicate with each other to exchange information to collaborate their actions. The distributed framework has been demonstrated to be a promising framework for handling the large sizes and the strong interaction of modern integrated chemical processes. Compared

to distributed process control, its dual problem - distributed state estimation, has received relatively less attention in the process control community.

Recently, there are some results on distributed moving horizon state estimation (DMHE). Moving horizon estimation (MHE) is an online optimization-based technique and can handle nonlinearities, constraints and optimality considerations [12, 13]. In [14], a DMHE algorithm was developed for nonlinear systems based on subsystem models. An observer-enhanced DMHE algorithm was developed in [15], where an auxiliary nonlinear observer is taken advantage of in the design of each local estimator. The auxiliary observer is used to calculate a reference state estimate based on which a confidence region is constructed every sampling time. Each local estimator optimizes its estimate within the confidence region. The observer-enhanced design is less sensitive to external noise compared with the auxiliary nonlinear observer. The convergence rate of the DMHE may be tuned by tuning the auxiliary observers. It was shown to be less dependent on the arrival cost, the estimation window size and have the potential to be used in output feedback control with provable closed-loop stability [15, 16]. On the other hand, time-scale multiplicity is a common feature of many systems. For chemical processes, it usually arises due to the strong coupling of physicochemical phenomena [17, 18]. A direct application of standard control or estimation methods without taking into account time-scale multiplicity to systems with different time scales may lead to ill-conditioning or even the loss of closed-loop stability [19, 20]. The singular perturbation theory is the standard tool for the analysis of systems with time-scale multiplicity [19, 20]. The majority of related results are on control system design for two-time-scale systems (e.g., [19, 20, 21]). Little attention has been given to state estimation of systems with time-scale multiplicity except in [22]. Therefore, it is much favorable to design a DMHE scheme that can handle the time-scale multiplicity in two-time-scale nonlinear systems.

Besides DMHE, there are several different ways to solve the distributed state estimation problem. The existing algorithms on distributed state estimation are primarily developed in four frameworks: the deterministic observer framework [31, 32], the Kalman filter (KF)

framework [33, 34, 35, 36, 11], the extended Kalman filter (EKF) framework [37] and the moving horizon estimation (MHE) framework [38, 39, 40]. It is worth mentioning that most of the existing algorithms were developed for linear systems with a few exceptions [41, 15, 42, 47, 113], where distributed moving horizon state estimation problems for nonlinear systems have been addressed. Moreover, in [81, 44], consensus-based estimators were designed for nonlinear systems to provide references for fault detection. Distributed state estimation was developed for fault diagnosis in interconnected nonlinear systems subject to uncertainties [45, 46]. In particular, the spatial-distribution-based subsystem decomposition, the communication scheme and the adopted distributed estimator framework in [81] provide inspirations to the present work. There is a prominent feature of all the above algorithms; that is, they require all the local estimators to be designed simultaneously and to be of the same type (e.g., all the local estimators should be MHEs [38]) and each of the developed algorithms is only compatible with one specific type of local estimators (e.g., distributed KF algorithms cannot be utilized to coordinate MHE-based local estimators). It is possible and sometimes favorable that different types of estimators are used for different subsystems. For example, an EKF-based estimator is used for a subsystem when no constraints should be considered and an MHE-based estimator is used for a subsystem to address constraints. Therefore, it is desirable to develop a more general approach, which can incorporate different types of local estimators together and provide improved estimation performance. Another feature of the existing distributed state estimation approaches is that in the design of these algorithms, no consideration is given to the potentially existing (decentralized) implementation of control/estimation algorithms in a process. If a decentralized state estimation algorithm has already been implemented in a process, the aforementioned distributed state estimation methods require a complete re-design of the existing implementation. This observation also motivates us to develop a general method that is capable of taking advantage of existing (decentralized) estimators together to provide improved overall estimation performance instead of performing a complete re-design.

Potential faults commonly encountered in complex and integrated process networks are another factor that can significantly affect the operating performance of these processes. In particular, faults (e.g., actuator faults and sensor faults) in some key components may result in control performance degradation, a higher risk of system failures, reduction in economic profits, or even catastrophic impacts on operation safety as well as the environment [62, 63, 64, 65, 66]. Since these undesired consequences may propagate from one single subsystem to the entire plant very quickly in modern process networks, FDI is especially important for these plants. During the past two decades, we have witnessed rapid progress in FDI. In the context of linear systems, model-based FDI approaches have been proposed. In [67, 68, 69], for instance, system redundancy is taken advantage of to create residuals for detection and identification of failures in different components. In another line of work, effort has been devoted to FDI of systems exhibiting significant nonlinearities. To name a few, in [70], detection and isolation of potential actuator faults involved in nonlinear systems was addressed based on a geometric approach. Assuming that all the states are measurable, a systematic approach was developed for FDI of actuator faults of nonlinear processes by examining the system structures [71]. More representative results are referred to [62, 74, 75, 76, 77]. Most approaches were developed within the centralized framework, and less attention has been given to the distributed FDI, which is more favorable and sometimes a necessity for complex and highly integrated plants due to its scalability and its capability in handling subsystem interaction [78].

Distributed fault detection problems were investigated for linear systems in Ding et al. [79], Shames et al. [80]. Distributed fault detection, diagnosis, isolation and fault tolerant control approaches were proposed for different nonlinear systems in [64, 81, 82, 83]. However, most of these results are based on the availability of measurements of the entire system states, which may not be satisfied in many applications. Therefore, state estimation based distributed FDI approaches are highly desirable from a practice perspective. For FDI purposes, it is desirable not only to ensure the convergence of the state estimation system,

but also to achieve a sufficiently fast convergence rate of the estimation error. In [78], the authors made an initial effort to design a distributed fault detection architecture based on a bank of observers with exponential convergence rates for input-output interconnected systems. We note that fault isolation was not considered in [78], but also needs to be addressed appropriately.

We have also noticed that subsystem decomposition for decentralized/distributed state estimation designs that is equivalently important for achieving satisfying state estimation performance has not received sufficient research attention. Actually, improper subsystem decomposition and local estimator configuration may lead to the increase of computational burdens, inaccurate estimation results or even deterioration of the observability of the entire system. In terms of control, subsystem decomposition and control structure configuration have been considered for decentralized/distributed control [89]. In [90], a decentralized control system structure selection and optimization approach was proposed based on a constrained genetic algorithm. In [91], an automatic structure selection approach was proposed for chemical processes within a decentralized control framework. A control structure selection approach for distributed model predictive control subject to model errors was reported in [92]. In [93, 95, 96], integer-optimization-based hierarchical clustering was taken advantage of to configure block decentralized control structures. The measure of relative sensitivity array was proposed in [94] for the selection of fully decentralized control structures. More results on control structure selection can be found in [98, 99, 100, 101, 102, 103, 104, 105, 137]. The aforementioned results primarily focus on control structure selection and configuration. Currently, no systematic methodology has been reported with respect to subsystem decomposition and configuration for distributed state estimation, on which we will also place emphasis.

It is much desirable if the subsystem decomposition and distributed state estimation can be jointly considered for large-scale processes, for example, the wastewater treatment plant. Wastewater treatment plants (WWTPs) are commonly used for the disposal of wastewater

to substantially reduce the environmental impacts of wastewater and to convert wastewater to reclaimed water [120, 121]. A WWTP which typically consists of several interconnected operating units, is a nonlinear process where complex physical and biological phenomena take place. The influent flow rate and composition to a WWTP may fluctuate significantly, which makes the control and monitoring of a WWTP challenging. In literature, several state estimation methods have been proposed for WWTPs. In [122], a state estimation method for WWTP was developed based on the singular perturbation theory for a WWTP with only 6 states. In [123], a centralized estimator was developed for a WWTP based on extended Kalman filter (EKF). In [125], a distributed EKF scheme of two subsystem estimators was developed. In [124, 126], a state estimation method was proposed based on model reduction for improved computational efficiency. While EKF was shown to give acceptable estimation performance in the above studies, it may give ad hoc results and is not robust to disturbance or poor initial guess due to the way nonlinearity is treated in EKF [127]. It is worthwhile to consider an estimation method like moving horizon estimation (MHE) that can handle the nonlinearity of WWTPs more appropriately and take constraints into consideration for further improved estimation performance. Initial attempts on applying MHE in state estimation of WWTPs were made in [128, 123], where centralized MHE estimators were developed for WWTPs described by the Activated Sludge Model No.1.

However, MHE is an optimization-based method and it has much higher computational complexity for nonlinear systems compared with EKF. Its high computational complexity is the main obstacle that prevents people from using MHE in nonlinear systems. This point is especially important for further synthesis of state estimation based MPC from an online implementation perspective. One effective way to reduce the computational complexity is to design MHE-based estimators within a distributed framework, which divides a large optimization problem into smaller sub-problems [15, 42]. Distributed state estimation can also be much favorable for WWTPs in terms of organizational complexity and fault tolerance [1, 47]. As mentioned earlier, subsystem decomposition is another key step for distributed

state estimation [103, 132, 61, 73]. Based on the above observations, we aim to propose a systematic subsystem decomposition and distributed MHE approach for the state estimation of WWTPs [130]. The developed distributed MHE scheme can provide online state estimates, which may be used in distributed MPC to form a distributed output-feedback control scheme for the WWTPs.

From an application point of view, the ultimate goal is to design distributed output-feedback control schemes for large-scale processes, which require the incorporation of distributed state estimation and distributed control in one integrated design. Regarding the subsystem decomposition for distributed output-feedback control, One possible solution is to perform subsystem decomposition for distributed estimation and distributed control separately and independently using existing methods. However, from the perspective of implementation, maintenance and communication, it is much more favorable if the local state estimators and local controllers are designed based on the same subsystem decomposition. However, a systematic approach to achieve this objective is not yet available.

The community detection concept originating from network theory provides a very promising way to address the considered problem [102, 103]. By means of the measure of modularity [142], community-based approaches have been proposed to find distributed control structures where the subsystems involving state, input and controlled output variables are made well-decoupled [132, 102]. In [103], the community-based method was applied to a benchmark chemical process to recommend different distributed control structures, and DMPC was designed based on each structure for extensive performance comparison. This motivates us to work on subsystem decomposition of nonlinear process networks for simultaneous distributed state estimation and distributed control by means of the concept of community structure detection.

1.2 Contributions and thesis outline

The rest of the thesis is organized as follows:

In Chapter 2, the scope is on the handling of time-scale multiplicity in state estimation. Specifically, we consider state estimation of a class of two-time-scale nonlinear systems. A system is first decomposed into a reduced-order fast system and several reduced-order slow subsystems. A fast MHE is designed for the fast system and a slow MHE is designed for each slow subsystem. The fast and slow MHEs form a DMHE scheme. Each MHE is designed via the observer-enhanced MHE technique [24]. It is discovered that the slow MHEs are entirely decoupled from the fast MHE which is a significant difference from control of two-time-scale systems. The decoupling ensures that only one-directional information transmission from the slow MHEs to the fast MHE is needed and the fast MHE does not send out any information. Sufficient conditions are derived under which the proposed DMHE is guaranteed to give ultimately bounded estimation error under bounded system disturbances and measurement noise. A reactor-separator process example is introduced to demonstrates the applicability and effectiveness of the method proposed in Chapter 2.

In Chapter 3, we consider the problem of distributed state estimation of nonlinear systems comprised of interconnected subsystems. We consider that a decentralized state estimation system comprised of local estimators potentially of different types has already been implemented for a nonlinear process network. In order to achieve improved estimation performance, the existing decentralized estimators may be connected together via a communication network to form a distributed state estimation network. We propose a systematic approach to take advantage of the existing decentralized estimators potentially of different types to form a distributed state estimation network without performing a complete re-design of the estimation system. Specifically, a compensator is designed for each subsystem, and is connected to the corresponding decentralized estimator to obtain an augmented estimator. The augmented estimators for the subsystems communicate with each other to exchange subsystem state estimates and measurements via a communication network every sampling time.

We derive sufficient conditions on the convergence and boundedness of the estimation error of the proposed distributed estimation network. Within the proposed distributed framework, estimation algorithms of local estimators for different subsystems can be independently selected. The proposed approach is demonstrated via the application to two chemical process examples and one hybrid-tank plant.

In Chapter 4, distributed output-feedback fault detection and isolation (FDI) of nonlinear cascade process networks is investigated. Based on the assumption that an exponentially convergent estimator exists for each subsystem, a distributed state estimation system is developed. In the distributed state estimation system, a compensator is designed for each subsystem to compensate for subsystem interaction and the estimators for subsystems communicate to exchange information. It is shown that when there is no fault, the estimation error of the distributed estimation system converges to zero in the absence of system disturbances and measurement noise. For each subsystem, a state predictor is also designed to provide subsystem state predictions. A residual generator is designed for each subsystem based on subsystem state estimates given by the distributed state estimation system and subsystem state predictions given by the predictor. A subsystem residual generator generates two residual sequences, which act as references for FDI. A distributed FDI mechanism is proposed based on residuals. The proposed approach is able to handle both actuator faults and sensor faults by evaluating the residual signals. A chemical process example is introduced to demonstrate the effectiveness of the distributed FDI mechanism.

In Chapter 5, we investigate the subsystem decomposition problem for distributed state estimation of nonlinear systems. A systematic procedure for subsystem decomposition for distributed state estimation is proposed. Key steps in the procedure include observability test of the entire system, observable states identification for each output measurement, relative degree analysis and sensitivity analysis between measured outputs and states. Considerations with respect to time-scale multiplicity and direct graph are discussed. A few examples are used to illustrate the applicability of the methods used in different steps. The effectiveness

of the entire distributed state estimation configuration procedure is also demonstrated via an application to a chemical process example used in coal handling and preparation plants.

In Chapter 6, we propose a subsystem decomposition approach and a distributed moving horizon estimation (MHE) method for wastewater treatment plants. The plant (described by Benchmark Simulation Model No.1) is decomposed into smaller subsystems based on structural closeness. Three subsystems are formed considering subsystem interaction and nonlinearity of the subsystems. An iterative distributed MHE scheme is proposed for the wastewater treatment plant. Innovation triggered evaluation of the local MHEs is used to reduce the computational complexity of the estimation scheme. Extensive simulations are performed to illustrate the effectiveness and applicability of the proposed subsystem decomposition and distributed estimation methods.

In Chapter 7, we address the problem of subsystem decomposition of general nonlinear process networks for simultaneous distributed state estimation and distributed control based on community structure detection. A systematic procedure based on modularity is proposed. A fast folding algorithm which approximately maximizes the modularity is used in the proposed procedure to find candidate subsystem configurations. Two criteria are proposed to determine the final subsystem configurations that are appropriate for simultaneous distributed estimation and control. Two chemical process examples of different complexities are used to illustrate the effectiveness and applicability of the proposed approach.

Chapter 8 summaries the contributions of this work and discusses future research directions.

Chapter 2

Distributed moving horizon state estimation of two-time-scale nonlinear systems

In this chapter, we focus on distributed moving horizon estimation (DMHE) for a class of two-time-scale nonlinear systems described in the framework of singularly perturbed systems. In Section 2.2, by taking advantage of the time-scale separation property, a two-time-scale system is first decomposed into a reduced-order fast system and a reduced-order slow system. The slow system is further decomposed into several interconnected slow subsystems. In Section 2.3, a local estimator is designed for each slow subsystem and for the reduced-order fast system. The slow subsystem estimators communicate with each other to exchange information and they are only required to send information to the fast system one-directionally. The fast system estimator does not send out any information. The local estimators are designed as observer-enhanced moving horizon estimators. Sufficient conditions on the convergence of the estimation error of the DMHE are derived and are given in Section 2.4. The application of the proposed DMHE to a chemical process example demonstrates its applicability and effectiveness via simulations in Section 2.5. This chapter is a revised version of [42].

2.1 Introduction

Complex and integrated systems are common occurrences in manufacturing industries (e.g., chemicals, petrochemicals and mineral processes). Model predictive control (MPC)) systems are widely used in the manufacturing industries to ensure the quality of products while maximizing economic profits and guaranteeing operation safety as well as environmental sustainability. Due to the medium to large scales of many systems, the centralized control framework is not practical in terms of computational burden, organizational complexity, and fault tolerance [1]. The above considerations motivate significant research interests in distributed MPC [1]. While there are extensive results on distributed MPC, less attention has been given to distributed or decentralized state estimation which is equally important and is closely related to distributed control. It should be pointed out that there are some algorithms on decentralized or distributed Kalman filtering (e.g, [11]). However, these algorithms typically do not account for system nonlinearity.

Recently, there are some results on distributed moving horizon state estimation (DMHE). Moving horizon estimation (MHE) is an online optimization-based technique and can handle nonlinearities, constraints and optimality considerations [12, 13]. In [14], a DMHE algorithm was developed for nonlinear systems based on subsystem models. An observer-enhanced DMHE algorithm was developed in [15], where an auxiliary nonlinear observer is taken advantage of in the design of each local estimator. The auxiliary observer is used to calculate a reference state estimate based on which a confidence region is constructed every sampling time. Each local estimator optimizes its estimate within the confidence region. The observer-enhanced design is less sensitive to external noise compared with the auxiliary nonlinear observer. The convergence rate of the DMHE may be tuned by tuning the auxiliary observers. It was shown to be less dependent on the arrival cost, the estimation window size and have the potential to be used in output feedback control with provable closed-loop stability [15, 16].

On the other hand, time-scale multiplicity is a common feature of many systems. For chemical processes, it usually arises due to the strong coupling of physicochemical phenomena

[17, 18]. A direct application of standard control or estimation methods without taking into account time-scale multiplicity to systems with different time scales may lead to ill-conditioning or even the loss of closed-loop stability [19, 20]. The singular perturbation theory is the standard tool for the analysis of systems with time-scale multiplicity [19, 20]. Within the singular perturbation framework, the original system is typically decomposed into reduced-order subsystems with "fast" and "slow" dynamics. The majority of related results are on control system design for two-time-scale systems (e.g., [19, 20, 21]). Little attention has been given to state estimation of systems with time-scale multiplicity. In [22], state estimation of a wastewater treatment plant was addressed via linearization in a centralized framework by neglecting the fast dynamics.

In this work, the scope is on the handling of time-scale multiplicity in state estimation. Specifically, we consider state estimation of a class of two-time-scale nonlinear systems. A system is first decomposed into a reduced-order fast system and several reduced-order slow subsystems. A fast MHE is designed for the fast system and a slow MHE is designed for each slow subsystem. The fast and slow MHEs form a DMHE scheme. Each MHE is designed via the observer-enhanced MHE technique [24]. It is discovered that the slow MHEs are entirely decoupled from the fast MHE which is a significant difference from control of two-time-scale systems. The decoupling ensures that only one-directional information transmission from the slow MHEs to the fast MHE is needed and the fast MHE does not send out any information. Sufficient conditions are derived under which the proposed DMHE is guaranteed to give ultimately bounded estimation error under bounded system disturbances and measurement noise. The effectiveness of the proposed method is demonstrated via the application to a chemical process.

2.2 Preliminaries

2.2.1 Notation

The operator $|\cdot|$ denotes the Euclidean norm of a vector and $|\cdot|_Q^2$ represents the square of the weighted Euclidean norm of a vector, defined as $|x|_Q^2 = x^T Q x$ where Q is a positive definite matrix. A function $f(x)$ is said to be Lipschitz with respect to its argument x if there exists a positive constant L_f^x such that $|f(x') - f(x'')| \leq L_f^x |x' - x''|$ holds for all x' and x'' in a given region of x and L_f^x is the associated Lipschitz constant. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and satisfies $\alpha(0) = 0$. A function $\beta(r, s)$ is said to be a class $\mathcal{KL}$ function if for each fixed s, $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to r, and for each fixed r, it is deceasing with respect to s, and $\beta(r, s) \to 0$ as $s \to \infty$. A function f on an interval is said to be concave if for any x and y in the interval and for any $\alpha \in [0, 1]$, $f((1-\alpha)x + \alpha y) \geq (1-\alpha)f(x) + \alpha f(y)$. The symbol $\text{diag}(v)$ denotes a diagonal matrix, in which the diagonal elements are the elements of vector v. The symbol A^+ denotes the pseudoinverse of a matrix (or vector) A. $\mathbb{I}$ denotes a set of integers defined as $\mathbb{I} = \{1, \ldots, m\}$.

2.2.2 System description

In this study, we consider a class of two-time-scale nonlinear systems that can be described in the framework of singularly perturbed systems as follows:

$$\dot{x}_s(t) = f(x_s(t), w_s(t), \epsilon) + \tilde{f}(x_s(t), x_f(t), \epsilon) \tag{2.1a}$$

$$\epsilon \dot{x}_f(t) = g(x_f(t), w_f(t), \epsilon) + \tilde{g}(x_s(t), x_f(t), \epsilon) \tag{2.1b}$$

$$y_s(t) = h_s(x_s(t)) + v_s(t) \tag{2.1c}$$

$$y_f(t) = h_f(x_f(t)) + v_f(t) \tag{2.1d}$$

where $x_s \in \mathbb{R}^{n_{x_s}}$ and $x_f \in \mathbb{R}^{n_{x_f}}$ are state vectors, $w_s \in \mathbb{R}^{n_{w_s}}$ and $w_f \in \mathbb{R}^{n_{w_f}}$ denote system disturbances, $y_s \in \mathbb{R}^{n_{y_s}}$ and $y_f \in \mathbb{R}^{n_{y_f}}$ are system outputs, $v_s \in \mathbb{R}^{n_{v_s}}$ and $v_f \in \mathbb{R}^{n_{v_f}}$ denote measurement noise, and ϵ is a small positive parameter reflecting the time-scale separation in the dynamics of the nonlinear system. The functions f and g depict, respectively, the dependence of the dynamics of x_s and x_f on themselves and associated system disturbances. The function $\tilde{f}$ characterizes the interaction between the dynamics of x_s and the state vector x_f. Similarly, $\tilde{g}$ depicts the interaction between the dynamics of x_f and x_s. It is assumed that functions f, g, $\tilde{f}$ and $\tilde{g}$ are all locally Lipschitz with respect to their arguments. Note that locally Lipschitz is a mild assumption on the continuity of the functions and it imposes limits on how fast the functions can change. The small parameter ϵ appears as a multiplier of the time derivative of state x_f, and the state x_f evolves much faster than the state x_s [135]. We will refer to x_s as the slow states and x_f as the fast states in the remainder. We assume that the measurements y_s and y_f are continuously available.

2.2.3 Two-time-scale decomposition

It is possible to decompose two-time-scale systems described in (1) into two separate reduced-order systems evolving in a fast and a slow time scales. This property will be taken advantage of in the design of the proposed distributed state estimation scheme.

First, we set $\epsilon = 0$ in (2.1) and obtain that:

$$\frac{dx_s(t)}{dt} = f(x_s(t), w_s(t), 0) + \tilde{f}(x_s(t), x_f(t), 0) \tag{2.2a}$$

$$0 = g(x_f(t), w_f(t), 0) + \tilde{g}(x_s(t), x_f(t), 0) \tag{2.2b}$$

We assume that there exists a unique isolated solution to the algebraic equation (2.2b):

$$x_f(t) = \hat{g}(x_s(t), w_f(t)) \tag{2.3}$$

for each pair of (x_s, w_f), and the partial derivatives $\partial\hat{g}/\partial x$ and $\partial\hat{g}/\partial w$ are sufficiently smooth. This assumption is a standard one in two-time-scale decomposition and is used to ensure that x_f can be uniquely expressed in terms of x_s and w_f [19]. Note that a control system is normally operated within an operating range and the assumption does not impose practical restrictions. Substituting (2.3) into (2.2a), the reduced-order slow system is obtained as follows:

$$\dot{\bar{x}}_s(t) = f(\bar{x}_s(t), w_s(t), 0) + \tilde{f}(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0) \tag{2.4}$$

Note that in (2.4), $\bar{x}_s$ is used to denote the state of the reduced-order slow system to indicate that the dynamics of the reduced-order slow system is (slightly) different from the dynamics of x_s in the original system (2.1).

To derive the reduced-order fast system, we define a fast time scale $\tau = \frac{t}{\epsilon}$ and introduce the deviation variable $e_f := x_f - \hat{g}(x_s, w_f)$. The fast system (2.1b) can be rewritten in the following form: $\frac{de_f}{d\tau} = g\Big(e_f + \hat{g}(x_s, w_f), w_f, \epsilon\Big) + \tilde{g}\Big(x_s, e_f + \hat{g}(x_s, w_f), \epsilon\Big) - \epsilon\frac{\partial\hat{g}(x_s, w_f)}{\partial w_f}\dot{w}_f - \epsilon\frac{\partial\hat{g}}{\partial x_s}\Big(f(x_s, w_s, \epsilon) + \tilde{f}(x_s, e_f + \hat{g}(x_s, w_f), \epsilon)\Big)$. Setting ϵ to be zero and defining $G(e_f, x_s, w_f) := g(e_f + \hat{g}(x_s, w_f), w_f, 0) + \tilde{g}(x_s, e_f + \hat{g}(x_s, w_f), 0)$, the reduced-order fast system is obtained as follows:

$$\frac{de_f(\tau)}{d\tau} = G(e_f, x_s, w_f) \tag{2.5}$$

2.2.4 Decomposition of slow dynamics

We consider that the slow dynamics can be further decomposed into several slow subsystems. However, we consider the fast dynamics as a whole and will design a single estimator for the entire fast system. The purpose of this approach is twofold: (a) to simplify the presentation of the proposed state estimation scheme since the subsystem decomposition can be applied to the fast system in a straightforward manner; and (b) to be consistent with the fact that typically only a relatively smaller portion of the entire states exhibits fast dynamics in chemical processes.

We consider that the slow dynamics can be decomposed into m slow subsystems with the i-th slow subsystem, $i = 1, \ldots, m$, described as follows:

$$\dot{x}_{si}(t) = f_i(x_{si}(t), w_{si}(t), \epsilon) + \tilde{f}_i(x_s(t), x_f(t), \epsilon) \tag{2.6a}$$

$$y_{si}(t) = h_{si}(x_{si}(t)) + v_{si}(t) \tag{2.6b}$$

where $x_{si} \in \mathbb{R}^{n_{x_{si}}}$ denotes the vector of states of slow subsystem i, $w_{si} \in \mathbb{R}^{n_{w_{si}}}$ denotes the disturbance associated with slow subsystem i. The function f_i is used to describe the dependence of the state vector of each subsystem on itself and the associated disturbance. The vector function $\tilde{f}_i$ characterizes the interaction between slow subsystem i and other slow subsystems as well as the fast system. $y_{si} \in \mathbb{R}^{n_{y_{si}}}$ is the vector of measured outputs of slow subsystem i and $v_{si} \in \mathbb{R}^{n_{v_{si}}}$ denotes the associated measurement noise. x_{si}, $i \in \mathbb{I}$, and x_f are assumed to be constrained as $x_{si} \in \mathbb{X}_{si}$, $x_f \in \mathbb{X}_f$ where $\mathbb{X}_{si}$, $i \in \mathbb{I}$, and $\mathbb{X}_f$ are convex compact sets. The system disturbances are constrained as $w_{si} \in \mathbb{W}_{si}$, $i \in \mathbb{I}$, and $w_f \in \mathbb{W}_f$, while the measurement noise is bounded as $v_{si} \in \mathbb{V}_{si}$ and $v_f \in \mathbb{V}_f$. The sets $\mathbb{W}_{si}$, $\mathbb{W}_f$, $\mathbb{V}_{si}$, $\mathbb{V}_f$ with $i \in \mathbb{I}$ are defined such that $|w_{si}| \leq \theta_{w_{si}}$, $|w_f| \leq \theta_{w_f}$, $|v_{si}| \leq \theta_{v_{si}}$ and $|v_f| \leq \theta_{v_f}$, where $\theta_{w_{si}}$, θ_{w_f}, $\theta_{v_{si}}$, and θ_{v_f} with $i \in \mathbb{I}$ are known positive real scalars. Note that the assumption of boundedness of subsystem states is based on the fact that many systems are regulated by control systems and are operated within specific bounded regions. Note also that assumption of bounded system disturbances and measurement noise is a common assumption and is needed to obtain deterministic results. Based on system (2.4) for the reduced-order slow system, the corresponding reduced-order slow subsystem i is:

$$\dot{\bar{x}}_{si}(t) = f_i(\bar{x}_{si}(t), w_{si}(t), 0) + \tilde{f}_i(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0) \tag{2.7}$$

where $\bar{x}_{si}$ denotes the state vector of slow subsystem i.

2.2.5 Nonlinear observers

In the proposed DMHE scheme, the local MHEs will be designed as observer-enhanced MHE developed in [24]. Note that the design of the auxiliary nonlinear observers is not within the scope of this work and there are extensive studies on the design of nonlinear observers (e.g., [25, 55, 27, 28]) with many successful applications to different areas including chemical processes (e.g., [55, 29]).

Consider the nominal reduced-order slow subsystem i without considering the interaction as follows:

$$\begin{aligned} \dot{\bar{x}}_{si}(t) =& f_i(\bar{x}_{si}(t), 0, 0) \\ \bar{y}_{si}(t) =& h_{si}(\bar{x}_{si}(t)) \end{aligned} \tag{2.8}$$

It is assumed that there exists a nonlinear observer for the subsystem (2.8) for $i \in \mathbb{I}$, as follows:

$$\dot{z}_{si}(t) = F_{si}(z_{si}(t), h_{si}(\bar{x}_{si}(t)),\ i \in \mathbb{I} \tag{2.9}$$

such that z_{si} asymptotically converges to $\bar{x}_{si}$ for $\bar{x}_{si} \in \mathbb{X}_{si}$, if $\tilde{f}_i\left(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0\right) \equiv 0$ and $w_{si}(t) \equiv 0$.

Next, consider the nominal reduced-order fast system:

$$\begin{aligned} \frac{de_f(\tau)}{d\tau} =& G(e_f, x_s, 0) \\ y_f(\tau) =& h_f(e_f(\tau) + \hat{g}(x_s(\tau)) \end{aligned} \tag{2.10}$$

It is also assumed that there exists a nonlinear observer for the the nominal system (2.10) in the following form:

$$\dot{z}_f(\tau) = F_f(z_f(\tau), x_s(\tau), h_f(e_f(\tau) + \hat{g}(x_s(\tau), 0))) \tag{2.11}$$

such that z_f asymptotically approaches e_f if $w_f \equiv 0$ for all τ. These assumptions imply that if we have $\tilde{f}_i\left(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0\right) \equiv 0$, $w_{si}(t) \equiv 0$ for all t and $w_f(t) \equiv 0$ for all τ, then

there exists a set of functions β_{si}, $i \in \mathbb{I}$, and β_f of class $\mathcal{KL}$ such that:

$$|z_{si}(t) - \bar{x}_{si}(t)| \leq \beta_{si}(|z_{si}(0) - \bar{x}_{si}(0)|, t),\ i \in \mathbb{I} \tag{2.12a}$$

$$|z_f(\tau) - e_f(\tau)| \leq \beta_f(|z_f(0) - e_f(0)|, \tau) \tag{2.12b}$$

where $z_{si}(0)$, $z_f(0)$ and $\bar{x}_{si}(0)$, $e_f(0)$ denote the initial conditions. It is assumed that F_{si}, $i \in \mathbb{I}$, and F_f are locally Lipschitz functions. It is also assumed that the interactions between the slow subsystems and the interactions between the slow and fast systems do not damage the collective observability of the subsystems/systems. The above assumptions on the existence of local auxiliary observers imply that the state of each subsystem/system is locally observable if the interaction is known and the entire system is locally observable.

2.3 DMHE design

A schematic of the proposed DMHE scheme is presented in Fig. 2.1. A local MHE is designed for each slow subsystem and the fast system. The MHEs associated with the slow subsystems communicate with each other and send information to the MHE of the fast system. However, there is no information transmitted from the MHE of the fast system to any of the MHEs of the slow subsystems. In the remainder, the MHE associated with the i-th slow subsystem will be referred to as s-MHE i, $i \in \mathbb{I}$, while the MHE associated with the fast system will be referred to as f-MHE. Since the states of a two-time-scale system evolve at different time scales, it is desirable to use different sampling periods in the local MHE designs for the slow and fast systems. In this work, we use Δ_s and Δ_f to denote the sampling periods for s-MHEs and f-MHE, respectively. Without loss of generality, we assume that Δ_s is integer multiple of Δ_f, i.e., $\Delta_s = n\Delta_f$ where n is a positive integer. In the description of the proposed design, we use $t_k := t_0 + k\Delta_s$ with $k \geq 0$ and $\tau_q := \tau_0 + q\Delta_f$ with $q \geq 0$ to denote the sampling instants of s-MHEs and f-MHE, respectively.

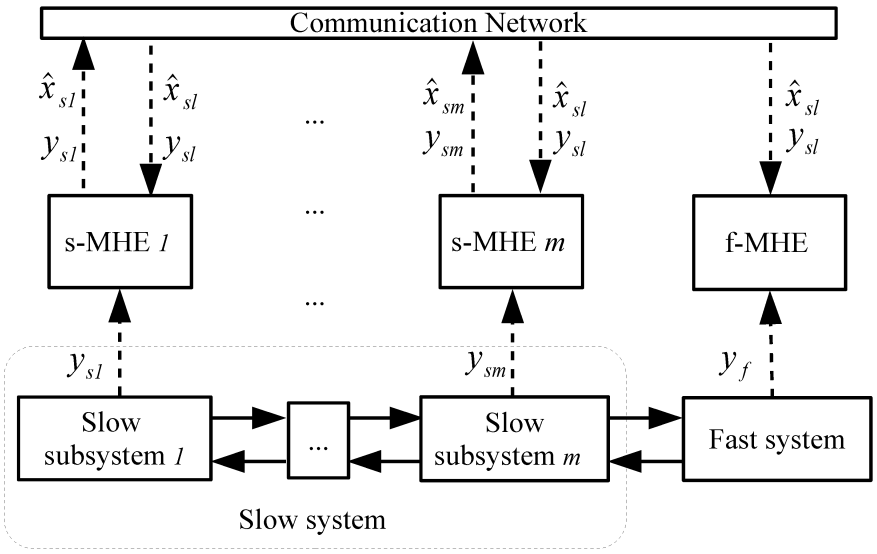

Figure 2.1: A schematic of the proposed distributed state estimation scheme (dashed lines indicate information flows and solid lines indicate interactions between subsystems).

2.3.1 Distributed state estimation algorithms

In the proposed DMHE scheme, the s-MHEs and the f-MHE are implemented independently except that the s-MHEs need to send information to the f-MHE. The implementation details of the DMHE are specified in the following two algorithms.

Algorithm 1. *Implementation of the s-MHEs.*

1. *At $t_0 = 0$, each s-MHE is initialized with output measurements $y_{sl}(0)$ and initial guesses $\hat{x}_{sl}(0)$, $l \in \mathbb{I}$, of all slow subsystems.*

2. *At $t_k > 0$, perform the following steps:*

   2.1. *Each s-MHE is updated with the output measurement of the corresponding slow subsystem; that is, s-MHE i receives $y_{si}(t_k)$, $i \in \mathbb{I}$.*

   2.2. *Each s-MHE requests and receives the output measurements and subsystem state estimates of the previous time instant from all other slow subsystems; that is, s-MHE i ($i \in \mathbb{I}$) receives $y_{sl}(t_{k-1})$ and $\hat{x}_{sl}(t_{k-1})$, $l \in \mathbb{I} \setminus \{i\}$.*

   2.3. *Based on the received information from the local subsystem as well as other slow*

subsystems, each s-MHE calculates the state estimate of the corresponding slow subsystem and sends the estimate to the f-MHE; that is, s-MHE i calculates $\hat{x}_{si}(t_k)$, $i \in \mathbb{I}$, and sends it to the f-MHE.

2.4. Go to Step 2.1 at the next slow system sampling time t_{k+1}.

Algorithm 2. *Implementation of the f-MHE.*

1. *At $\tau_0 = 0$, the f-MHE is initialized with the output measurements of the slow subsystems $y_{sl}(0)$, $l \in \mathbb{I}$, the output measurement of the fast system $y_f(0)$, the initial guesses of the s-MHEs $\hat{x}_{sl}(0)$, $l \in \mathbb{I}$, and the initial guess of the f-MHE $\hat{x}_f(0)$.*
2. *At $\tau_q > 0$, carry out the following steps:*

   2.1. The f-MHE receives the measured output of the fast system; that is, the f-MHE receives $y_f(\tau_q)$.

   2.2. Based on the latest received information of the slow subsystems and the fast system, the f-MHE calculates the estimate of the state of the fast system $\hat{x}_f(\tau_q)$.

   2.3. Go to Step 2.1 at the next fast system sampling time τ_{q+1}.

From the two algorithms, it can be seen that the s-MHEs communicate with themselves and send out information to the f-MHE while the f-MHE does not send out any information to the s-MHEs.

Remark 1. *Note that in the proposed design, the s-MHEs are decoupled from the f-MHE in the sense that the performance of the s-MHEs does not depend on the f-MHE. It allows us to treat slow and fast estimators essentially separately. This property is also different from most of the results in the control of two-time-scale systems in which the controllers of the fast dynamics and of the slow dynamics are normally mutually coupled via system states because control actions are fedback to the system. Note also that the s-MHEs communicate with themselves and send out information to the f-MHE. The communicated information is*

used in interaction compensation. The distributed estimators are coupled via information flow which renders the overall design a distributed state estimation scheme.

2.3.2 Reference state estimate calculation for s-MHEs

In this subsection, we show how to augment nonlinear observers (2.9) to account for the interactions between slow subsystems. The augmented nonlinear observers will be used to generate reference state estimates for s-MHEs every slow sampling period. The augmented observer i at time t_k is designed as follows:

$$\begin{aligned}\dot{z}_{si}(t) = F_{si}(z_{si}(t), y_{si}(t_{k-1})) + \tilde{f}_i(\hat{x}_s(t_{k-1}), \hat{g}(\hat{x}_s(t_{k-1}), 0), 0) \\ + \sum_{l\in\mathbb{I}} K_{si,l}(\hat{x}_{sl})(y_{sl}(t_{k-1}) - h_{sl}(\hat{x}_{sl}(t_{k-1})))\end{aligned} \tag{2.13}$$

At t_k, the augmented observer is initialized as $z_{si}(t_{k-1}) = \hat{x}_{si}(t_{k-1})$, where $\hat{x}_{si}(t_{k-1})$, $i \in \mathbb{I}$, is the optimal state estimate of $x_{si}(t_{k-1})$ obtained by s-MHE i at t_{k-1} and $K_{si,l}$, $l \in \mathbb{I}$, are gain matrices designed as follows:

$$K_{si,l} = \frac{\partial \tilde{f}_i}{\partial \bar{x}_{sl}} \left(\frac{\partial h_{sl}}{\partial \bar{x}_{sl}}\right)^{+}\Bigg|_{\hat{x}_{sl}(t_{k-1})}, \ \forall l \in \mathbb{I} \tag{2.14}$$

We note that the augmented observer is evaluated prior to the evaluation of s-MHE i, $i \in \mathbb{I}$, in order to generate a reference state estimate at each sampling time. $\hat{x}_s(t_{k-1})$ is an approximation of $x_s(t)$ for $t \in [t_{k-1}, t_k)$ in (2.13).

2.3.3 Design of s-MHEs

Based on the reference state estimate $z_{si}(t_k)$ given by observer (2.13), the proposed s-MHE i at t_k is designed as:

$$\min_{\tilde{x}_{si}(t_{k-N_s}),\ldots,\tilde{x}_{si}(t_k)} \left\{ \sum_{j=k-N_s}^{k-1} |w_{si}(t_j)|^2_{Q_{si}^{-1}} + \sum_{j=k-N_s}^{k} |v_{si}(t_j)|^2_{R_{si}^{-1}} \right\} \tag{2.15a}$$

$$\begin{aligned} \text{s.t. } \dot{\tilde{x}}_{si}(t) &= f_i(\tilde{x}_{si}(t), w_{si}(t_j), 0) \\ &\quad + \tilde{f}_i(\hat{x}_s(t_j), \hat{g}(\hat{x}_s(t_j), 0), 0), \\ &\quad t \in [t_j, t_{j+1}],\ j = k - N_s, \ldots, k-1 \end{aligned} \tag{2.15b}$$

$$v_{si}(t_j) = y_{si}(t_j) - h_{si}(\tilde{x}_{si}(t_j)),\ j = k - N_s, \ldots, k \tag{2.15c}$$

$$\begin{aligned} &w_{si}(t_j) \in \mathbb{W}_{si},\ v_{si}(t_j) \in \mathbb{V}_{si},\ \tilde{x}_{si}(t_j) \in \mathbb{X}_{si}, \\ &\quad j = k - N_s, \ldots, k-1 \end{aligned} \tag{2.15d}$$

$$|\tilde{x}_{si}(t_k) - z_{si}(t_k)| \leq \kappa_{si} |y_{si}(t_k) - h_{si}(z_{si}(t_k))| \tag{2.15e}$$

where $\tilde{x}_{si}$ denotes the prediction of x_{si} within the optimization problem; Q_{si} and R_{si} represent the covariance matrices of w_{si} and v_{si}, respectively; N_s is the estimation horizon; $\hat{x}_s$ denotes the optimal estimate of x_s obtained by the s-MHEs and κ_{si}, $i \in \mathbb{I}$, is a positive constant. The design of κ_{si} will be made clear in following discussion.

Once the optimization problem (2.15) is solved, the optimal solution is denoted as $\{\tilde{x}^*_{si}(t_{k-N}),\ \ldots,\ \tilde{x}^*_{si}(t_k)\}$, of which the last element $\tilde{x}^*_{si}(t_k)$ is adopted as the optimal estimate of $x_{si}(t_k)$ and is denoted as $\hat{x}_{si}(t_k)$; that is, $\hat{x}_{si}(t_k) = \tilde{x}^*_{si}(t_k)$. In (2.15), (2.15a) is the cost function for s-MHE i to be minimized. (2.15b) and (2.15c) describe the reduced-order slow subsystem i. $\hat{x}_s(t_j)$ approximates $x_s(t)$ for $t \in [t_j, t_{j+1})$ to compensate for the interactions in the system. Constraint (2.15d) explicitly bounds system disturbance w_{si} and measurement noise v_{si}. Constraint (2.15e) is utilized to calculate a confidence region, within which the estimate of subsystem state is optimized.

The feasibility of the optimization problem (2.15) is guaranteed since the two constraints

(2.15d) and (2.15e) are not imposed simultaneously and do not conflict with each other. Note also that constraint (2.15e) takes into account the boundedness properties of the system and measurement noise via the confidence region constructed via the right-hand-side of the inequality. Note also that the use of reduced-order slow subsystems in the design can avoid ill-conditioning of the estimation problem as recognized in control relevant literature (e.g., [19, 20, 21]).

2.3.4 Reference state estimate calculation for f-MHE

For the reduced-order fast system, observer (2.11) is used to calculate a reference state estimate every fast sampling time. Specifically, at τ_q, the observer takes the following form:

$$\dot{z}_f(\tau) = F_f(z_f(\tau), \hat{x}_s(\tau_{q-1}), y_f(\tau_{q-1})) \tag{2.16a}$$

$$z_f(\tau_{q-1}) = \hat{e}_f(\tau_{q-1}) \tag{2.16b}$$

where $\hat{x}_s(\tau_{q-1})$ is an approximation of $x(\tau)$ for $\tau \in [\tau_{q-1}, \tau_q)$. Note that $\hat{x}_s(\tau_{q-1})$ indicates the latest available information on $\hat{x}_s$ at each fast sampling instant τ_q. It does not necessarily mean the actual $\hat{x}_s$ at τ_{q-1} because s-MHEs do not update at every fast sampling time.

2.3.5 Design of f-MHE

In the f-MHE design, we do not estimate the original fast state x_f directly. Instead, we obtain an optimal estimate of the reduced-order fast system state e_f, based on which an estimate of x_f is calculated. Based on the reference state estimate $z_f(\tau_q)$ generated by observer (2.16),

the proposed f-MHE at τ_q is designed as follows:

$$\min_{\tilde{e}_f(\tau_{q-N_f}),\ldots,\tilde{e}_f(\tau_q)} \left\{ \sum_{j=q-N_f}^{q-1} |w_f(\tau_j)|^2_{Q_f^{-1}} + \sum_{j=q-N_f}^{q} |v_f(\tau_j)|^2_{R_f^{-1}} \right\} \tag{2.17a}$$

$$\text{s.t.} \quad \dot{\tilde{e}}_f(\tau) = G(\tilde{e}_f(\tau), \hat{x}_s(\tau_j), w_f(\tau_j)),$$

$$\tau \in [\tau_j, \tau_{j+1}],\ j = q - N_f, \ldots, q-1 \tag{2.17b}$$

$$v_f(\tau_j) = y_f(\tau_j) - h_f(\tilde{e}_f(\tau_j) + \hat{g}(\hat{x}_s(\tau_j), w_f(\tau_j))),$$

$$j = q - N_f, \ldots, q \tag{2.17c}$$

$$w_f(\tau_j) \in \mathbb{W}_f,\ v_f(\tau_j) \in \mathbb{V}_f,\ j = q - N_f, \ldots, q-1 \tag{2.17d}$$

$$|\tilde{e}_f(\tau_q) - z_f(\tau_q)| \leq \kappa_f |y_f(\tau_q) - h_f(z_f(\tau_q) + \hat{g}(\hat{x}_s(\tau_q), w_f(\tau_{q-1})))| \tag{2.17e}$$

where $\tilde{e}_f$ denotes the estimate of e_f within the optimization problem; Q_f and R_f denote the covariance matrices of w_{si} and v_{si}, respectively; N_f is the estimation horizon of f-MHE and κ_f is a positive constant.

The optimal solution to (2.17) is denoted as $\{\tilde{e}_f^*(\tau_{q-N_f}),$ $\ldots, \tilde{e}_f^*(\tau_q)\}$, and the last element is the optimal state estimate for time instant τ_q. That is, $\hat{e}_f(\tau_q) = \tilde{e}_f^*(\tau_q)$. In the optimization problem (2.17), (2.17a) is the cost function for f-MHE to be minimized. Constraints (2.17b) and (2.17c) are the reduced-order fast system (2.5). $\hat{x}_s(\tau_j)$ denotes the latest information of $\hat{x}_s$ available at τ_j. (2.17d) describes the constraints on w_f and v_f. (2.17e) is used to calculate a confidence region, and the estimate of the fast system state is only allowed to be optimized within the region. Note that in (2.17), we assume w_f takes piece-wise constant values between two fast sampling instants. This is the reason that $w_f(\tau_{q-1})$ is used in (2.17e). Note also that the optimal solution to (2.17), $\{\tilde{e}_f^*(\tau_{q-N_f}), \ldots, \tilde{e}_f^*(\tau_q)\}$, is associated with an optimal disturbance trajectory $\{w_f^*(\tau_{q-N_f}), \ldots, w_f^*(\tau_{q-1})\}$. Based on the optimal estimate $\hat{e}_f(\tau_q)$ and $w_f^*(\tau_{q-1})$, the optimal estimate of the original fast system state x_f is calculated as $\hat{x}_f(\tau_q) = \hat{e}_f(\tau_q) +$

$\hat{g}(\hat{x}_s(\tau_q), w_f^*(\tau_{q-1}))$.

2.4 Stability analysis

In this section, we perform stability analysis of the proposed distributed state estimation scheme. We first focus on the stability of the s-MHEs and then on the f-MHE.

2.4.1 Stability analysis of s-MHEs

We define a function $\gamma_i(s)$ for each $i \in \mathbb{I}$ as $\gamma_i(s) := (L_{\tilde{f}_i}^{\hat{g}} L_{\hat{g}}^{w_f} \theta_{w_f} + L_{F_{si}}^{y_{si}} \theta_{v_{si}} + L_{f_i}^{w_{si}} \theta_{w_{si}} + \sum_{l\in\mathbb{I}} \mathcal{K}_{si,l} \theta_{v_{sl}})s + L_{F_{si}}^{y_{si}} L_{h_{si}} \bar{\mathcal{M}}_{si} s^2/2 + \sum_{l\in\mathbb{I}} (L_{f_i}^{x_{sl}} L_{\tilde{f}_i}^{\hat{g}} L_{\hat{g}}^{x_s}) \bar{\mathcal{M}}_{sl} s^2/2$, where $L_{\tilde{f}_i}^{\hat{g}}$ is the Lipschitz constant of $\tilde{f}_i$ with respect to its second argument, $L_{\hat{g}}^{w_f}$ is the Lipschitz constant of $\hat{g}$ with respect to its second argument, $L_{F_{si}}^{y_{si}}$ is the Lipschitz constant of F_{si} defined in (2.9) with respect to its second argument, $L_{f_i}^{w_{si}}$ is the Lipschitz constant of f_i with respect to its second argument, $L_{h_{si}}$ is the Lipschitz constant of h_{si} with respect to its argument, $L_{f_i}^{x_{sl}}$ is the Lipschitz constant of f_i with respect to its first argument, $L_{\hat{g}}^{x_s}$ is the Lipschitz constant of $\hat{g}$ with respect to its first argument, $\mathcal{K}_{si,l}$ $\forall i, l \in \mathbb{I}$, is a positive constant such that $|K_{si,l}| \leq \mathcal{K}_{si,l}$, $\bar{\mathcal{M}}_{s,i}$ is a finite constant denoting the upper bound of $|\dot{\hat{x}}_{si}|$ (i.e. $|\dot{\hat{x}}_{si}| \leq \bar{\mathcal{M}}_{si}$), $\forall i \in \mathbb{I}$. Note that subsystem states are within convex compact sets, and the Lipschitz properties are assumed to be held in these sets and the operating regions of the observers locally.

Proposition 1. *[c.f. [15]] Consider the augmented observer (2.13) for the reduced-order slow subsystem i, $i \in \mathbb{I}$, during the time interval $t \in [t_k,\ t_{k+1}]$ with initial condition $z_{si}(t_k) = \bar{x}_{si}(t_k)$ and output measurement $y_{si}(t_k)$. If each gain matrix $K_{si,l}$ is determined following (2.14) and is bounded such that $|K_{si,l}| \leq \mathcal{K}_{si,l}$, $i \in \mathbb{I}$, then the estimation error e_{si}^z between the observer state z_{si} and the state of the reduced-order slow subsystem i in (2.7) in one slow sampling period Δ_s is bounded as follows:*

$$|e_{si}^z(t_{k+1})| \leq \beta_{si}(|e_{si}^z(t_k)|, \Delta_s) + \gamma_i(\Delta_s) + \sum_{l\in\mathbb{I}} L_{i,l} |e_{sl}^z(t_k)|^2 \Delta_s \tag{2.18}$$

where $e_{si}^z = z_{si} - \bar{x}_{si}$, $L_{i,l} = (\mathcal{H}_l^{\tilde{f}_i} + \mathcal{K}_{si,l}\mathcal{H}_l^{h_s})$ *with* $\mathcal{H}_l^{\tilde{f}_i}$, *and* $\mathcal{H}_l^{h_s}$, *for all* $i, l \in \mathbb{I}$, *being finite positive constants that are associated with the second-order terms generated by Taylor expansions of* $\tilde{f}_i$ *and* h_s, *respectively.*

Proposition 2. *Consider the reduced-order slow subsystem* i *described by (2.7). The deviation of the state* $\bar{x}_{si}$ *of the reduced-order slow subsystem* i *from the corresponding state* x_{si} *of the actual system described by (2.1) in one slow sampling period* Δ_s *is bounded as follows:*

$$|e_{si}^x(\Delta_s)| \leq \phi_i(\epsilon, \Delta_s) \tag{2.19}$$

where $e_{si}^x = \bar{x}_{si} - x_{si}$, $\phi_i(\epsilon, s) = ((L_{f_i}^\epsilon + L_{\tilde{f}_i}^\epsilon)\epsilon + 2\mathcal{N}_{\tilde{f}_i})s + \int_0^{\Delta_s} |f_i(x_{si}(t) + e_{si}^x(t), w_{si}(t), 0) - f_i(x_{si}(t), w_{si}(t), 0)|dt$, $L_{f_i}^\epsilon$ *and* $L_{\tilde{f}_i}^\epsilon$ *are Lipschitz constants of* f_i *and* $\tilde{f}_i$, *and* $\mathcal{N}_{\tilde{f}_i}$ *(*$i \in \mathbb{I}$*) is a positive constant such that* $|\tilde{f}_i(\cdot, \cdot, 0)| \leq \mathcal{N}_{\tilde{f}_i}$.

Proof: The time derivative of the error $e_{si}^x(t)$ is:

$$\begin{aligned}\dot{e}_{si}^x(t) =& f_i(\bar{x}_{si}(t), w_{si}(t), 0) + \tilde{f}_i(\bar{x}_s(t), \hat{g}(\bar{x}_s(t), w_f(t)), 0) \\ &- f_i(x_{si}(t), w_{si}(t), \epsilon) - \tilde{f}_i(x_s(t), x_f(t), \epsilon)\end{aligned}$$

Adding and subtracting $f_i(x_{si}(t), w_{si}(t), 0)$, and taking into account the Lipschitz properties of functions f_i and $\tilde{f}_i$, and the fact that $|\tilde{f}_i(\cdot, \cdot, 0)| \leq \mathcal{N}_{\tilde{f}_i}$, the following inequality can be obtained:

$$\begin{aligned}|\dot{e}_{si}^x(t)| \leq& |f_i(\bar{x}_{si}(t), w_{si}(t), 0) - f_i(x_{si}(t), w_{si}(t), 0)| \\ &+ (L_{f_i}^\epsilon + L_{\tilde{f}_i}^\epsilon)\epsilon + 2\mathcal{N}_{\tilde{f}_i}\end{aligned} \tag{2.20}$$

where $L_{f_i}^\epsilon$ and $L_{\tilde{f}_i}^\epsilon$ are the Lipschitz constants of f_i, $\tilde{f}_i$ with respect to ϵ. Given that $\bar{x}_{si} = x_{si} + e_{si}^x$ and the definition of function $\phi_i(\epsilon, s)$, Proposition 2 can be proved by integrating (2.20) from $t = 0$ to $t = \Delta_s$. $\square$

Now we are in a position to present sufficient conditions on the convergence and ultimate boundedness of the estimation error of the slow system with s-MHEs.

Theorem 1. *Consider system (2.1) with outputs y_{si}, $i \in \mathbb{I}$, of the corresponding slow subsystems at each time instant t_k, $k \geq 0$. If the developed s-MHEs are implemented following Algorithm 1 with s-MHE i, $i \in \mathbb{I}$, designed as in (2.15) based on nonlinear observers satisfying constraint (2.12a) and all the assumptions in Proposition 1 and Proposition 2 hold, and if $\phi_i(\epsilon, \Delta_s) \leq d_i(\epsilon)$ all the time for all $i \in \mathbb{I}$ and there exist a class of concave functions $\mathcal{Q}_{si}$, $i \in \mathbb{I}$, satisfying:*

$$\mathcal{Q}_{si}(s) \geq \beta_{si}(s, \Delta_s) \tag{2.21}$$

for all $0 \leq s \leq \rho_{si}$, and if there exist positive scalars ϱ_{si} satisfying $0 \leq \varrho_{si} \leq \rho_{si}$, and positive scalars $a_{si} \geq 1$, $b_{si} > 0$, $\xi_{si} > 0$ such that

$$\varrho_{si} - a_{si}(\mathcal{Q}_{si}(\varrho_{si}) + \gamma_i(\Delta_s) + \sum_{l \in \mathbb{I}} L_{i,l}\rho_{sl}^2 \Delta_s) - b_{si}\theta_{v_{si}} - d_i(\epsilon) \geq \xi_{si} \tag{2.22}$$

for all $i \in \mathbb{I}$, and if κ_{si} for all $i \in \mathbb{I}$ are picked such that $0 \leq \kappa_{si} \leq \min\{(a_{si}-1)/L_{h_{si}}, b_{si}\}$, then the estimation error $|e_{si}| = |\hat{x}_{si} - x_{si}|$, $i \in \mathbb{I}$, is ultimately bounded within a small region such that: $\lim_{t\to\infty} \sup|e_{si}(t)| \leq \rho_{si}^{min}$ for $|e_{si}^{\bar{x}}(0)| + |e_{si}^{x}(0)| \leq \rho_{si}$, $i \in \mathbb{I}$, where $\rho_{si}^{min} := \max\{|e_{si}^{\bar{x}}(t+\Delta_s)| + |e_{si}^{x}(t+\Delta_s)| : |e_{si}^{\bar{x}}(t)| + |e_{si}^{x}(t)| \leq \varrho_{si}\}$. This also implies that the estimation error of the entire slow system is ultimately bounded.

Proof: We first study the error between the estimated states $\hat{x}_{si}$ given by s-MHE i and the states of reduced-order slow subsystem i. This error is denoted as $e_{si}^{\bar{x}} = \hat{x}_{si} - \bar{x}_{si}$. Considering the constraint in (2.15e) and the Lipschitz property of function h_{si} and the upper bound $\theta_{v_{si}}$ for measurement noise v_{si}, we derive that

$$|\hat{x}_{si}(t_{k+1}) - z_{si}(t_{k+1})| \leq L_{h_{si}}\kappa_{si}|\bar{x}_{si}(t_{k+1}) - z_{si}(t_{k+1})| + \kappa_{si}\theta_{v_{si}} \tag{2.23}$$

Considering the fact that $|\hat{x}_{si} - \bar{x}_{si}| \leq |\hat{x}_{si} - z_{si}| + |z_{si} - \bar{x}_{si}|$, the fact that $e_{si}^{\bar{x}}(t_k) = e_{si}^{z}(t_k)$, and the property of class $\mathcal{KL}$ function β_{si}, we can obtain from (2.23) that

$$
\begin{aligned}
|e^{\bar{x}}_{si}(t_{k+1})| \leq (1+L_{h_{si}}\kappa_{si})\Big(&\beta_{si}\left(|e^{\bar{x}}_{si}(t_k)|+|e^{x}_{si}(t_k)|,\Delta_s\right)\\
&+\gamma_i(\Delta_s)+\sum_{l\in\mathbb{I}}L_{i,l}(|e^{\bar{x}}_{sl}(t_k)|+|e^{x}_{sl}(t_k)|)^2\Delta_s\Big)+\kappa_{si}\theta_{v_{si}}
\end{aligned}
\tag{2.24}
$$

Next, we focus on the actual estimation error $|e_{si}|= |x_{si}-\hat{x}_{si}|$ between the actual state of the original slow subsystem i (i.e., x_{si}) and the state estimate for the slow subsystem i (i.e., $\hat{x}_{si}$). If (2.21) holds, based on (2.24), Proposition 2 and the assumption that $\phi_i(\epsilon,\Delta_s)\leq d_i(\epsilon)$, it is derived that

$$
\begin{aligned}
&|e^{\bar{x}}_{si}(t_{k+1})|+|e^{x}_{si}(t_{k+1})|\\
&\quad\leq (1+L_{h_{si}}\kappa_{si})\Big(\mathcal{Q}_{si}\left(|e^{\bar{x}}_{si}(t_k)|+|e^{x}_{si}(t_k)|\right)+\gamma_i(\Delta_s)\\
&\qquad+\sum_{l\in\mathbb{I}}L_{i,l}(|e^{\bar{x}}_{sl}(t_k)|+|e^{x}_{sl}(t_k)|)^2\Delta_s\Big)+\kappa_{si}\theta_{v_{si}}+d_i(\epsilon)
\end{aligned}
\tag{2.25}
$$

for $|e^{\bar{x}}_{si}(t_k)|+|e^{x}_{si}(t_k)|\leq \rho_{si}$, $i\in\mathbb{I}$.

If there exist positive constants ϱ_{si}, ρ_{si} and ξ_{si} that satisfy (2.22), it can be obtained that:

$$
|e^{\bar{x}}_{si}(t_{k+1})|+|e^{x}_{si}(t_{k+1})|\leq |e^{\bar{x}}_{si}(t_k)|+|e^{x}_{si}(t_k)|-\xi_{si} \tag{2.26}
$$

for all $\varrho_{si}\leq |e^{\bar{x}}_{si}|+|e^{x}_{si}|\leq \rho_{si}$. Inequality (2.26) implies that $|e^{\bar{x}}_{si}|+|e^{x}_{si}|$ is decreasing if $\varrho_{si}\leq |e^{\bar{x}}_{si}|+|e^{x}_{si}|$ and will eventually become smaller than ϱ_{si} within finite sampling periods. Once $|e^{\bar{x}}_{si}|+|e^{x}_{si}|< \varrho_{si}$, $|e^{\bar{x}}_{si}|+|e^{x}_{si}|$ will not necessarily decrease at every sampling instant, but will be bounded within a small region such that $|e^{\bar{x}}_{si}|+|e^{x}_{si}|\leq \rho^{min}_{si}$ because of the definition of ρ^{min}_{si}.

Further, based on the triangular property $|e_{si}|\leq |e^{\bar{x}}_{si}|+|e^{x}_{si}|$, it is obtained that the estimation error $|e_{si}|$ is ultimately bounded as: $\lim_{t\to\infty}\sup|e_{si}(t)|\leq \rho^{min}_{si}$ after sufficient sampling periods. The ultimate boundedness of $|e_{si}|$, $i\in\mathbb{I}$, implies the ultimate boundedness of $|e_s|$. This proves Theorem 1. □

In condition (2.22) of Theorem 1, the function $\mathcal{Q}_{si}(\cdot)$ $(i\in\mathbb{I})$ is a concave approximation of $\beta_{si}(\cdot,\Delta_s)$ $(i\in\mathbb{I})$ which reflects the convergence rate of the associated auxiliary observer.

The term $\gamma_i(\Delta_s)$ $(i \in \mathbb{I})$ characterizes the effects of the slow sampling time Δ_s, system disturbances and partially measurement noise. The term $\sum_{l\in\mathbb{I}} L_{i,l}\rho_{si}^2\Delta_s$ $(i \in \mathbb{I})$ describes the effects caused by subsystem interactions after interaction compensation. The term $b_{si}\theta_{v_{si}}$ $(i \in \mathbb{I})$ characterizes the effect of measurement noise. And $d_i(\epsilon)$ $(i \in \mathbb{I})$ provides an upper bound on the model error between the actual slow subsystem i and the reduced-order slow subsystem i. Note that conditions (2.21) and (2.22) can be satisfied simultaneously as long as the convergence rates of the auxiliary observers are sufficiently large, the sampling time Δ_s and the measurement noise upper bound $\theta_{v_{si}}$ $(i \in \mathbb{I})$ and the time separation indication ϵ are sufficiently small. Note also that as long as $a_{si} \geq 1$ and $b_{si} > 0$, the parameter κ_{si} associated with s-MHE i $(i \in \mathbb{I})$ satisfying the condition stated in Theorem 1 can always be found. When bigger a_{si} and b_{si} exist, bigger κ_{si} can be picked which implies that the s-MHE i can optimize the estimate within a bigger region. An extreme case is $\kappa_{si} = 0$, which leads to $\tilde{x}_{si}^* = z_{si}$.

The assumption in Theorem 1 that $\phi_i(\epsilon, \Delta_s) \leq d_i(\epsilon)$ $(i \in \mathbb{I})$ imposes a restriction on the system. It requires that the difference between the reduced-order model and the original system should not grow unbounded. This assumption can normally be satisfied when the original and the reduced-order systems are stable and the value of ϵ is sufficiently small. In Theorem 1, the model-plant-mismatch between a reduced-order slow subsystem and the actual subsystem is explicitly considered. This can be seen from (2.22) in which $d_i(\epsilon)$ is involved. The model-plant-mismatch affects the size of the region in which the estimation error e_{si} is ultimately bounded. This can be seen from the definition of ρ_{si}^{min}.

Remark 2. *In Theorem 1, the assumption on the existence of a concave function satisfying (2.21) is made to simplify the proof of Theorem 1. The results can still be proved without this assumption but the proof would be more involved. Specifically, the class $\mathcal{KL}$ function $\beta_{si}(|e|, \Delta)$ is a strictly increasing function with respect to $|e|$. $\beta_{si}(|e|, \Delta)$ with fixed Δ could be a convex increasing function or a concave increasing function of $|e|$. It requires different approaches to handle the two difference cases. To simplify the proof and unify the conditions,*

the assumption on the existence of a concave function is introduced.

Remark 3. *In our work, an approach based on the worst case scenario is used in the derivation of the conditions in Theorem 1. In the analysis, a subsystem treats the estimation error of other interacting subsystems as external disturbances. Note that the error is compensated for using the information communicated in the calculation of reference state estimates in the augmented nonlinear observers. The augmented observers can completely remove the first- and second-order interaction effects (when expanded using Taylor series) between subsystems. This ensures that the worst case scenario approach is not very conservative. Also note that while the conditions in Theorem 1 (as well as Theorem 2) are stated in terms of subsystems, they are required to be satisfied simultaneously for all the subsystems. One implication of these conditions is that the connected distributed estimators subject to uncompensated high-order interaction effects, process and measurement noise and the model approximation error due to the reduced-order models should be convergent.*

2.4.2 Stability analysis of f-MHE

Next, we conduct stability analysis of the proposed f-MHE. We define $\mu(s) := n\sum_{l\in\mathbb{I}} L_{F_f}^{x_s}\mathcal{M}_{sl}s^2 + L_{F_f}^{y_f}L_{h_f}\mathcal{M}_f s^2/2 + \Big(L_{F_f}^{y_f}L_{h_f}L_{\hat{g}}^{w_f}\theta_{w_f} + L_{F_f}^{y_f}\theta_{v_f} + L_G^{w_f}\theta_f\Big)s + \sum_{l\in\mathbb{I}} L_{F_f}^{y_f}L_{h_f}L_{\hat{g}}^{x_s}\mathcal{M}_{sl}s^2/2$ where $L_{F_f}^{x_s}$ is the Lipschitz constant of F_f with respect to its second argument, $L_{F_f}^{y_f}$ is the Lipschitz constant of F_f with respect to its third argument, L_{h_f} is the Lipschitz constant of h_f with respect to its argument, $L_G^{w_f}$ is the Lipschitz constant of G with respect to its third argument, $\mathcal{M}_{sl}$ is a finite constant that denotes the upper bound of $\dot{x}_{sl}$ such that $|\dot{x}_{sl}|\leq \mathcal{M}_{sl}$ for all $x_{sl} \in \mathbb{X}_{sl}$ with $l \in \mathbb{I}$, $\mathcal{M}_f$ is a finite constant such that $|\dot{e}_f|\leq \mathcal{M}_f$.

Proposition 3. *Consider the nonlinear observer (2.16) designed for the reduced-order fast system (2.5), within one sampling period from τ_q to τ_{q+1}, the state deviation generated by*

the nonlinear observer for the fast system is bounded as follows:

$$|\varepsilon_{z_f}(\tau_{q+1})| < \beta_f(|\varepsilon_{z_f}(\tau_q)|, \Delta_f) + \sum_{l\in\mathbb{I}} L_{F_f}^{x_s}|e_{sl}(\tau_{q-n+1})|\Delta_f + \mu(\Delta_f) \tag{2.27}$$

where $\varepsilon_{z_f}(\tau) = z_f(\tau) - e_f(\tau)$.

Proof: In this proof, we focus on the time interval $\tau \in [\tau_q,\ \tau_{q+1}]$, f-MHE may not receive any update on $\hat{x}_s$ from the s-MHEs because of the difference in the sampling times. The latest received $\hat{x}_s$ will be used to approximate $\hat{x}_s(\tau_q)$. To analyze the effect induced by the approximation, we take into account the worst case scenario; that is, s-MHEs do not update for a period of $(n-1)\Delta_f$ from τ_{q-n+2} until τ_{q+1}. Let us consider the nonlinear observer designed for the local f-MHE of Eq. (2.17) corresponding to the reduced-order fast system, and define $\varepsilon_{z_f} = z_f - e_f$ as the estimation error generated by the nonlinear observer for the fast system. We consider the time interval from $\tau = \tau_q$ to $\tau = \tau_{q+1}$ with initial condition $z_f(\tau_q) = \hat{e}_f(\tau_q)$, the derivative of the estimation error ε_{z_f} is calculated as follows:

$$\dot{\varepsilon}_{z_f}(\tau) = F_f(z_f(\tau), \hat{x}_s(\tau_{q-n+1}), y_f(\tau_q)) - G(e_f(\tau), x_s(\tau), w_f(\tau)) \tag{2.28}$$

We can obtain an inequality from (2.28) by taking advantage of the Lipschitz properties of F_f and G, $|v_f(\tau)| \leq \theta_{v_f}$, $|w_f(\tau)| \leq \theta_{w_f}$. If we integrate the obtained inequality from $\tau = \tau_q$ to $\tau = \tau_{q+1}$, we can have that

$$\begin{aligned}
\left|\varepsilon_{z_f}(\tau_{q+1})\right| < \beta_f(|\varepsilon_{z_f}(\tau_q)|, \Delta_f) &+ \sum_{l\in\mathbb{I}} L_{F_f}^{x_s}|e_{sl}(\tau_{q-n+1})|\Delta_f \\
&+ n\sum_{l\in\mathbb{I}} L_{F_f}^{x_s}\mathcal{M}_{sl}\Delta_f^2 + L_{F_f}^{y_f}L_{h_f}\mathcal{M}_f\Delta_f^2/2 \\
&+ L_{F_f}^{y_f}L_{h_f}L_{\hat{g}}^{w_f}\theta_{w_f}\Delta_f + L_{F_f}^{y_f}\theta_{v_f}\Delta_f + L_G^{w_f}\theta_{w_f}\Delta_f \\
&+ \sum_{l\in\mathbb{I}} L_{F_f}^{y_f}L_{h_f}L_{\hat{g}}^{x_s}\mathcal{M}_{sl}\Delta_f^2/2
\end{aligned} \tag{2.29}$$

Given the definition of function $\mu(s)$ with $s = \Delta_f$, (2.29) can be re-written as (2.27) in Proposition 3. This completes the proof. $\square$

Theorem 2. *Consider the reduced-order fast system in (2.5), if the proposed f-MHE is implemented following the MHE design in (2.17) based on a local deterministic nonlinear observer satisfying (2.12b) and conditions in Proposition 3, and if the s-MHEs are designed satisfying conditions in Theorem 1, and if there exist positive constants ϱ_f and ρ_f that satisfy $0 \leq \varrho_f \leq \rho_f$, and there exists a concave function $\mathcal{Q}_f$ such that*

$$\mathcal{Q}_f(|\varepsilon_f|) \geq \beta_f(|\varepsilon_f|, \Delta_f) \tag{2.30}$$

for all $|\varepsilon_{z_f}| \leq \rho_f$, and if there exist constants $a_f \geq 1$, $b_f \geq 0$ and $\xi_f \geq 0$ such that

$$\varrho_f - a_f\Big(\mathcal{Q}_f(\varrho_f) + \sum_{l\in\mathbb{I}} L_{F_f}^{x_s}\rho_{sl}\Delta_f + \mu(\Delta_f)\Big) - b_f(\sum_{l\in\mathbb{I}} L_{h_f}L_{\hat{g}}^{x_s}\rho_{sl} + \theta_{v_f}) \geq \xi_f \tag{2.31}$$

and if κ_f is selected subject to the following condition: $0 \leq \kappa_f \leq \min\{(a_f - 1)/L_{h_f}, b_f\}$, then the error $|\theta_f|$ between the estimated state x_f based on f-MHE and the actual fast system state x_f is ultimately bounded as

$$\lim_{\tau\to\infty} \sup|\theta_f(\tau)| \leq \rho_f^{min} + \sum_{l\in\mathbb{I}} L_{\hat{g}}^{x_s}\rho_{sl}^{min} \tag{2.32}$$

with $\rho_f^{min} = \max\{|\varepsilon_f(\tau + \Delta_f)|\colon |\varepsilon_f(\tau)| \leq \varrho_f\}$ for $|\varepsilon_f(0)| \leq \rho_f$.

Proof: It is straightforward from (2.17e) that:

$$|\hat{e}_f(\tau_{q+1}) - z_f(\tau_{q+1})| \leq \kappa_f|y_f(\tau_{q+1}) - h_f(z_f(\tau_{q+1}) + \hat{g}(\hat{x}_s(\tau_{q+1}), w_f(\tau_{q+1})))|$$

From the Lipschitz property of h_f and $\hat{g}$ and triangle inequality $|\hat{e}_f - e_f| \leq |\hat{e}_f - z_f| + |e_f - z_f|$, we have

$$\begin{aligned} |\hat{e}_f(\tau_{q+1}) - e_f(\tau_{q+1})| \leq (1 + L_{h_f}\kappa_f)|e_f(\tau_{q+1}) - z_f(\tau_{q+1})| \\ + \sum_{l\in\mathbb{I}} \kappa_f L_{h_f} L_{\hat{g}}^{x_s} |e_{sl}(\tau_{q+1})| + \kappa_f \theta_{v_f} \end{aligned} \tag{2.33}$$

From Proposition 3 and (2.33), and the fact that $e_{z_f}(\tau_q) = e_f(\tau_q)$, if (2.30) holds then we can derive that

$$\begin{aligned} |\varepsilon_f(\tau_{q+1})| <& (1 + L_{h_f}\kappa_f)\Big(\mathcal{Q}_f(|\varepsilon_f(\tau_q)|) + \sum_{l\in\mathbb{I}} L_{F_f}^{x_s} |e_{sl}(\tau_{q-n+1})|\Delta_f + \mu(\Delta_f)\Big) \\ &+ \sum_{l\in\mathbb{I}} \kappa_f L_{h_f} L_{\hat{g}}^{x_s} |e_{sl}(\tau_{q+1})| + \kappa_f \theta_{v_f} \end{aligned} \tag{2.34}$$

We know $|e_{si}(0)| \leq \rho_{si}$, $i \in \mathbb{I}$ and $|e_{si}|$ is ultimately bounded as described in Theorem 1. Therefore, we have $|e_{si}| \leq \rho_{si}$. If there exists a constant ϱ_f satisfying (2.31) and κ_f is selected as stated in Theorem 2, then for any estimation error ε_f satisfying $\varrho_f \leq |\varepsilon_f| \leq \rho_f$ we can obtain:

$$\begin{aligned} \varepsilon_f - (1 + L_{h_f}\kappa_f)\Big(\mathcal{Q}_f(|\varepsilon_f|) + \sum_{l\in\mathbb{I}} L_{F_f}^{x_s} |e_{sl}(\tau_{q-n+1})|\Delta_f \\ + \mu(\Delta_f)\Big) - \sum_{l\in\mathbb{I}} \kappa_f L_{h_f} L_{\hat{g}}^{x_s} |e_{sl}(\tau_{q+1})| - \kappa_f \theta_{v_f} \geq \xi_f \end{aligned} \tag{2.35}$$

Based on (2.34) and (2.35), we obtain that $|\varepsilon_f(\tau_{q+1})| < |\varepsilon_f(\tau_q)| - \xi_f$ if $\varrho_f \leq |\varepsilon_f| \leq \rho_f$. This implies that the error given by f-MHE for the fast system is decreasing if $|\varepsilon_f| \geq \varrho_f$. Moreover, the estimation error $|\varepsilon_f|$ will become smaller than ϱ_f within finite steps, and once $|\varepsilon_f| < \varrho_f$ is satisfied, then it is no longer guaranteed that the estimation error $|\varepsilon_f|$ will further decrease. However, it will remain to satisfy $|\varepsilon_f| \leq \rho_f^{min}$ with the definition $\rho_f^{min} = \max\{|\varepsilon_f(\tau + \Delta_f)|: |\varepsilon_f(\tau)| \leq \varrho_f\}$, i.e., $\lim_{t\to\infty} \sup|\varepsilon_f| \leq \rho_f^{min}$.

Finally, we consider the estimation error θ_f between state $\hat{x}_f$ obtained by the DMHE and actual state x_f, which can be expressed as $|\theta_f| = |\hat{x}_f - x_f| = |\hat{e}_f + \hat{g}(\hat{x}_s, w_f) - e_f - \hat{g}(x_s, w_f)| \leq |\varepsilon_f| + \sum_{l\in\mathbb{I}} L_{\hat{g}}^{x_s} |e_{sl}|$. Since $|\varepsilon_f|$ and $|e_{sl}|$, $l \in \mathbb{I}$, are proved to decrease and will ultimately remain to satisfy $|\varepsilon_f| \leq \rho_f^{min}$ and $|e_{sl}| \leq \rho_{sl}^{min}$, respectively, the estimation error $|\theta_f|$ is ultimately

bounded as $\lim_{t\to\infty} \sup|\theta_f| \leq \rho_f^{min} + \sum_{l\in\mathbb{I}} L_{\hat{g}}^{x_s} \rho_{sl}^{min}$. $\square$

In condition (2.31) of Theorem 2, the function $\mathcal{Q}_f(\cdot)$ is a concave approximation of $\beta_f(\cdot, \Delta_f)$. The term $\mu(\Delta_f)$ characterizes the effects of the fast sampling time Δ_f. The term $\sum_{l\in\mathbb{I}} L_{F_f}^{x_s} \rho_{sl} \Delta_f$ characterizes the coupling between the slow dynamics and the fast dynamics. The term $b_f(\sum_{l\in\mathbb{I}} L_{h_f} L_{\hat{g}}^{x_s} \rho_{sl} + \theta_{v_f})$ characterizes the effect of measurement noise and the impact of the s-MHEs on the f-MHE. Note that conditions (2.30) and (2.31) can be satisfied simultaneously as long as the convergence rate of the auxiliary observer for the fast dynamics is sufficiently large, the sampling time Δ_f and the measurement noise upper bound θ_{v_f} are sufficiently small. Note that while the performance of the s-MHEs does not depend on the f-MHE, the performance of the f-MHE depends on the s-MHEs as can be seen from (2.32).

Remark 4. *In Proposition 3, the local Lipschitz assumptions on the functions ensure that (2.29) can be retrieved via integration. Proposition 3 proves that within a fast sampling time Δ_f, the estimation error ε_{z_f} induced by the nonlinear observer for the fast system is bounded as described in (2.27), which may also be interpreted using input-to-state stability with the sampling time Δ_f and the error in the slow states $e_{sl}(\tau_{q-n+1})$ as inputs and the estimation error given by the observer ε_{z_f} as the state.*

Remark 5. *The separation of the two time scales in the system dynamics is indicated by the small positive parameter ϵ, which may be understood as the ratio of the time constants of the fast dynamics and slow dynamics. The proposed design relies on an obvious separation of the two time scales (i.e., ϵ should be sufficiently small). Note that the conditions in Theorems 1 and 2 implicitly impose a restriction on the maximum value of ϵ.*

2.5 Application to a chemical process

A chemical process consisting of two continuous stirred tank reactors (CSTRs) and a flash separator connected in series is considered. Pure material A is fed into the two CSTRs, in which two reactions occur, i.e. $A \to B$ and $B \to C$. The outlet of the second CSTR is fed

into the flash separator at a flow rate F_2. The overhead of the separator is condensed and passed to a downstream unit at flow rate F_r and the bottom product stream is removed at flow rate F_{outlet}. Each tank is equipped with a jacket to heat or cool the tank. The dynamic model obtained via mass and energy balances is as follows:

$$\frac{dx_f^A}{dt} = \frac{F_{10}}{V_1}(x_{A10} - x_f^A) - k_1 e^{\frac{-E_1}{RT_f}} x_f^A \tag{2.36a}$$

$$\frac{dx_f^B}{dt} = \frac{F_{10}}{V_1}(x_{B10} - x_f^B) + k_1 e^{\frac{-E_1}{RT_f}} x_f^A - k_2 e^{\frac{-E_2}{RT_f}} x_f^B \tag{2.36b}$$

$$\frac{dT_f}{dt} = \frac{F_{10}}{V_1}(T_{10} - T_f) + \frac{-\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_f}} x_f^A + \frac{-\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_f}} x_f^B + \frac{Q_1}{\rho c_p V_1} \tag{2.36c}$$

$$\frac{dx_{s1}^A}{dt} = \frac{F_1}{V_2}(x_f^A - x_{s1}^A) + \frac{F_{20}}{V_2}(x_{A20} - x_{s1}^A) - k_1 e^{\frac{-E_1}{RT_{s1}}} x_{s1}^A \tag{2.36d}$$

$$\frac{dx_{s1}^B}{dt} = \frac{F_1}{V_2}(x_f^B - x_{s1}^B) + \frac{F_{20}}{V_2}(x_{B20} - x_{s1}^B) + k_1 e^{\frac{-E_1}{RT_{s1}}} x_{s1}^A - k_2 e^{\frac{-E_2}{RT_{s1}}} x_{s1}^B \tag{2.36e}$$

$$\begin{aligned}\frac{dT_{s1}}{dt} =& \frac{F_1}{V_2}(T_f - T_{s1}) + \frac{F_{20}}{V_2}(T_{20} - T_{s1}) + \frac{-\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_{s1}}} x_{s1}^A \\ &+ \frac{-\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_{s1}}} x_{s1}^B + \frac{Q_2}{\rho c_p V_2}\end{aligned} \tag{2.36f}$$

$$\frac{dx_{s2}^A}{dt} = \frac{F_2}{V_3}(x_{s1}^A - x_{s2}^A) - \frac{(F_r + F_p)}{V_3}(x_{Ar} - x_{s2}^A) \tag{2.36g}$$

$$\frac{dx_{s2}^B}{dt} = \frac{F_2}{V_3}(x_{s1}^B - x_{s2}^B) - \frac{(F_r + F_p)}{V_3}(x_{Br} - x_{s2}^B) \tag{2.36h}$$

$$\frac{dT_{s2}}{dt} = \frac{F_2}{V_3}(T_{s1} - T_{s2}) + \frac{Q_3}{\rho c_p V_3} + \frac{(F_r + F_p)}{\rho c_p V_3}(x_{Ar}\Delta H_{vapA} + x_{Br}\Delta H_{vapB} + x_{Cr}\Delta H_{vapC}) \tag{2.36i}$$

where $x_f^A,\ x_{s1}^A,\ x_{s2}^A$ denote the mass fractions of A in tank 1, 2, 3; $x_f^B,\ x_{s1}^B,\ x_{s2}^B$ are the mass fractions of B in tank 1, 2, 3; $x_f^C,\ x_{s1}^C,\ x_{s2}^C$ are the mass fractions of C in tank 1, 2, 3; $x_{Ar},\ x_{Br},\ x_{Cr}$, respectively, represent the mass fractions of $A,\ B,\ C$ in the overhead stream; $T_f,\ T_{s1},\ T_{s2}$ are the temperatures in tank 1, 2, 3; $T_{10},\ T_{20}$ denote the feed stream temperatures; $F_1,\ F_2$ are the effluent flow rates from reactors; $F_{10},\ F_{20}$ denote steady feed stream flow rates; $F_r,\ F_p$ are, respectively, the overhead and bottom stream flow rates in the sep-

arator; $V_1,\ V_2,\ V_3$ denote the volumes of tank 1, 2, 3; $E_1,\ E_2$ denote the activation energy for reactions; $k_1,\ k_2$ are the pre-exponential values for reactions; $\Delta H_1,\ \Delta H_2$ denote the heats of reaction for reactions; ΔH_{vapi} denotes the evaporating enthalpies for material i, $i = A,\ B,\ C$; $Q_1,\ Q_2,\ Q_3$ denote the heat inputs/removals; and $c_p,\ R,\ \rho$, respectively, denote the heat capacity, gas constant and solution density. The parameters of the process variables are given in Table 2.1. We assume that the amount of reaction occurring in the separator is negligible. We note that the mass fractions of the overhead stream $x_{Ar},\ x_{Br},\ x_{Cr}$ are only dependent on the material composition in the flash separator.

The entire system is divided into three subsystems: the first CSTR, of which the system states evolve at a relatively fast time-scale, is the fast system; the second CSTR and the separator constitute the slow system, which is further divided into two slow subsystems. Specifically, $x_f = [x_f^A,\ x_f^B,\ T_f]^T$ is the state vector for the fast system, while $x_{s1} = [x_{s1}^A,\ x_{s1}^B,\ T_{s1}]^T$ and $x_{s2} = [x_{s2}^A,\ x_{s2}^B,\ T_{s2}]^T$ denote the state vectors of the two slow subsystems, respectively. It is assumed that the temperatures $T_f,\ T_{s1}$ and T_{s2} are measurable, and are considered as the system outputs (i.e. $y_f = T_f,\ y_{s1} = T_{s1}$ and $y_{s2} = T_{s2}$). The objective is to estimate the entire system state based on the three outputs using the proposed two-time-scale DMHE. In the simulations, constant heat inputs to the three tanks are considered; that is, $Q = [Q_1,\ Q_2,\ Q_3]^T = [3.0 \times 10^6\ KJ/h,\ 1.0 \times 10^6\ KJ/h,\ 3.0 \times 10^6\ KJ/h]^T$.

The local auxiliary observers assumed in (2.9) and (2.11) are designed following [30]. Random noise in the process and measurements is generated following normal distribution. In local MHE designs, the estimation horizons for f-MHE and s-MHEs are $N_f = N_s = 5$. The weighting matrices in the cost function for f-MHE and s-MHEs are $Q_f = Q_{s1} = Q_{s2} = diag([1,\ 1.44,\ 3600])$, $R_f = R_{s1} = R_{s2} = 8.314$. The parameters κ_f in constraint (2.17e) and $\kappa_{s1},\ \kappa_{s2}$ in constraint (2.15e) are tuned following [24] to be $\kappa_f = [0.0041\ 0.0141\ 0.3580]^T$, $\kappa_{s1} = [0.0039\ 0.0140\ 0.3000]^T$ and $\kappa_{s2} = [0.0025\ 0.0180\ 0.4440]^T$, respectively.

We consider three different schemes: (I) the proposed DMHE; (II) distributed nonlinear observers in which the augmented nonlinear observers are used and connected in a dis-

Table 2.1: Process parameters

$F_{10} = 12.0\ \frac{m^3}{h}$	$\Delta H_1 = -6.0 \cdot 10^4\ \frac{KJ}{kmol}$
$F_{20} = 3.0\ \frac{m^3}{h}$	$\Delta H_2 = -7.0 \cdot 10^4\ \frac{KJ}{kmol}$
$F_r = 13.4\ \frac{m^3}{h}$	$\Delta H_{vapA} = -3.53 \cdot 10^4\ \frac{KJ}{kmol}$
$F_p = 0.5\ \frac{m^3}{h}$	$\Delta H_{vapB} = -1.57 \cdot 10^4\ \frac{KJ}{kmol}$
$V_1 = 1.0\ m^3$	$\Delta H_{vapC} = -4.07 \cdot 10^4\ \frac{KJ}{kmol}$
$V_2 = 5.0\ m^3$	$k_1 = 2.77 \times 10^3\ s^{-1}$
$V_3 = 4.0\ m^3$	$k_2 = 2.6 \times 10^3\ s^{-1}$
$\alpha_A = 3.5$	$c_p = 4.2\ KJ/kg \cdot K$
$\alpha_B = 1.0$	$R = 8.314\ KJ/kmol \cdot K$
$\alpha_C = 0.5$	$\rho = 1000.0\ kg/m^3$
$T_{10} = 300\ K$	$x_{A10} = 1.0$
$T_{20} = 300\ K$	$x_{B10} = 0$
$E_1 = 5.0 \cdot 10^4\ \frac{KJ}{kmol}$	$x_{A20} = 1.0$
$E_2 = 6.0 \cdot 10^4\ \frac{KJ}{kmol}$	$x_{B20} = 0$

tributed fashion the same as Scheme I; and (III) a DMHE scheme in which local estimators are designed based on regular MHE. In the two DMHE schemes, the arrival cost in each subsystem MHE is included and is approximated using an extended Kalman filtering approach neglecting the interactions between subsystems. First, a set of simulations is carried out to verify the effectiveness of the proposed DMHE (Scheme I) and to illustrate the advantage of Scheme I over the auxiliary nonlinear observers (Scheme II). The f-MHE is evaluated at a fast sampling time $\Delta_f = 18s$ while all the s-MHEs are evaluated at $\Delta_s = 36s$. The results are presented in Fig. 2.2. We see that the proposed DMHE is able to track the actual state trajectories of the chemical process. The corresponding mean and maximum values of the Euclidean norm of the normalized estimation error of the proposed DMHE are 0.6642 and 2.1753, respectively. We note that the ability to use different sampling periods in the f-MHE and s-MHEs contributes to the applicability (especially the computational efficiency) of the proposed DMHE. We also see that the auxiliary nonlinear observers have rather poor

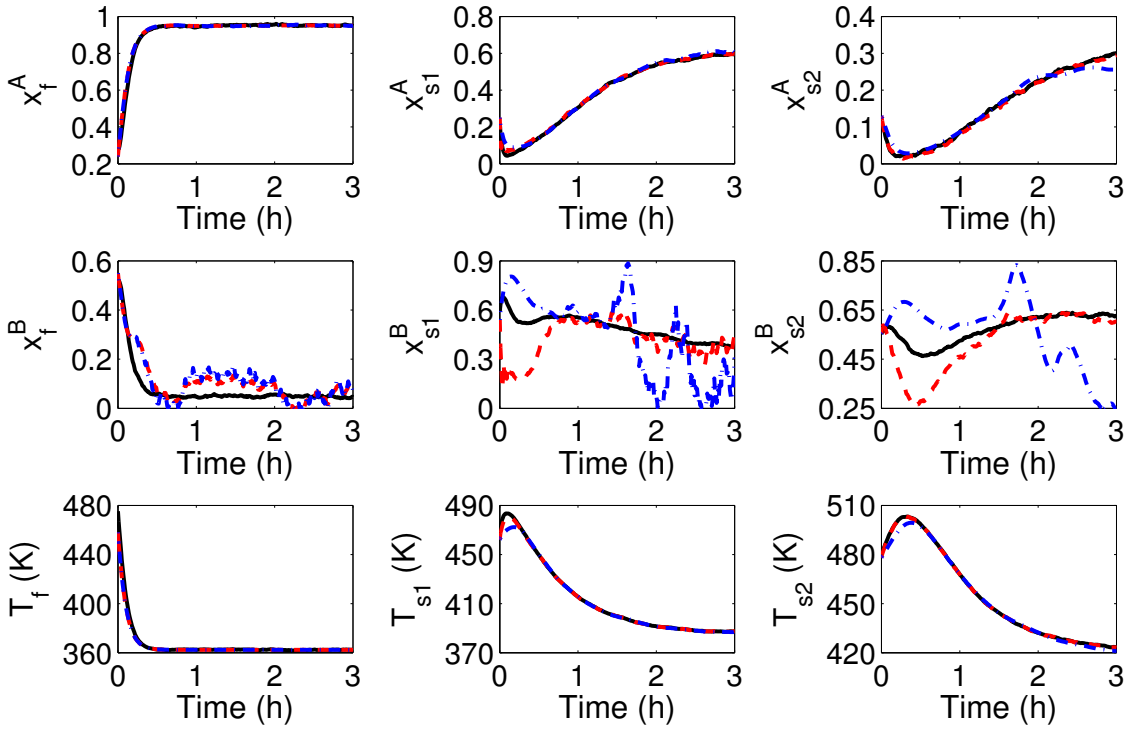

Figure 2.2: Trajectories of the actual process states (solid lines), state estimates calculated by the proposed DMHE (dashed lines) and state estimates calculated by Scheme II (dash-dotted lines) with $\Delta_f = 18s$ and $\Delta_s = 36s$.

performance in estimating the concentrations of B. In Scheme II, the mean and maximum values of the normalized estimation error norms are 0.9146 and 2.0720, respectively. The trajectories of the normalized estimation error norms for Scheme I and Scheme II are given in Fig. 2.3. It is seen that the mean error norm of Scheme II is much larger than that of Scheme I. The maximum error norms for the two schemes are similar. Note that the maximum error of Scheme I is primarily due to the initial stage. The results imply that Scheme I is able to give overall more accurate estimates compared to Scheme II. This is because the proposed DMHE takes into account the noise information in an optimization framework of MHE.

We also conduct simulations to demonstrate the less dependence of the proposed method (Scheme I) on the size of the estimation horizon compared to the counterpart based on regular MHE (Scheme III). The mean values of the normalized error norms of the two schemes with different estimation horizon sizes are presented in Fig. 2.4. We see that the proposed scheme gives good performance even with a very small horizon and its performance is much less

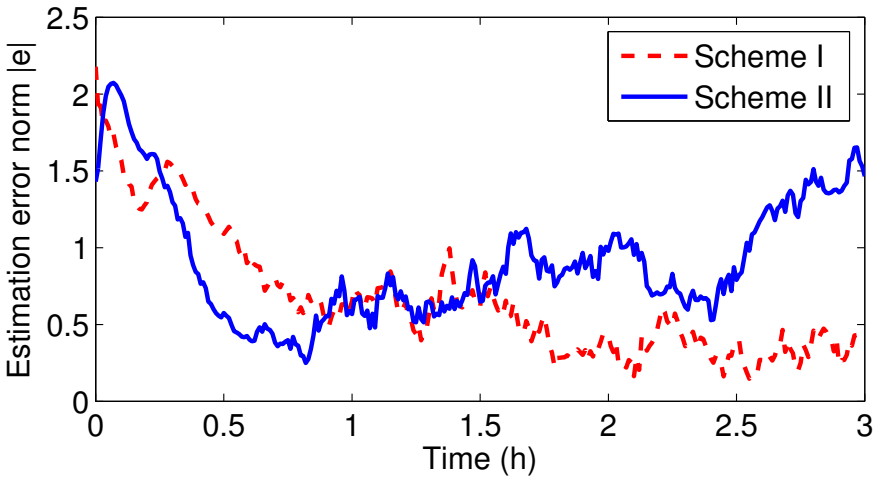

Figure 2.3: Trajectories of the normalized estimation errors of Scheme I (dashed line) and Scheme II (solid line).

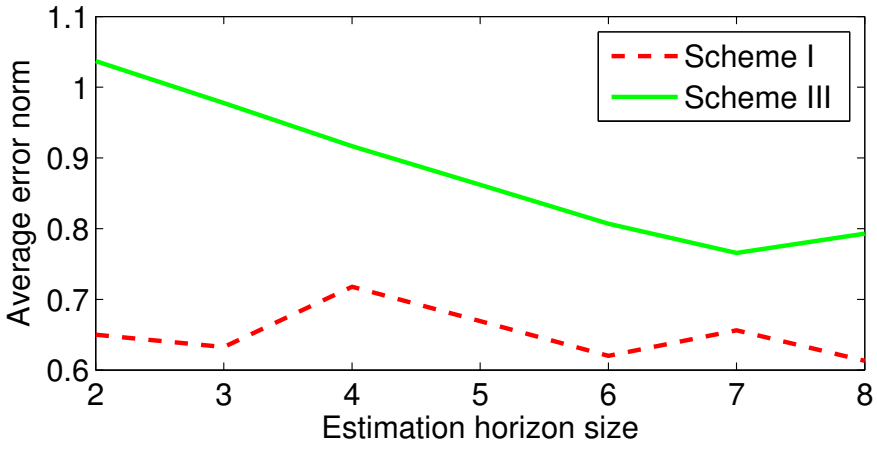

Figure 2.4: Average normalized estimation error norms of Scheme I (dashed line) and scheme III (solid line) with different estimation horizon sizes.

sensitive to the horizon size. Also, the proposed scheme gives smaller average estimation error than Scheme III. The results also imply that the proposed scheme could be much more computationally efficient since a much smaller horizon can be used.

Next, we study the effects when the same sampling time is used in both the f-MHE and the s-MHEs (i.e., $\Delta_f = \Delta_s = 18s$). The proposed DMHE can still track the actual system states. The mean and maximum values of the normalized error norms are 0.5416 and 2.0737, respectively. Compared with the previous simulations, the estimation performance in terms of the maximum and mean error values has been improved by about 18% and 4%,

respectively. However, the performance improvement is at the cost of much higher communicational and computational burdens. When the s-MHEs use $\Delta_s = 18s$, they are evaluated twice frequently compared with the case with $\Delta_s = 36s$. This implies that the number of communication times between the s-MHEs increases by 100%. The average simulation time over 10 runs for a period of three hours under the proposed DMHE with $\Delta_s = 36s$ is about $382s$ while the corresponding value when $\Delta_s = 18s$ is about $692s$ which is increased by 81%.

We also study the effects of the developed one-directional communication strategy. We carry out simulations in the context of bi-directional communication, i.e., the f-MHE and the s-MHEs send their state estimates to each other. Compared to the one-directional communication, the s-MHEs will have more accurate information of the fast system states (i.e., the estimate calculated by the f-MHE) and will use it (instead of steady-state values of the fast dynamics) in state estimation of the slow system states. In this set of simulations, the corresponding mean and maximum values of the error norm are 0.6683 and 2.1089. These values are very close to the ones of Scheme I with only one-directional communication from the s-MHEs to the f-MHE. This implies that extra information from f-MHE to the s-MHEs does not help to improve the estimation performance significantly, which further demonstrates the effectiveness of the proposed DMHE.

2.6 Summary

In this chapter, we developed a distributed state estimation method based on MHE for a class of two-time-scale nonlinear systems. The nonlinear system was decomposed into a fast system and several slow subsystems. In the proposed DMHE, a one-directional communication strategy was adopted. Stability analysis was carried out for the proposed distributed scheme. The effectiveness of the proposed approach was illustrated via the application to a chemical process.

Chapter 3

Forming distributed state estimation network from decentralized estimators

In this chapter, we assume that a decentralized state estimation system already exists for the nonlinear system, where the local estimators can be of different types. In order to achieve improved estimation performance, the existing decentralized estimators may be connected together via a communication network to form a distributed state estimation network. We propose a systematic approach to take advantage of the existing decentralized estimators potentially of different types to form a distributed state estimation network without performing a complete re-design of the estimation system. Specifically, a compensator is designed for each subsystem, and is connected to the corresponding decentralized estimator to obtain an augmented estimator. The augmented estimators for the subsystems communicate with each other to exchange subsystem state estimates and measurements via a communication network every sampling time. In Section 3.2.4, we derive sufficient conditions on the convergence and boundedness of the estimation error of the proposed distributed estimation network. In Section 3.3, the proposed approach is demonstrated via the application to two chemical process examples and one hybrid-tank plant. This chapter is a revised version of [47, 129].

3.1 Preliminaries

3.1.1 Notation

The operator $|\cdot|$ denotes the Euclidean norm of a vector. $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. A function $f(s)$ is said to be Lipschitz with respect to its argument s if there exists a positive constant L_f^s such that $|f(s_1) - f(s_2)| \leq L_f^s |s_1 - s_2|$ for s_1 and s_2. Further, a function $f(s)$ is locally Lipschitz if it is Lipschitz on every compact subset of $\mathbb{S}$ which itself is a compact set. A continuous function $\alpha : [0, a) \to [0, \infty)$ belongs to class $\mathcal{K}$ if it is strictly increasing and satisfies initial condition $\alpha(0) = 0$. A function $\beta(r, s)$ is said to belong to class $\mathcal{KL}$ if for any fixed s, $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to its first argument r, and for any fixed r, it is deceasing with respect to the second argument s, and $\beta(r, s) \to 0$ as $s \to \infty$. A^+ denotes a matrix (or a vector) being the Moore-Penrose pseudoinverse of matrix (or vector) A. $\mathbb{I}$ represents a set containing finite integers defined as $\mathbb{I} = \{1, \ldots, m\}$. We use $\{s(t_n)\}_{n=0}^k$ to denote a sequence of s from $n = 0$ to $n = k$ (i.e., $s(t_0),\ s(t_1), \ldots, s(t_k)$). The operator diag (v) denotes a diagonal matrix in which the diagonal elements are the elements in the vector v.

3.1.2 System description

In this study, we take into account a class of nonlinear systems comprised of m interconnected subsystems described in the following form:

$$\begin{aligned} \dot{x}(t) &= \mathcal{F}\left(x(t), w(t)\right) \\ y(t) &= h(x(t)) + v(t) \end{aligned} \tag{3.1}$$

where $x(t) \in \mathbb{R}^{n_x}$ represents the state vector, $w(t) \in \mathbb{R}^{n_w}$ represents the disturbances affecting the dynamics of the system states, $y \in \mathbb{R}^{n_y}$ denotes the measured output vector of the entire system, and $v(t) \in \mathbb{R}^{n_v}$ denotes the measurement noise of the entire system. The function $\mathcal{F}$ characterizes the dynamics of x. h denotes a nonlinear vector function, of which the

form is dependent on the relationship between the system states and the available output measurements.

It is assumed that the i-th subsystem of system (2.1), $i \in \mathbb{I}$, can be described as follows:

$$\begin{aligned} \dot{x}_i(t) &= f_i\left(x_i(t), w_i(t)\right) + \tilde{f}_i\left(x_i(t), X_i(t)\right) \\ y_i(t) &= h_i\left(x_i(t)\right) + v_i(t) \end{aligned} \tag{3.2}$$

where $x_i \in \mathbb{R}^{n_{x_i}}$ denotes the vector of state variables, $w_i \in \mathbb{R}^{n_{w_i}}$ represents the disturbances directly affecting the dynamics of the i-th subsystem, $y_i \in \mathbb{R}^{n_{y_i}}$ denotes the measured output of subsystem i, while $v_i \in \mathbb{R}^{n_{v_i}}$ denotes the measurement noise associated with the i-th subsystem. Vector function f_i characterizes the dependence of the dynamics of x_i on itself while the vector function $\tilde{f}_i$ characterizes the interaction between subsystem i and other subsystems. It is assumed that the functions $\tilde{f}_i$, $i \in \mathbb{I}$, are locally Lipschitz with respect to their arguments on the compact sets of interest. $\tilde{f}_i$ and h_i are assumed to belong to class C^2 (i.e., having 2 continuous derivatives) over a domain of interest. h_i, $i \in \mathbb{I}$, denotes a nonlinear vector function. $X_i(t) \in \mathbb{R}^{n_{X_i}}$ denotes a vector of the states of the subsystems other than subsystem i that are involved in characterizing the interaction. In the remainder, we will utilize $\mathbb{I}_i$, $i \in \mathbb{I}$, to denote the set of subsystem indices, of which the corresponding subsystem states are involved in X_i. For example, if X_1 involves states of subsystem 3 and subsystem 5, then $\mathbb{I}_1 = \{3,\ 5\}$. We call each subsystem l, $l \in \mathbb{I}_i$ an upstream interacting subsystem of subsystem i. Accordingly, we call subsystem i a downstream interacting subsystem of each subsystem l. We further introduce symbol $\mathbb{J}_i$ to denote a set that is defined as $\mathbb{J}_i := \mathbb{I}_i \cup \{i\}$. We assume that all the integers contained in the set $\mathbb{I}_i$, $i \in \mathbb{I}$, are known. The system states and the disturbances are assumed to be bounded as follow.

Assumption 1. *Subsystem states x_i, $i \in \mathbb{I}$, satisfy the following constraint:*

$$x_i \in \mathbb{X}_i \tag{3.3}$$

where $\mathbb{X}_i$, $i \in \mathbb{I}$, is a compact set. Moreover, the system disturbances and measurement noise are bounded, such that $w_i \in \mathbb{W}_i$ and $v_i \in \mathbb{V}_i$ are satisfied $\forall i \in \mathbb{I}$. The sets $\mathbb{W}_i$ and $\mathbb{V}_i$, $i \in \mathbb{I}$, are defined as

$$\begin{aligned} \mathbb{W}_i :=& \{w_i \in \mathbb{R}^{n_{w_i}} \ \ s.t. \ |w_i| \leq \theta_{w_i}, \theta_{w_i} > 0\}, \\ \mathbb{V}_i :=& \{v_i \in \mathbb{R}^{n_{v_i}} \ \ s.t. \ |v_i| \leq \theta_{v_i}, \theta_{v_i} > 0\} \end{aligned} \tag{3.4}$$

where θ_{w_i} and θ_{v_i} are known scalars.

Assumption 2. *$\frac{\partial h_i}{\partial x_i}|_{x_i = z_i(t_k)}$ is a full column (or full row) rank matrix for $t_{k \geq 0}$ where $z_i(t_k)$ denotes the state estimate of x_i, $i \in \mathbb{I}$, at time instant t_k given by the distributed state estimation scheme to be specified.*

Remark 6. *The assumption (Assumption 1) on the boundedness of each subsystem state within a compact set is motivated by the fact that most processes are maintained in bounded operating regions by control systems. This assumption does imply that the system considered is stable (through the regulation of a control system). The size of the compact set depends on the properties of the system (and the corresponding control system). Note that the boundedness of each subsystem state also implies that each vector X_i, $i \in \mathbb{I}$, is bounded within a compact set.*

Remark 7. *It is highly possible to achieve decomposed subsystems in the form of (3.2) providing that sufficient measurements are available, which is also a prerequisite for state estimation. More specifically, in the ordinary differential equations (ODEs) describing the dynamics of x_i, we can form f_i using the terms containing only local state x_i, and then arrange the remaining terms characterizing interacting dynamics in $\tilde{f}_i$, $i \in \mathbb{I}$.*

In the subsystem model (3.2), we consider that the disturbances (uncertainties) only exist in the local dynamics of each subsystem. Note that the design and the analysis will be very similar if we consider that the interacting term f_i in (3.2) is also affected by uncertainties.

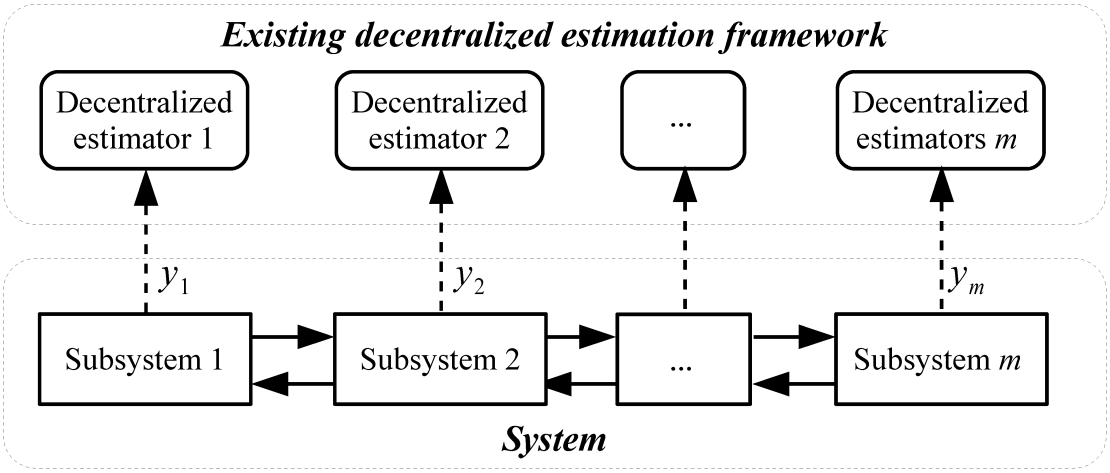

Figure 3.1: A schematic of the existing decentralized estimation system.

3.1.3 Problem formulation and assumptions

In this work, we provide an alternative approach to design distributed state estimation networks by taking fully advantage of existing decentralized estimation schemes. Specifically, we consider that for system (3.1), there exists a decentralized state estimation system, which consists of m decentralized local estimators for the m subsystems as shown in Figure 3.1. In this decentralized system, each local estimator is developed for a subsystem in a decentralized manner. The local estimators are evaluated independently without sharing information with each other. Constant values (e.g., the steady-state value of X_i or zero) are used in each decentralized estimator for handling subsystem interaction in a relatively conservative way. Note that the existing decentralized local estimators can be of different types. To achieve improved overall estimation performance, we aim at developing a systematic approach, which can connect the existing decentralized local estimators of different types together to form a distributed state estimation system instead of performing a complete re-design.

Specifically, we assume that the measurements of the subsystems are sampled synchronously at each sampling instant $\{t_{n\geq 0}\}$, where $t_n := t_0 + n\Delta$ with $t_0 = 0$ denoting the initial time, Δ being a constant sampling period and n representing positive integers. It is also assumed that for subsystem i, $i \in \mathbb{I}$, there exists a decentralized local estimator,

which takes advantage of sampled output measurements and is described in a discrete-time framework as follows:

$$z_i^o(t_k) = G_i(\ \{y_i(t_n)\}_{n=0}^{k}, z_i^o(t_0), X_{is}) \tag{3.5}$$

where z_i^o denotes the state estimate given by the i-th decentralized estimator, $\{y_i(t_n)\}_{n=0}^{k}$ is the output measurement sequence of subsystem i from the initial time t_0 to t_k, X_{is} denotes a vector containing constant values (e.g., the steady-state value of X_i or zero) that may be used to compensate for the interaction between subsystems, G_i is a vector field describing the discrete-time decentralized estimator in a general form. Estimator (3.5) is evaluated every sampling time starting from initial time t_0 to provide state estimates for the corresponding subsystem. The decentralized estimators are assumed to satisfy the following assumption in this work.

Assumption 3. *For subsystem i, $i \in \mathbb{I}$, in the form of (3.2), there already exists a decentralized estimator described as in (3.5), such that if the subsystem state trajectory of x_i are bounded as shown in (3.3), and if all the states of the subsystems that directly affect the dynamics of subsystem i are at the constant values used in the design of the decentralized estimators (i.e., $X_i(t) \equiv X_{is}$ for $i \in \mathbb{I}$ and $t \in [t_0, t_k]$), then there exists a class $\mathcal{KL}$ function β_i, class $\mathcal{K}$ functions γ_i and ϕ_i, such that the estimation error of decentralized estimator i in (3.5) is bounded as follows:*

$$\begin{aligned} |z_i^o(t_k) - x_i(t_k)| \leq & \beta_i(|z_i^o(t_0) - x_i(t_0)|, t_k - t_0) \\ & + \gamma_i\Big(\max_{\tau} |w_i(\tau)|\Big) + \phi_i\Big(\max_{\tau} |v_i(\tau)|\Big) \end{aligned} \tag{3.6}$$

The above assumption characterizes an upper bound on the estimation error for the i-th subsystem given by the decentralized local estimator in the presence of external disturbances and measurement noise without considering the dynamics of subsystem interaction. Assumption 3 is to describe the property of a commonly-used decentralized estimator design method. Specifically, Eq.(3.6) characterizes the convergence and boundedness of the decen-

tralized estimators. It does not remove the interaction between subsystems, nor simplifies the problem. It simply requires that the decentralized estimators are fairly well designed. In these designs, a decentralized estimator is designed for a local operating subsystem assuming that the interaction between the local operating subsystem and its interacting subsystems remains constant (e.g., at steady state or 0). For decentralized estimators, the interaction between subsystems is not considered or considered in conservative ways. Assumption 3 can be satisfied by some existing nonlinear state estimation methods, including moving horizon estimators [48, 49, 24] and extended Kalman filters [141]. Note that the above designs are based on certain mild assumptions. In particular, one major assumption of the MHE-based designs in [48, 49] is that the nonlinear system should be incrementally input/output-to-state stable (i-IOSS). While it is argued in [49] that it is in general difficult to verify whether a system is i-IOSS or not, this criterion is not a restrictive as it can be interpreted as a detectability criterion of nonlinear systems [50]. Note also that if certain local estimators were based on an EKF algorithm, an implicit requirement is that the the initial estimation error should be within a certain neighborhood of the origin.

3.1.4 Illustrative example

In this section, we provide a benchmark example to illustrate the ideas presented in the previous subsection. We consider a chemical process that is comprised of two well-mixed non-isothermal continuous stirred tank reactors (CSTRs) with different volumes that are connected in a cascade fashion. CSTR 2 is connected to CSTR 1 through the feed stream from the first CSTR (i.e., the outlet stream of CSTR 1 at flow rate F_1). Three paralleled irreversible elementary exothermic reactions take place in the two reactors, i.e., $A \to B$, $A \to C$ and $A \to D$. A denotes the reactant, B denotes the desired product while C and D represent two undesirable side products. A fresh feed flow with pure material A is fed into the first reactor at flow rate F_0, temperature T_0 and molar concentration C_{A0}. The effluent from the first reactor accounts for a feed stream to the second reactor, i.e., an inlet fed to CSTR 2

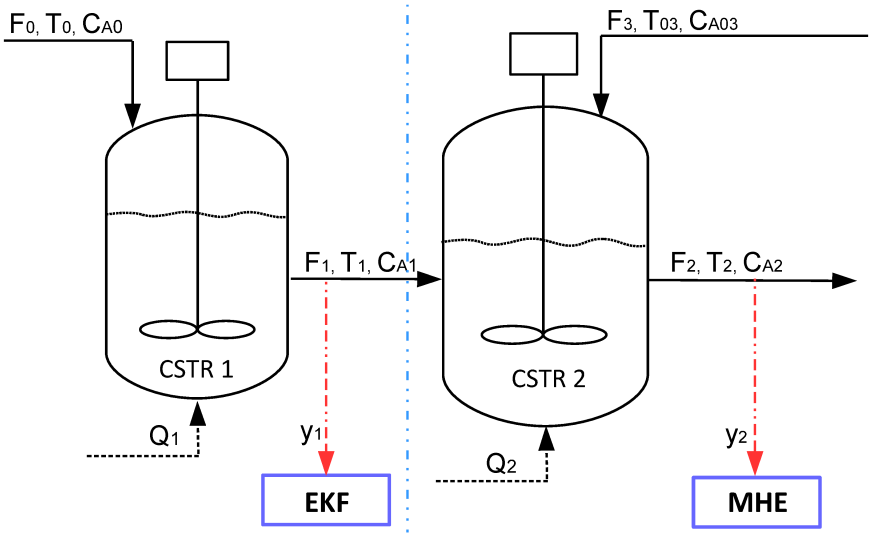

Figure 3.2: A schematic of the cascade chemical process and designed decentralized estimators.

at flow rate F_1, temperature T_1 and concentration C_{A1}. Another fresh feed stream containing pure material A is also fed at flow rate F_3, temperature T_{03} and molar concentration C_{A03} to CSTR 2 for reactions. The effluent of CSTR 2 is discharged for further processing. In this chemical process, each CSTR is equipped with a jacket to add/remove heat to/from the corresponding vessel. Based on typical modeling assumptions, mass balances and energy balances, we obtain a model that consists of four differential equations as follows to describe the dynamics of the process:

$$\begin{aligned}
\dot{T}_1 &= \frac{F_0}{V_1}(T_0 - T_1) + \sum_{i=1}^{3} \mathcal{M}_i(T_1) C_{A1} + \frac{Q_1}{\rho c_p V_1} \\
\dot{C}_{A1} &= \frac{F_0}{V_1}(C_{A0} - C_{A1}) - \sum_{i=1}^{3} \mathcal{N}_i(T_1) C_{A1} \\
\dot{T}_2 &= \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_3}{V_2}(T_{03} - T_2) + \sum_{i=1}^{3} \mathcal{M}_i(T_2) C_{A2} + \frac{Q_2}{\rho c_p V_2} \\
\dot{C}_{A2} &= \frac{F_1}{V_2}(C_{A1} - C_{A2}) + \frac{F_3}{V_2}(C_{A03} - C_{A2}) - \sum_{i=1}^{3} \mathcal{N}_i(T_2) C_{A2}
\end{aligned} \tag{3.7}$$

with $\mathcal{N}_i(T_j) = k_i e^{\frac{-E_i}{RT_j}}$ and $\mathcal{M}_i(T_j) = -\frac{\Delta H_i}{\rho c_p}\mathcal{N}_i(T_j)$, $j = 1,\ 2$. This model is established based on the chemical process introduced in [51]. The definitions of the variables in system model (3.7) are presented in Table 3.1. It is assumed that the temperatures T_1 and T_2 of the two CSTRs are measurable, while the molar concentrations C_{A1} and C_{A2} of material A in the two vessels are unmeasurable and need to be estimated. It is also assumed that there is an unmeasured but bounded disturbance in the flow rate F_3.

This system is decomposed into two subsystems. For subsystem i, $x_i = [x_{i,1}\ x_{i,2}]^{\mathrm{T}} = [T_i\ C_{Ai}]^{\mathrm{T}}$, is the state vector, $y_i = T_i + v_i$ is the output measurement, $w_i = [w_{i,1}\ w_{i,2}]^{\mathrm{T}}$ is the vector of additive disturbances, $i = 1,\ 2$. For the first subsystem, the vector field f_1 is with the following form:

$$
\begin{aligned}
f_1 &= \begin{bmatrix} f_{1,1}(x_1, w_1) \\ f_{1,2}(x_1, w_1) \end{bmatrix} \\
&= \begin{bmatrix} \frac{F_0}{V_1}(T_0 - x_{1,1}) + \sum_{i=1}^{3} \mathcal{M}_i(x_{1,1})x_{1,2} + \frac{Q_1}{\rho c_p V_1} \\ \frac{F_0}{V_1}(C_{A0} - x_{1,2}) - \sum_{i=1}^{3} \mathcal{N}_i(x_{1,1})x_{1,2} \end{bmatrix} + \begin{bmatrix} w_{1,1} \\ w_{1,2} \end{bmatrix}
\end{aligned}
$$

and $h_1(x_1) = [1\ \ 0]\, x_1$. Since the dynamics of the first subsystem is not affected by subsystem 2, X_1 is a zero vector and $\tilde{f}_1$ is a zero vector field. In terms of the second subsystem, we

Table 3.1: Process variables

$C_{A1},\ C_{A2}$	molar concentrations of A in reactors 1, 2
$C_{A0},\ C_{A3}$	molar concentrations of pure A in $F_0,\ F_3$
$T_1,\ T_2$	temperatures in reactors 1, 2
$T_0,\ T_{03}$	feed stream temperatures to reactors 1, 2
$F_1,\ F_2$	effluent flow rates from reactors 1 and 2
$F_0,\ F_3$	steady-state feed stream flow rates to reactors 1 and 2
$V_1,\ V_2$	volumes of reactor 1, 2
$E_1,\ E_2,\ E_3$	activation energies of the three reactions
$k_1,\ k_2,\ k_3$	pre-exponential values for the three reactions
$\Delta H_1,\ \Delta H_2,\ \Delta H_3$	enthalpies of the three reactions
$Q_1,\ Q_2$	heat inputs/removals into/from reactor 1, 2
$c_p,\ \rho$	heat capacity and solution density of fluid
R	gas constant

have:

$$
\begin{aligned}
f_2 &= \begin{bmatrix} f_{2,1}(x_2, w_2) \\ f_{2,2}(x_2, w_2) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{F_1}{V_2} T_2 + \frac{F_3}{V_2}(T_{03} - x_{2,1}) + \sum_{i=1}^{3} \mathcal{M}_i(x_{2,1}) x_{2,2} + \frac{Q_2}{\rho c_p V_2} \\ -\frac{F_1}{V_2} x_{2,2} + \frac{F_3}{V_2}(C_{A03} - x_{2,2}) - \sum_{i=1}^{3} \mathcal{N}_i(T_2) x_{2,2} \end{bmatrix} \\
&\quad + \begin{bmatrix} w_{2,1} \\ w_{2,2} \end{bmatrix} \\
\tilde{f}_2 &= \begin{bmatrix} \tilde{f}_{2,1}(x_2, X_2) \\ \tilde{f}_{2,2}(x_2, X_2) \end{bmatrix} = \begin{bmatrix} \frac{F_1}{V_2} x_{1,1} \\ \frac{F_1}{V_2} x_{1,2} \end{bmatrix}
\end{aligned}
$$

and $h_2(x_2) = [1 \ \ 0]\, x_2$.

Let us assume that decentralized state estimators have been developed for this process such that a local extended Kalman filter (EKF) is designed for CSTR 1 for reduced computational complexity and a moving horizon estimator (MHE) is designed for CSTR 2 in

order to take into account the boundedness information of the disturbance in F_3. As a typical practice, the steady-state value of the interaction between the two CSTRs is used in the decentralized estimators to compensate for the interaction. When operating close to the steady-state of the process, this composite EKF-MHE estimation system works well and provides a balance between estimation accuracy (by taking into account boundedness information of F_3) and computational cost (by using an EKF for CSTR 1). A schematic of the process together with the estimators for the reactors are presented in Figure 3.2.

Now, let us suppose that due to some reasons (e.g., increased market competition), we would like to enhance the estimation performance by enabling information exchange between the two subsystems (i.e., to adopt a distributed estimation system) such that the interaction between the two CSTRs can be compensated for in a less conservative fashion. However, it is not possible to take advantage of the existing estimators of two different types based on existing approaches. Therefore, we aim at developing a systematic approach to connect existing decentralized estimators which may be of different types to form a distributed scheme to achieve improved performance.

3.2 Forming distributed state estimation network

In this section, we present the proposed approach to connect existing decentralized local estimators in the form of (3.5) together to form a distributed state estimation network to achieve improved estimation performance.

3.2.1 Proposed distributed state estimation network

A schematic of the proposed distributed state estimation network is presented in Figure 3.3. In this design, each subsystem is equipped with an augmented estimator (AE) that consists of an existing decentralized estimator and a compensator, the design of which will be specified in the next subsection. Each augmented estimator is evaluated every sampling time.

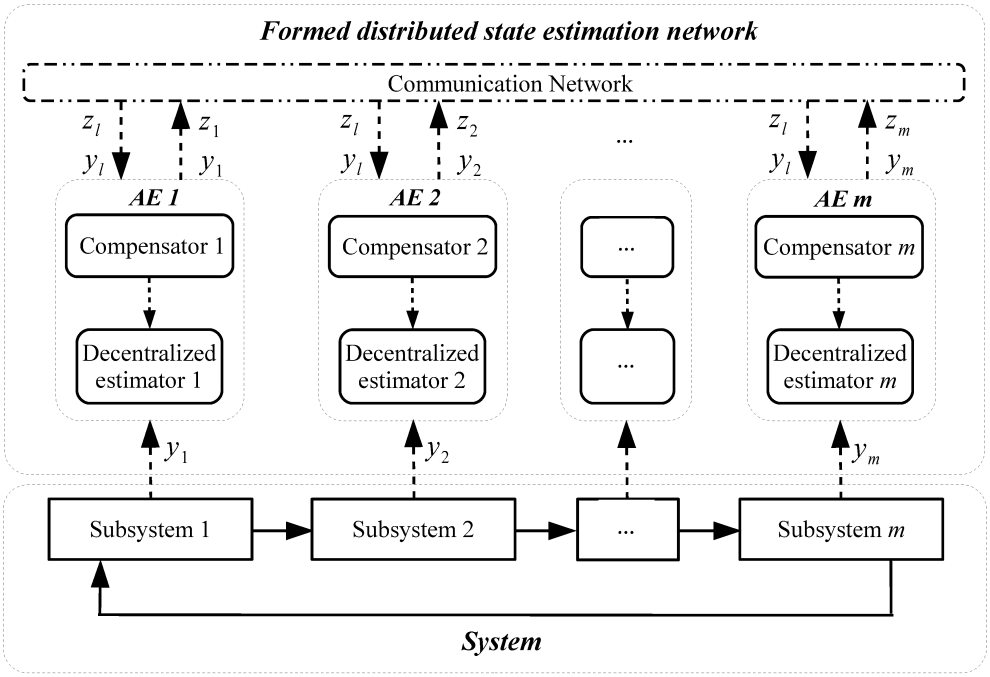

Figure 3.3: A schematic of the proposed distributed state estimation network.

Each augmented estimator receives its subsystem's output measurement every sampling instant, and communicates with the interconnected subsystems every sampling time to exchange subsystem state estimates and subsystem measurements. Specifically, the i-th augmented estimator, $\forall i \in \mathbb{I}$, receives the state estimates and the measurements of subsystem j, $j \in \mathbb{I}_i$, via the communication network. In the meantime, y_i, $i \in \mathbb{I}$, is sampled and is sent out together with its subsystem estimate z_i, $i \in \mathbb{I}$, to the augmented estimators that are corresponding to the downstream interacting subsystems of subsystem i (i.e., all the subsystems of which the dynamics are directly affected by x_i).

3.2.2 Design of subsystem compensators

To explicitly take into account the dynamics of the subsystem interaction, we design a compensator for each subsystem. The compensator for the i-th subsystem is designed as in

the following form:

$$\begin{aligned} z_i^c(t_{k+1}) =& \tilde{f}_i(z_i(t_k), Z_i(t_k))\Delta - \tilde{f}_i(z_i(t_k), X_{is})\Delta \\ &+ \sum_{l\in\mathbb{I}_i} K_{i,l}(z_l)(y_l(t_k) - h_l(z_l(t_k))\Delta \end{aligned} \tag{3.8}$$

where z_i^c is a compensation signal generated by the designed compensator, $z_i(t_k)$ represents the state estimate of subsystem i obtained at time instant t_k, $Z_i(t_k)$ denotes a vector which is an estimate of the state vector X_i at time instant t_k, while $K_{i,l}$ with $i \in \mathbb{I}$, $l \in \mathbb{I}_i$, are correction gains. Compensator i is designed based on the interaction model $\tilde{f}_i$ as well as the output measurements and subsystem estimates of subsystem i and its upstream interacting subsystems. It tracks the dynamics which is the difference between the complete subsystem dynamics and the local dynamics of the same subsystem.

Specifically, in compensator (3.8), the first two terms on the right-hand-side of (3.8) are used to approximate the deviation of the subsystem interaction of subsystem i, $i \in \mathbb{I}$, from the associated constant vector X_{is}. Since we can only use subsystem state estimates in the compensators instead of the actual values, the deviation between the actual subsystem interaction and the constant vector X_{is} used in the i-th decentralized estimator cannot be completely eliminated. In other words, an error caused by the use of subsystem estimates of the interacted subsystems (instead of the corresponding actual subsystem states) still exists. To handle this, we use the third term on the right-hand-side of Eq.(3.8) to further compensate for the error caused by the use of the estimated state (instead of the actual state) in the interaction model. The gain $K_{i,l}$ depends on z_l and is re-evaluated every sampling time as follows:

$$K_{i,l}(t_k) = \frac{\partial \tilde{f}_i}{\partial x_l}\left(\frac{\partial h_l}{\partial x_l}\right)^+\Bigg|_{x_l=z_l(t_k)} \tag{3.9}$$

for $i \in \mathbb{I}$ and $l \in \mathbb{J}_i$.

Note that the pseudo-inverse may sometimes lead to ill-conditioned matrices, which de-

pends on the available measurements and the output measurement function h_i, $i \in \mathbb{I}$. If ill-conditioning is not avoidable, one approach to overcome it in the proposed distributed estimation design is to use a sufficiently small sampling period Δ. The convergence of the estimates is still ensured providing that the condition in Theorem 3 which will be introduced later is satisfied.

3.2.3 Design of augmented estimators

We take advantage of the existing decentralized local estimators and the developed compensators to construct augmented estimators for the subsystems. Specifically, for subsystem i, $i \in \mathbb{I}$, the existing decentralized estimator in (3.5) is connected to the designed compensator i in (3.8). An augmented nonlinear estimator for subsystem i, $i \in \mathbb{I}$, is formulated and described as follows:

$$\begin{aligned} z_i(t_{k+1}) = {} & G_i(\{y_i(t_n)\}_{n=0}^{k+1}, z_i(t_0), X_{is}) + \tilde{f}_i(z_i(t_k), Z_i(t_k))\Delta \\ & - \tilde{f}_i(z_i(t_k), X_{is})\Delta + \sum_{l \in \mathbb{I}_i} K_{i,l}(z_l(t_k))(y_l(t_k) - h_l(z_l(t_k))\Delta \end{aligned} \tag{3.10}$$

where z_i denotes the state estimate given by the i-th augmented estimator, and $z_i(t_0)$ denotes its initial condition. The augmented estimator for subsystem i, $i \in \mathbb{I}$, comprises two parts: the first term on the right-hand-side of (3.10) comes from the existing decentralized local estimator, while the remaining part of the right-hand-side of (3.10) is the contribution from the compensator. Note that each augmented estimator should be treated as one system. The augmented estimators need to be initialized and the initialization is independent of the decentralized estimators.

Remark 8. *The proposed distributed state estimation method does not require the existence of a steady state. It has the potential to be used for applications in different fields (e.g., chemical, mechanical and electrical engineering). The use of X_{is} in the proposed method is to describe the commonly used decentralized estimator design approaches. We may consider*

X_{is} as 0 (if interaction between subsystem is totally ignored in the design of decentralized estimators) or a constant (e.g., the average value of the interaction) or the steady state (if it exists and is used in the design of decentralized estimators).

3.2.4 Stability analysis

In this section, we investigate the stability properties of the proposed distributed state estimation network consisting of augmented estimators for local subsystems. Specifically, the evolution of the estimation error generated by augmented estimator i, $i \in \mathbb{I}$, in one sampling period with $K_{i,l}$ determined as in (3.9) is first studied in Proposition 1, then sufficient conditions on the convergence and ultimate boundedness of the estimation error of the entire system state are derived.

We note that $|\dot{x}_i|$, $i \in \mathbb{I}$, is bounded based on the assumptions that the system functions are Lipschitz and the subsystem states are bounded made in Section 3.1.2. We denote the upper bound of $|\dot{x}_i|$ by $\mathcal{M}_i$, $i \in \mathbb{I}$, which is a positive constant (i.e., $|\dot{x}_i| \leq \mathcal{M}_i$). Also, we assume that $K_{i,l}$, is bounded such that $|K_{i,l}| \leq \mathcal{K}_{i,l}$, with $\mathcal{K}_{i,l}$ being a positive constant for all $i \in \mathbb{I}$, $l \in \mathbb{J}_i$. This assumption requires that the Jacobians of $\tilde{f}_i$ and h_l, $i \in \mathbb{I}$, $l \in \mathbb{J}_i$, with respect to their arguments exist. Before presenting the main results, we define a function for each i, $i \in \mathbb{I}$, as follows:

$$\lambda_i(s) = L_{\tilde{f}_i}^{x_i} \mathcal{M}_i s^2 + L_{\tilde{f}_i}^{X_i} \sum_{l \in \mathbb{I}_i} \mathcal{M}_l s^2/2 + \sum_{l \in \mathbb{J}_i} \mathcal{K}_{i,l} \theta_{v_l} s + \mathcal{K}_{i,i} \theta_{v_i} s \tag{3.11}$$

In (3.11), $L_{\tilde{f}_i}^{x_i}$ denotes the Lipschitz constant of function $\tilde{f}_i$ with respect to its first argument, $L_{\tilde{f}_i}^{X_i}$ represents the Lipschitz constant of function $\tilde{f}_i$ with respect to its second argument. Further, let symbol $\mathbb{L}$ denote the largest set that contains only the indices of subsystems, of which the output equations satisfy that $\frac{\partial h_l}{\partial x_l}|_{x_l = z_l(t_k)}$ are full-column rank for $t_{k \geq 0}$. This implies that for any subsystem l, $l \in \mathbb{I}$, if $\frac{\partial h_l}{\partial x_l}|_{x_l = z_l(t_k)}$ is full-column rank for $t_{k \geq 0}$, then l is an element of the set $\mathbb{L}$ (i.e., $l \in \mathbb{L}$). Also, for each $l \in \mathbb{L}$, $\frac{\partial h_l}{\partial x_l}|_{x_l = z_l(t_k)}$ is full-column rank for

$t_{k\geq 0}$.

Proposition 4. *Consider the augmented estimator (3.10) which is designed based on a decentralized local estimator in the form of (3.5) satisfying Assumption 1 for subsystem i, $i \in \mathbb{I}$. If each $K_{i,l}$ for $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$ is determined following (3.9) and is bounded such that $|K_{i,l}| \leq \mathcal{K}_{i,l}$, and if $|\dot{x}_i|$ is bounded as $|\dot{x}_i| \leq \mathcal{M}_i$, $\forall i \in \mathbb{I}$, then the estimation error between the state estimate z_i and the actual subsystem state x_i of the i-th subsystem within one sampling time Δ is bounded for all $x_i \in \mathbb{X}_i$, $i \in \mathbb{I}$, as follows:*

$$\begin{aligned} |e_i(t_{k+1})| \leq\ & \alpha_i(|e_i(t_k)|) + \gamma_i\Big(\max_{\tau} |w_i(\tau)|\Big) + \phi_i\Big(\max_{\tau} |v_i(\tau)|\Big) \\ & + \sum_{p \in \mathbb{J}_i \setminus \mathbb{L}} \psi_{i,p}^{\min}(z_p)\, |e_p(t_k)|\, \Delta + \psi_{i,i}^{\min}(z_i)\, |e_i(t_k)|\, \Delta \\ & + \sum_{l \in \mathbb{J}_i} \Big(H_{X_i}^{\tilde{f}_i} + K_{i,l} H_{x_l}^{h_l}\Big) |e_l(t_k)|^2\, \Delta + H_{x_i}^{\tilde{f}_i} |e_i(t_k)|^2\, \Delta + \lambda_i(\Delta) \end{aligned} \tag{3.12}$$

where $e_i = z_i - x_i$, $i \in \mathbb{I}$, $\alpha_i(\cdot) = \beta_i(\cdot, \Delta)$, γ_i and ϕ_i are class $\mathcal{K}$ functions as defined in (3.6), $\psi_{i,p}^{\min}(z_p) := \min \left\| \frac{\partial \tilde{f}_i}{\partial x_p}\Big|_{x_p = z_p(t_k)} - K_{i,p}(z_p)\, \frac{\partial h_p}{\partial x_p}\Big|_{x_p = z_p(t_k)} \right\|_F$ for $i \in \mathbb{I}$ and $p \in \mathbb{J}_i \setminus \mathbb{L}$, $H_{X_i}^{\tilde{f}_i}$, $H_{x_i}^{\tilde{f}_i}$ and $H_{x_l}^{h_l}$ are positive constants related to the Taylor expansions of $\tilde{f}_i$ and h_l, respectively.

Proof: We consider the augmented estimator i, $i \in \mathbb{I}$, of (3.10) and focus on the time interval $t \in [t_k,\ t_{k+1}]$. The estimation error e_i for subsystem i at time t_{k+1} is calculated as follows:

$$\begin{aligned} e_i(t_{k+1}) &= z_i(t_{k+1}) - x_i(t_{k+1}) \\ &= G_i(\, \{y_i(t_n)\}_{n=0}^{k+1}, z_i(t_0), X_{is}) + \tilde{f}_i\left(z_i(t_k), Z_i(t_k)\right) \Delta \\ &\quad - \tilde{f}_i\left(z_i(t_k), X_{is}\right) \Delta + \sum_{l \in \mathbb{I}_i} K_{i,l}(z_l)(y_l(t_k) - h_l(z_l(t_k))\Delta \\ &\quad - x_i(t_k) - \int_{t_k}^{t_{k+1}} f_i(x_i(\tau), w_i(\tau)) d\tau \\ &\quad - \int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(\tau), X_i(\tau)) d\tau \end{aligned} \tag{3.13}$$

Subtracting/adding $\int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(\tau), X_{is}) d\tau$, $\int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(t_k), X_i(t_k)) d\tau$ and $\int_{t_k}^{t_{k+1}} \tilde{f}_i\left(x_i(t_k), X_{is}\right) d\tau$

from/to (3.13), it is derived that

$$
\begin{aligned}
e_i(t_{k+1}) = &\Big(G_i(\,\{y_i(t_n)\}_{n=0}^{k+1}\,, z_i(t_0), X_{is}) - x_i(t_k) - \int_{t_k}^{t_{k+1}} \left(f_i\left(x_i(\tau), w_i(\tau)\right) + \tilde{f}_i\left(x_i(\tau), X_{is}\right)\right)d\tau\Big) \\
&+ \Big(\tilde{f}_i\left(z_i(t_k), Z_i(t_k)\right)\Delta - \int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(t_k), X_i(t_k))d\tau + \int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(t_k), X_i(t_k))d\tau \\
&\quad - \int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(\tau), X_i(\tau))d\tau + \sum_{l\in\mathbb{J}_i\cap\mathbb{L}} K_{i,l}(z_l)(y_l(t_k) - h_l(z_l(t_k)))\Delta \\
&\quad + \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} K_{i,p}(z_p)(y_p(t_k) - h_p(z_p(t_k)))\Delta\Big) \\
&+ \Big(\int_{t_k}^{t_{k+1}} \tilde{f}_i\left(x_i(\tau), X_{is}\right)d\tau - \int_{t_k}^{t_{k+1}} \tilde{f}_i\left(x_i(t_k), X_{is}\right)d\tau + \int_{t_k}^{t_{k+1}} \tilde{f}_i\left(x_i(t_k), X_{is}\right)d\tau \\
&\quad - \tilde{f}_i\left(z_i(t_k), X_{is}\right)\Delta - K_{i,i}(z_i)(y_i(t_k) - h_i(z_i(t_k)))\Delta\Big)
\end{aligned}
$$

Taking into account that $y_i(t) = h_i(x_i(t)) + v_i(t)$ and $|v_i(t_k)| \leq \theta_{v_i}$, $\forall i \in \mathbb{I}$, it is derived from the above equality based on the Lipschitz properties of $\tilde{f}_i$ that:

$$
\begin{aligned}
|e_i(t_{k+1})| \leq &\Big|G_i(\,\{y_i(t_n)\}_{n=0}^{k+1}\,, z_i(t_0), X_{is}) - x_i(t_k) \\
&\qquad - \int_{t_k}^{t_{k+1}} \left(f_i\left(x_i(\tau), w_i(\tau)\right) + \tilde{f}_i\left(x_i(\tau), X_{is}\right)\right)d\tau\Big| \\
&\quad + \Big|\tilde{f}_i\left(z_i(t_k), Z_i(t_k)\right)\Delta - \int_{t_k}^{t_{k+1}} \tilde{f}_i(x_i(t_k), X_i(t_k))d\tau \\
&\qquad + \sum_{l\in\mathbb{J}_i\cap\mathbb{L}} K_{i,l}(z_l)(h_l(x_l(t_k)) - h_l(z_l(t_k)))\Delta \\
&\qquad + \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} K_{i,p}(z_p)(h_p(x_p(t_k)) - h_p(z_p(t_k)))\Delta\Big| \\
&\quad + \Big|\int_{t_k}^{t_{k+1}} \tilde{f}_i\left(x_i(t_k), X_{is}\right)d\tau - \tilde{f}_i\left(z_i(t_k), X_{is}\right)\Delta \\
&\qquad - K_{i,i}(z_i)(h_i(x_i(t_k)) - h_i(z_i(t_k)))\Delta\Big| \\
&\quad + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{X_i} \int_{t_k}^{t_{k+1}} |x_l(\tau) - x_l(t_k)|\, d\tau + \sum_{l\in\mathbb{J}_i} K_{i,l}(z_l)\theta_{v_l}\Delta \\
&+ 2L_{\tilde{f}_i}^{x_i} \int_{t_k}^{t_{k+1}} |x_i(\tau) - x_i(t_k)|\, d\tau + K_{i,i}(z_l)\theta_{v_i}\Delta
\end{aligned}
\tag{3.14}
$$

where $L_{\tilde{f}_i}^{x_i}$ and $L_{\tilde{f}_i}^{X_i}$ are the Lipschitz constants with respect to the first and second arguments of $\tilde{f}_i$, respectively. Performing Taylor-series expansions, we obtain the following inequalities:

$$\begin{aligned}\tilde{f}_i(x_i(t_k), X_i(t_k)) = \tilde{f}_i(z_i(t_k), Z_i(t_k)) + \left.\frac{\partial \tilde{f}_i}{\partial x_i}\right|_{x_i=z_i(t_k)} (x_i(t_k) - z_i(t_k)) \\ + \sum_{l\in\mathbb{I}_i} \left.\frac{\partial \tilde{f}_i}{\partial x_l}\right|_{x_l=z_l(t_k)} (x_l(t_k) - z_l(t_k)) + H.O.T_{X_i}^{\tilde{f}_i}\end{aligned} \tag{3.15a}$$

$$\tilde{f}_i(x_i(t_k), X_{is}) = \tilde{f}_i(z_i(t_k), X_{is}) + \left.\frac{\partial \tilde{f}_i}{\partial x_i}\right|_{x_i=z_i(t_k)} (x_i(t_k) - z_i(t_k)) + H.O.T_{x_i}^{\tilde{f}_i} \tag{3.15b}$$

$$h_l(x_l(t_k) = h_l(z_l(t_k)) + \left.\frac{\partial h_l}{\partial x_l}\right|_{x_l=z_l(t_k)} (x_l(t_k) - z_l(t_k)) + H.O.T_{x_l}^{h_l} \tag{3.15c}$$

where $H.O.T_{X_i}^{\tilde{f}_i}$, $H.O.T_{x_i}^{\tilde{f}_i}$ and $H.O.T_{x_l}^{h_l}$ denote the Taylor remainders [52] of the Taylor-series expansions of functions $\tilde{f}_i$ and h_l. The remainders satisfy the following constraints:

$$\begin{aligned} H.O.T_{X_i}^{\tilde{f}_i} &\leq \sum_{l\in\mathbb{J}_i} H_{X_i}^{\tilde{f}_i} \left|x_l(t_k) - z_l(t_k)\right|^2, \\ H.O.T_{x_i}^{\tilde{f}_i} &\leq H_{x_i}^{\tilde{f}_i} \left|x_i(t_k) - z_i(t_k)\right|^2, \\ H.O.T_{x_l}^{h_l} &\leq H_{x_l}^{h_l} \left|x_l(t_k) - z_l(t_k)\right|^2 \end{aligned} \tag{3.16}$$

for all $x_i \in \mathbb{X}_i$ with $H_{X_i}^{\tilde{f}_i}$, $H_{x_i}^{\tilde{f}_i}$, $i \in \mathbb{I}$, and $H_{x_l}^{h_l}$, $l \in \mathbb{J}_i$, being finite positive constants when $|x_i(t_k) - z_i(t_k)|$ is bounded.

Let us define $S_i(t_k) := \tilde{f}_i\left(z_i(t_k), Z_i(t_k)\right) - \tilde{f}_i\left(x_i(t_k), X_i(t_k)\right) + \sum_{l\in\mathbb{J}_i\cap\mathbb{L}} K_{i,l}(z_l)(h_l(x_l(t_k)) - h_l(z_l(t_k))) + \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} K_{i,p}(z_p)\left(h_p(x_p(t_k)) - h_p(z_p(t_k))\right)$ and define $P_i(t_k) := \tilde{f}_i(x_i(t_k), X_{is}) - \tilde{f}_i(z_i(t_k), X_{is}) - K_{i,i}(z_i)(h_i(x_i(t_k)) - h_i(z_i(t_k)))$. Taking (3.15) into consideration, it is calcu-

lated that:

$$
\begin{aligned}
S_i(t_k) = &- \left.\frac{\partial \tilde{f}_i}{\partial x_i}\right|_{x_i=z_i(t_k)} (x_i(t_k) - z_i(t_k)) - \sum_{l\in\mathbb{I}_i} \left.\frac{\partial \tilde{f}_i}{\partial x_l}\right|_{x_l=z_l(t_k)} (x_l(t_k) - z_l(t_k)) \\
&+ \sum_{l\in\mathbb{J}_i\cap\mathbb{L}} K_{i,l}(z_l) \left.\frac{\partial h_l}{\partial x_l}\right|_{x_l=z_l(t_k)} (x_l(t_k) - z_l(t_k)) \\
&+ \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} K_{i,p}(z_p) \left.\frac{\partial h_p}{\partial x_p}\right|_{x_p=z_p(t_k)} (x_p(t_k) - z_p(t_k)) + \sum_{l\in\mathbb{J}_i} K_{i,l}(z_l) H.O.T_{x_l}^{h_l} - H.O.T_{X_i}^{\tilde{f}_i}
\end{aligned}
\tag{3.17a}
$$

$$
P_i(t_k) = \left.\frac{\partial h_l}{\partial x_l}\right|_{x_i=z_i(t_k)} (x_i(t_k) - z_i(t_k)) - K_{i,i}(z_i)(h_i(x_i(t_k)) - h_i(z_i(t_k))) + H.O.T_{x_i}^{\tilde{f}_i} \tag{3.17b}
$$

If $K_{i,l}$ in (3.17a) and $K_{i,i}$ in (3.17b) are calculated following (3.9), then based on (3.16) and (3.17), in the worst case scenario, we can obtain that

$$
\begin{aligned}
&|S_i(t_k)| \le \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} \psi_{i,p}^{\min} |x_p(t_k) - z_p(t_k)| + \sum_{l\in\mathbb{J}_i} \left(H_{X_i}^{\tilde{f}_i} + K_{i,l} H_{x_l}^{h_l}\right) |x_l(t_k) - z_l(t_k)|^2 \\
&|P_i(t_k)| \le \psi_{i,i}^{\min} |x_i(t_k) - z_i(t_k)| + H_{x_i}^{\tilde{f}_i} |x_i(t_k) - z_i(t_k)|^2
\end{aligned}
\tag{3.18}
$$

From (3.14) and (3.18), we derive the following inequality:

$$
\begin{aligned}
|e_i(t_{k+1})| \le &\Big| G_i(\{y_i(t_n)\}_{n=0}^{k+1}, z_i(t_0), X_{is}) - x_i(t_k) \\
&\quad - \int_{t_k}^{t_{k+1}} \left(f_i\left(x_i(\tau), w_i(\tau)\right) + \tilde{f}_i\left(x_i(\tau), X_{is}\right)\right) d\tau \Big| \\
&+ \sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} \psi_{i,p}^{\min} |e_p(t_k)| \Delta + \psi_{i,i}^{\min}(t_k) |e_i(t_k)| \Delta \\
&+ \sum_{l\in\mathbb{J}_i} \left(H_{X_i}^{\tilde{f}_i} + K_{i,l} H_{x_l}^{h_l}\right) |e_l(t_k)|^2 \Delta + H_{x_i}^{\tilde{f}_i} |e_i(t_k)|^2 \Delta \\
&+ 2L_{\tilde{f}_i}^{x_i} \int_{t_k}^{t_{k+1}} |x_i(\tau) - x_i(t_k)| \, d\tau + K_{i,i}(z_l)\theta_{v_i}\Delta \\
&+ \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{X_i} \int_{t_k}^{t_{k+1}} |x_l(\tau) - x_l(t_k)| \, d\tau + \sum_{l\in\mathbb{J}_i} K_{i,l}(z_l)\theta_{v_l}\Delta
\end{aligned}
\tag{3.19}
$$

where $e_l(t_k) = z_l(t_k) - x_l(t_k)$, $l \in \mathbb{J}_i$. Consider the term $\int_{t_k}^{t_{k+1}} |x_i(\tau) - x_i(t_k)|\, d\tau$, we have that $\int_{t_k}^{t_{k+1}} |x_i(\tau) - x_i(t_k)|\, d\tau \leq \int_{t_k}^{t_{k+1}} \mathcal{M}_i(\tau - t_k) d\tau = \int_0^{\Delta} \mathcal{M}_i \tilde{\tau} d\tilde{\tau} = \frac{1}{2}\mathcal{M}_i \Delta^2$ considering the boundedness of $|\dot{x}_i|$, $i \in \mathbb{I}$. Based on the property of the existing decentralized estimator for subsystem i that is stated in Assumption 3, and taking into account the boundedness of $K_{i,l}$ for all $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$, the following inequality can be obtained:

$$
\begin{aligned}
|e_i(t_{k+1})| \leq\ & \beta_i\left(|e_i(t_k)|, \Delta\right) + \gamma_i\Big(\max_{\tau} |w_i(\tau)|\Big) + \phi_i\Big(\max_{\tau} |v_i(\tau)|\Big) \\
& + \sum_{p \in \mathbb{J}_i \setminus \mathbb{L}} \psi_{i,p}^{\min} |e_p(t_k)|\, \Delta + \psi_{i,i}^{\min}(t_k)\, |e_i(t_k)|\, \Delta \\
& + \sum_{l \in \mathbb{J}_i} \left(H_{X_i}^{\tilde{f}_i} + \mathcal{K}_{i,l} H_{x_l}^{h_l}\right) |e_l(t_k)|^2\, \Delta + H_{x_i}^{\tilde{f}_i} |e_i(t_k)|^2\, \Delta \\
& + L_{\tilde{f}_i}^{x_i} \mathcal{M}_i \Delta^2 + L_{\tilde{f}_i}^{X_i} \sum_{l \in \mathbb{I}_i} \mathcal{M}_l \Delta^2/2 + \sum_{l \in \mathbb{J}_i} \mathcal{K}_{i,l} \theta_{v_l} \Delta + \mathcal{K}_{i,i} \theta_{v_i} \Delta
\end{aligned}
\tag{3.20}
$$

Given a fixed sampling period Δ and the definition of function λ_i, this completes the proof of Proposition 4. $\square$

Proposition 4 provides the upper bound of the (open-loop evolution of the) estimation error for each subsystem within one sampling time Δ. Note that measurement sampling and information exchange occur at each sampling instant, and Δ is the time interval within which there is no measurement feedback or information exchange between subsystems. From Eq.(3.19) to Eq.(3.20), we find the upper bound on the estimation error evolution over one sampling time without feedback. In what follows, we provide sufficient conditions on the convergence and ultimate boundedness of the estimation error of the proposed distributed state estimation network.

Theorem 3. *Consider nonlinear system (3.1) consisting of m subsystems with subsystem measurements y_i, $i \in \mathbb{I}$. If Assumption 1 to Assumption 3 are all satisfied, and if augmented estimators are designed following (3.10) with correction gains $K_{i,l}$ ($i \in \mathbb{I}$ and $l \in \mathbb{I}_i$) determined following (3.9) and bounded as $|K_{i,l}| \leq \mathcal{K}_{i,l}$, and if $|\dot{x}_i|$ is bounded as $|\dot{x}_i| \leq \mathcal{M}_i$, $\forall i \in \mathbb{I}$, and if the initial error satisfies $|e_i(t_0)| \leq \rho_i$, $i \in \mathbb{I}$, and if there exist positive scalars ϱ_i, ρ_i*

such that $0 < \varrho_i < \rho_i$, *and positive constants* $\epsilon_i > 0$, $\forall i \in \mathbb{I}$, *such that:*

$$\begin{aligned}\varrho_i - \Big(\mathcal{Q}_i(|\varrho_i|) + \gamma_i(\max_{\tau}|w_i(\tau)|) + \phi_i(\max_{\tau}|v_i(\tau)|)\Big) \\ - \Big(\sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} \psi_{i,p}^{\max}\rho_p\Delta + \psi_{i,i}^{\max}\rho_i\Delta\Big) - \Big(\sum_{l\in\mathbb{J}_i} R_{il}\rho_l^2\Delta + H_{x_i}^{\tilde{f}_i}\rho_i^2\Delta + \lambda_i(\Delta)\Big) \geq \epsilon_i\end{aligned} \tag{3.21}$$

where $\mathcal{Q}_i$ *is a concave function satisfying* $\mathcal{Q}_i(|e_i|) \geq \alpha_i(|e_i|)$ *for all* $|e_i| \leq \rho_i$, *and* $R_{il} := H_{X_i}^{\tilde{f}_i} + \mathcal{K}_{i,l}H_{x_l}^{h_l}$, $i \in \mathbb{I}$, $l \in \mathbb{J}_i$, $\psi_{i,p}^{\max} := \max\{\psi_{i,p}^{\min}(z_p) : z_p \in \mathbb{Z}_p\}$, $i \in \mathbb{I}$, $p \in \mathbb{J}_i \setminus \mathbb{L}$ *with* $\mathbb{Z}_i$ *denoting the smallest compact set that bounds* z_i *such that* $z_i \in \mathbb{Z}_i$ *for all* $x_i \in \mathbb{X}_i$ *and* $|z_i - x_i| \leq \rho_i$, $i \in \mathbb{I}$, *then the estimation error* $|e_i| = |z_i - x_i|$ *given by augmented estimator in (3.10) is a decreasing sequence for all* $\varrho_i \leq |e_i| \leq \rho_i$, $i \in \mathbb{I}$ *and is ultimately bounded as follows:*

$$\lim_{k\to\infty}\sup|e_i(t_k)| \leq \rho_i^{\min} \tag{3.22}$$

for $i \in \mathbb{I}$ *with* $\rho_i^{\min} \leq \rho_i$ *being defined as* $\rho_i^{\min} := \max\{|e_i(t_k+\Delta)|: |e_i(t_k)| \leq \varrho_i\}$ *for all* $|e_i(t_0)| \leq \rho_i$ *and* $x_i \in \mathbb{X}_i$. *The ultimate boundedness on the estimation error for each subsystem* i, $i \in \mathbb{I}$, *also implies that the entire system state estimation error is ultimately bounded.*

Proof: We first prove that the estimation error $|e_i| = |z_i - x_i|$ of augmented estimator i in the form of (3.10) for subsystem i, $i \in \mathbb{I}$, is decreasing and ultimately bounded after finite sampling periods. Then, we show that the estimation error of the entire system is a decreasing sequence and is eventually bounded.

Specifically, taking into account the property of the concave function $\mathcal{Q}_i$, $i \in \mathbb{I}$, if there exist scalars ϱ_i and ρ_i, $i \in \mathbb{I}$, satisfying (3.21), then the following inequality

$$\begin{aligned}|e_i| - \Big(\mathcal{Q}_i(|e_i|) + \gamma_i(\max_{\tau}|w_i(\tau)|) + \phi_i(\max_{\tau}|v_i(\tau)|)\Big) \\ - \Big(\sum_{p\in\mathbb{J}_i\setminus\mathbb{L}} \psi_{i,p}^{\max}|e_p|\,\Delta + \psi_{i,i}^{\max}|e_i|\,\Delta\Big) \\ - \Big(\sum_{l\in\mathbb{J}_i} R_{il}|e_l|^2\,\Delta + H_{x_i}^{\tilde{f}_i}|e_i|^2\,\Delta + \lambda_i(\Delta)\Big) \geq \epsilon_i\end{aligned} \tag{3.23}$$

holds for all $\varrho_i \leq |e_i| \leq \rho_i$, $i \in \mathbb{I}$, and $|e_l| \leq \rho_l$, $l \in \mathbb{J}_i$. Taking into consideration (3.12), (3.23), the property that $\mathcal{Q}_i(|e_i|) \geq \alpha_i(|e_i|)$ for all $|e_i| \leq \rho_i$ and the initial condition that $|e_i(0)| \leq \rho_i$, we can derive the following inequality:

$$|e_i(t_{k+1})| \leq |e_i(t_k)| - \epsilon_i \tag{3.24}$$

for all $\varrho_i \leq |e_i| \leq \rho_i$, $i \in \mathbb{I}$. Based on (3.24), if $|e_i| \geq \varrho_i$ for the time interval between t_0 and t_k, then $|e(t_q)|$, $t_0 < t_q \leq t_k$ is a decreasing sequence described as follows:

$$|e_i(t_q)| \leq |e_i(t_0)| - q\epsilon_i \tag{3.25}$$

for all $\varrho_i \leq |e_i(t_q)| \leq \rho_i$. This implies that $|e_i|$ decreases at every sampling time and will eventually become smaller than ϱ_i within finite steps. Once condition $|e_i| < \varrho_i$ is satisfied, $|e_i|$ will not decrease at every next sampling instant. However, it will remain to satisfy $|e_i(t)| \leq \rho_i^{\min}$ because of the nature of $\rho_i^{\min}$; that is, $\lim_{k\to\infty} \sup |e_i(t_k)| \leq \rho_i^{\min}$.

The boundedness of $|e_i|$, $i \in \mathbb{I}$, implies the ultimate boundedness of the entire system state estimation error, which can be demonstrated by the inequality $\lim_{k\to\infty} \sup |e(t_k)| \leq \sum_{i=1}^{m} \rho_i^{\min}$. This completes the proof of Theorem 3. □

Remark 9. *In condition (3.21) of Theorem 3, the function $\mathcal{Q}_i(\cdot)$ is used as a concave approximation of $\alpha_i(\cdot)$, $i \in \mathbb{I}$. The term $H_{x_i}^{\tilde{f}_i} |\rho_i|^2 \Delta$ characterizes the influence of the i-th subsystem itself involved in the interaction term $\tilde{f}_i$ after interaction compensation. The term $\sum_{l\in\mathbb{J}_i} R_{il} |\rho_l|^2 \Delta$ depicts the effect of the subsystem interaction between subsystem i and its upstream interacting subsystems after interaction compensation. The function $\lambda_i(\Delta)$ characterizes the effects on the upper bound of the subsystem estimation error that are caused by the communication frequency and the magnitudes of the subsystem disturbances and measurement noise. Note that condition (3.21) can be satisfied if the convergence speed of the existing decentralized estimator for each subsystem i is sufficiently fast, the communication frequency is sufficiently high (i.e., Δ is sufficiently small) and the upper bounds on subsystem*

disturbances and measurement noise are sufficiently small.

Remark 10. *It is noted that while the conditions in Theorem 1 are derived in a constructive way, they are conservative since the worst case scenario is considered. The conditions reveal the interplay between the parameters and give us directions on how to obtain a successful design. This is actually the case with many results for nonlinear systems in the literature that consider the worst case scenarios (e.g., using Lipschitz properties and upper bounds of uncertainties) as discussed and illustrated in [53, 54]. One prerequisite for the satisfaction of Theorem 3 is that the initial error is not arbitrarily large ($|e_i(t_0)| \leq \rho_i$) due to the condition $\mathcal{Q}_i(|e_i|) \geq \alpha_i(|e_i|)$, $i \in \mathbb{I}$. Moreover, an implication of condition (3.21) is that the ultimate upper bound of the estimation error depends on a few factors including the magnitudes of the disturbances and measurement noise as well as the communication interval between estimators. In particular, smaller upper bounds on the system disturbances and measurement noise and the use of a smaller communication interval (i.e., Δ) can lead to a smaller value of ϱ_i, such that the condition in (3.21) can be satisfied for $\epsilon_i > 0$, $i \in \mathbb{I}$. Correspondingly, the ultimate upper bound of the estimation error given by the distributed scheme decreases. In practice, the design parameters should be tuned or estimated following these conditions based on simulations. This will be illustrated in Section 3.3.2.*

Remark 11. *Note that the class $\mathcal{K}$ function $\alpha_i(|e_i|)$ could be either a convex increasing function or a concave increasing function with respect to its argument. The two different case scenarios should be handled via different approaches which will increase the complexity for analysis. In order to simplify the proof and unify the conditions, we introduce the assumption on the existence of a concave function that satisfies the condition $Q_i(|e_i|) \geq \alpha_i(|e_i|)$ for all $|e_i| \leq \rho_i$, $i \in \mathbb{I}$.*

Remark 12. *We note that the proposed approach can be extended to the case that local estimators are continuous-time estimators (e.g., [55, 30, 56, 57]). In this case, we need to assume that the output measurements of each subsystem are immediately and continuously*

available to the corresponding estimator, while the subsystem estimators communicate to exchange information at discrete-time instants. We point out that this extension of the proposed approach provides an alternative way to design nonlinear observers for systems with moderately large scales (which is in general a challenging task based on existing centralized approaches) by connecting small-size local observers together. This approach will be illustrated via the application to a chemical process in Section 3.3.2.

Remark 13. *In this work, we consider an autonomous system model in Eq.(3.1) for analytical convenience. If the input is known and the corresponding decentralized estimators satisfy Assumption 1, the proposed approach can be extended to handle manipulated inputs in a straightforward manner.*

Remark 14. *Communication delays and packet dropouts are important factors that may affect the performance of a distributed state estimation network. Taking into account these factors will make the proposed approach more applicable, and they will be investigated in the future work.*

3.3 Applications of the proposed approach

In this section, we apply the proposed approach to three process examples to illustrate its effectiveness.

3.3.1 Application to the illustrative example

First, we revisit the illustrative example introduced in Section 3.1.4 to show that we can achieve much improved estimation performance when the proposed distributed state estimation approach is applied. The values of the variables and parameters of this process are presented in Table 3.2. Accordingly, we calculate the values of the Lipschitz constants with respect to the nonlinear functions of the model (3.7) as $L_{\tilde{f}_1}^{X_1} = 0$ (as the dynamics of the first subsystem is not affected by the second subsystem) and $L_{\tilde{f}_2}^{X_2} = 10.2621$. We consider two

Table 3.2: Parameters of process variables

$F_0 = 38.996\ \mathrm{m}^3/\mathrm{h}$	$\Delta H_1 = -5.0 \times 10^4\ \mathrm{kJ/kmol}$
$F_1 = 38.996\ \mathrm{m}^3/\mathrm{h}$	$\Delta H_2 = -5.2 \times 10^4\ \mathrm{kJ/kmol}$
$F_2 = 65.021\ \mathrm{m}^3/\mathrm{h}$	$\Delta H_3 = -5.05 \times 10^4\ \mathrm{kJ/kmol}$
$F_3 = 26.025\ \mathrm{m}^3/\mathrm{h}$	$k_1 = 3.0 \times 10^6\ \mathrm{h}^{-1}$
$V_1 = 5.1\ \mathrm{m}^3$	$k_2 = 3.0 \times 10^5\ \mathrm{h}^{-1}$
$V_2 = 3.8\ \mathrm{m}^3$	$k_3 = 3.0 \times 10^5\ \mathrm{h}^{-1}$
$T_0 = 300.0\ \mathrm{K}$	$E_1 = 5.0 \times 10^4\ \mathrm{kJ/kmol}$
$T_{03} = 300.0\ \mathrm{K}$	$E_2 = 7.5 \times 10^4\ \mathrm{kJ/kmol}$
$R = 8.314\ \mathrm{kJ/kmol \cdot K}$	$E_3 = 7.53 \times 10^4\ \mathrm{kJ/kmol}$
$C_{A0} = 3.9\ \mathrm{kmol/m}^3$	$\rho = 1000.0\ \mathrm{kg/m}^3$
$C_{A03} = 2.05\ \mathrm{kmol/m}^3$	$c_p = 0.231\ \mathrm{kJ/(kg \cdot K)}$
$Q_1 = 1.02 \times 10^4\ \mathrm{kJ/h}$	$Q_2 = 9.8 \times 10^3\ \mathrm{kJ/h}$
$R = 8.314\ \mathrm{kJ/kmol \cdot K}$	

different state estimation schemes: the decentralized composite EKF-MHE scheme discussed in Section 3.1.4, and a formed distributed state estimation scheme based on the existing decentralized scheme. In the simulations for the two schemes, both the EKF and the MHE estimators are evaluated every 7.2s; that is, $\Delta = 7.2$s. We assume that the measured outputs of the two CSTRs are subject to random noise, which is generated following normal distribution with zero mean and standard deviation 0.4 for both the two subsystems. Random process disturbances in the dynamics of the process are also added. In addition, a bounded disturbance in the feed stream flow F_3 to reactor 2 is considered. In the simulations, this disturbance is generated following normal distribution with zero mean value and standard deviation $0.5\mathrm{m}^3/\mathrm{h}$, and is constrained by $-0.9\ \mathrm{m}^3/\mathrm{h}$ and $0.95\ \mathrm{m}^3/\mathrm{h}$. For the EKF designed for the first CSTR, two weighting matrices associated with the process disturbances and the measurement noise are selected as: $Q_1 = \mathrm{diag}(\,[25, 0.04]^T\,)$ and $R_1 = 0.16$, respectively. The initial error covariance matrix is picked as $P(0) = \mathrm{diag}(\,[100, 100]^T\,)$. In terms of the MHE for the second CSTR, the size of the estimation horizon is $N = 8$, the two weighting

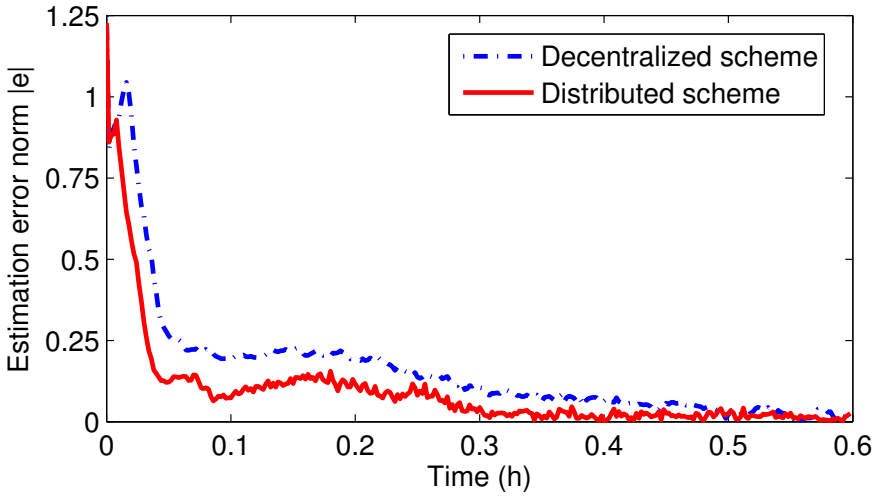

Figure 3.4: Trajectories of the normalized estimation error norms of different schemes for the second CSTR of the cascade chemical process.

matrices associated with the disturbances and measurement noise for the MHE are selected as: $Q_2 = \text{diag}(\,[16, 0.04]^T\,)$ and $R_2 = 0.16$, respectively. Note that no arrival cost is used in the MHE.

We assume that the initial state of the process is: $x_0 = [280.7160\text{K}\ 2.2977\text{kmol/m}^3\ 270.8933\text{K}\ 2.2837\text{kmol/m}^3]^T$, and the initial guess for both the decentralized composite EKF-MHE scheme and the formed distributed scheme is: $z_0 = [300.7100\text{K}\ 2.6900\text{kmol/m}^3\ 322.3500\text{K}\ 3.2930\text{kmol/m}^3]^T$.

We also make performance comparisons of different schemes, including the proposed distributed scheme, the existing decentralized scheme, and a distributed MHE scheme with auxiliary observers proposed in [15]. To evaluate the estimation performance of different schemes, we use the Euclidean norm of the normalized estimation error defined as $|e|$. Specifically, the normalized estimation error norm is calculated at each time t_k as follows:

$$|e(t_k)| = \sqrt{\sum_{i \in n_x} \left(\frac{z_i(t_k) - x_i(t_k)}{\max\limits_{t_n \in [t_0, \infty)} (z_i(t_n) - x_i(t_n))} \right)^2} \tag{3.26}$$

We note that when applying the above criterion to different state estimation schemes, the

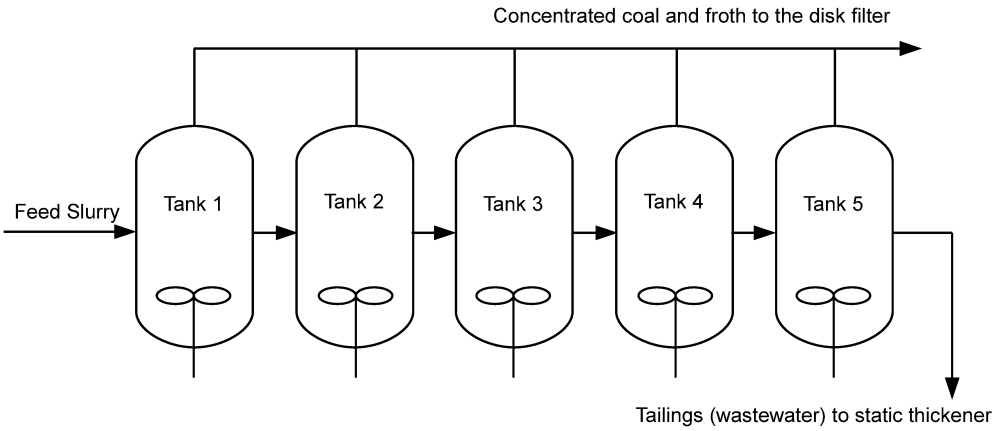

Figure 3.5: A schematic of the forth-flotation-unit process.

maximum estimation error for each system state is picked as the largest value for the corresponding state among all the considered schemes to make the normalized errors unbiased.

Due to the cascade nature of the process, the estimates for the first CSTR given by the EKF estimator are exactly the same for both the decentralized scheme and the distributed scheme. To make comparisons on the estimation performance, we only need to take into account the state estimates for the second reactor given by the two schemes. The trajectories of the normalized estimation error norms of the second CSTR based on the three considered schemes are presented in Figure 3.4. The mean values of the normalized estimation error norms of the proposed distributed scheme, the distributed MHE in [15] and the existing decentralized scheme are 0.0935, 0.0917 are 0.1614, respectively. The results indicate that the formed distributed state estimation scheme can improve the estimation performance for the second reactor significantly by approximately 42.07% compared to the existing decentralized scheme in terms of the mean estimation error norm. Also, the two distributed estimation schemes give very close performance. While the performance of the two distributed schemes is similar, the proposed approach is 40.92% more computationally efficient compared with the distributed MHE design in [15].

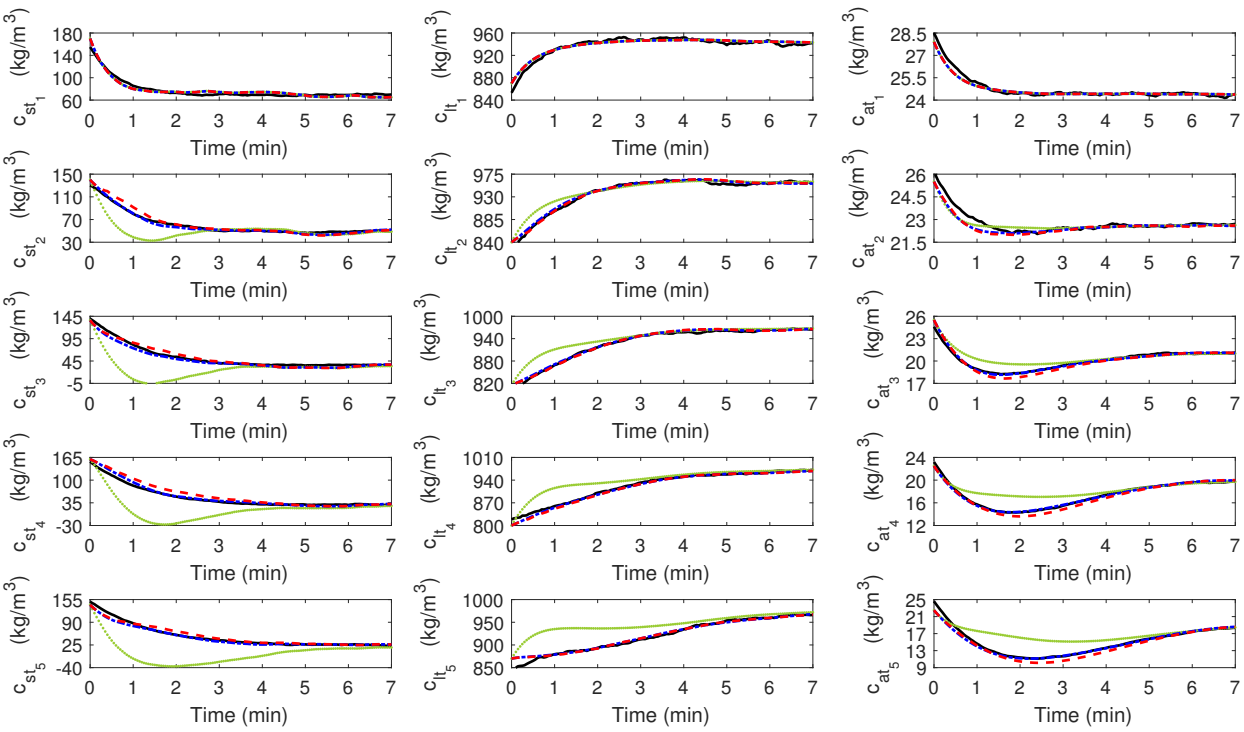

Figure 3.6: Trajectories of the actual states (solid black lines), the state estimates of the decentralized scheme (dotted green lines), the state estimates of the distributed scheme with $\Delta = 30s$ (dashed red lines), and the state estimates of the distributed scheme with $\Delta = 6s$ (dash-dotted blue lines).

3.3.2 Application to a froth flotation process example

As mentioned in Remark 12, the proposed method can be used in the design of nonlinear observers for moderately large-scale nonlinear processes. To illustrate this point, we apply the proposed approach to a process of froth flotation units that is used to clean and recover fine coal produced by Coal Handling and Preparation Plants (CHPPs). The froth-flotation-unit process is comprised of five interconnected tanks. A schematic diagram of this process is presented in Figure 3.5. A model description is presented in Appendix I at the end of this chapter.

It is assumed that the states c_{lt_i} and c_{at_i}, $i \in \mathbb{I}$, are the measured outputs. The entire process is divided into five subsystems according to the five vessels. We consider that for each vessel, there exists a decentralized estimator developed based on the high-gain observer design in [30], such that each decentralized estimator can provide state estimates for the

Table 3.3: Mean and maximum error norms of the three different schemes.

	Decentralized scheme	Distributed scheme with $\Delta = 30$s	Distributed scheme with $\Delta = 6$s
Mean norm	1.6117	0.7285	0.5658
Max norm	3.0163	2.2078	2.2078

corresponding subsystem based on continuous measurements of c_{lt_i} and c_{at_i}. To achieve improved estimation performance, we apply the proposed approach to the existing decentralized scheme to form a distributed state estimation network, where five augmented estimators that exchange information with each other every sampling period Δ are designed.

A set of simulations is carried out. Specifically, we consider three different schemes: the existing decentralized state estimation scheme; the formed distributed state estimation scheme with a low communication frequency ($\Delta = 30s$); the formed distributed state estimation scheme with a high communication frequency ($\Delta = 6s$). In the simulations, random measurement noise and system disturbances that follow normal distribution are added. The initial state of the entire process is assumed to be: $x_0 = [155\text{kg/m}^3,\ 853\text{kg/m}^3,\ 28.5\text{kg/m}^3,\ 130\text{kg/m}^3,\ 825.5\text{kg/m}^3,\ 26.2\text{kg/m}^3,\ 139.5\text{kg/m}^3,\ 797\text{kg/m}^3,\ 24.6\text{kg/m}^3,\ 153\text{kg/m}^3,\ 819\text{kg/m}^3,\ 23.2\text{kg/m}^3,\ 149\text{kg/m}^3,\ 845\text{kg/m}^3,\ 24.6\text{kg/m}^3]^T$, while the initial guess for the distributed state estimation system is set to be: $z_0 = [170\ \text{kg/m}^3,\ 870\text{kg/m}^3,\ 27.9\text{kg/m}^3,\ 140\text{kg/m}^3,\ 840\text{kg/m}^3,\ 25.5\text{kg/m}^3,\ 136\text{kg/m}^3,\ 815\text{kg/m}^3,\ 25.5\text{kg/m}^3,\ 160\text{kg/m}^3,\ 800\text{kg/m}^3,\ 22.5\text{kg/m}^3,\ 140\text{kg/m}^3,\ 870\text{kg/m}^3,\ 22.5\text{kg/m}^3]^T$.

The actual system states, together with the state estimates of the three different schemes are shown in Figure 3.6. The results imply that within the existing decentralized framework, the state estimates of all the states can converge to the actual system states when the entire process is at the steady state. However, during the transient part, the estimation errors are relatively large, which leads to unsatisfying overall estimation performance. By contrast, the distributed state estimation schemes with different communication frequencies are both able to provide much improved state estimation performance during the transient

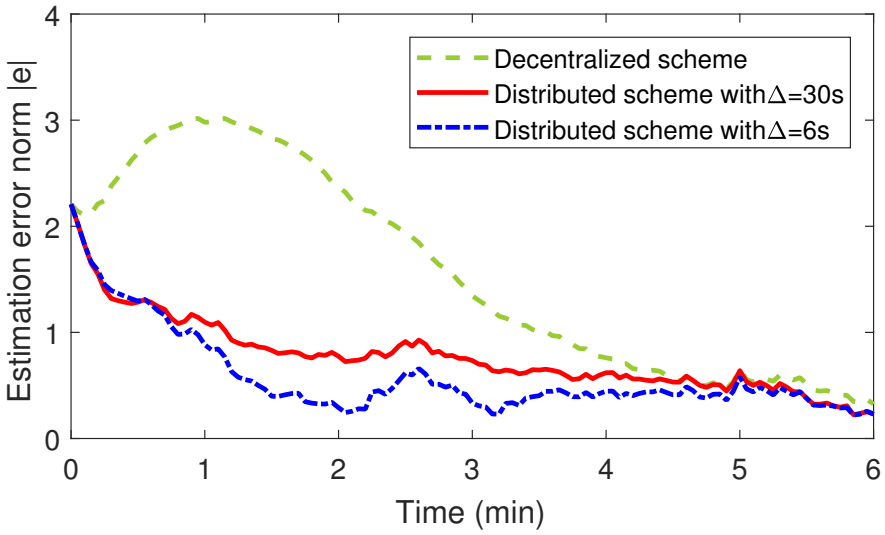

Figure 3.7: Trajectories of the normalized estimation errors of the decentralized scheme (dashed green lines), the distributed scheme with $\Delta = 30$s (solid red lines), and the state estimates of the distributed scheme with $\Delta = 6$s (dash-dotted blue lines).

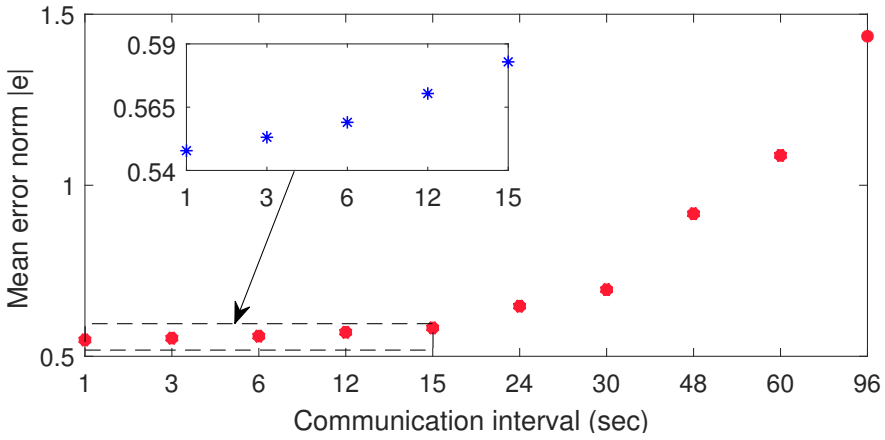

Figure 3.8: Mean estimation errors of the distributed scheme with different communication intervals.

part compared with the decentralized scheme. This is because that the dynamics of the subsystem interaction is explicitly taken into account via the compensators designed for the subsystems. To further compare the estimation performance of the three schemes, the trajectories of the Euclidean norms of the normalized estimation errors for the three schemes are given in Figure 3.7, while the corresponding mean and maximum values of the Euclidean error norms are presented in Table 3.3.

The results further verify that the formed distributed state estimation network obtained

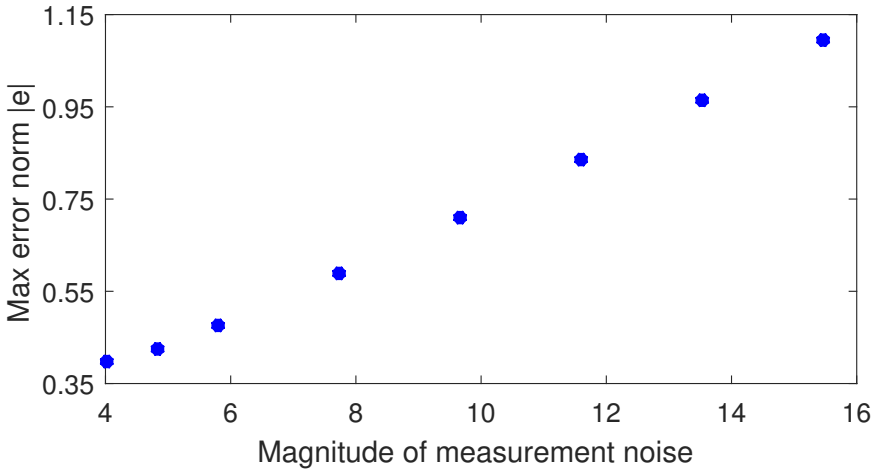

Figure 3.9: The relationship between the magnitude of measurement noise and the mean value of the estimation error norm subject to $\sqrt{\sum_{i=1}^{5} \theta_{w_i}^2} = 7.4559$.

via the proposed approach can give much improved estimation performance. Both the mean value and the maximum value of the estimation error norm of each distributed scheme are significantly smaller than the corresponding values of the decentralized counterpart. Moreover, by increasing the communication frequency, the overall estimation performance can be further improved. By reducing the sampling period Δ from 30s to 6s, the mean value of the normalized estimation error norm has decreased by 22.33%. Note that the maximum values of the estimation error norms of the two distributed schemes are the same. This is because that the maximum errors of the two schemes both occur at the initial stage and the initial estimation errors of the two schemes are identical.

To further demonstrate the relationship between the communication frequency and the estimation accuracy, we also carry out simulations for the proposed distributed approach with different communication intervals between subsystems. The mean estimation error norms at different communication intervals are shown in Figure 3.8. The result implies that in the presence of disturbances and measurement noise, increasing the communication frequency can improve the estimation accuracy.

Finally, simulations are performed to reveal the relationship between the magnitudes of process disturbances/measurement noise and the ultimate upper bound of the estimation er-

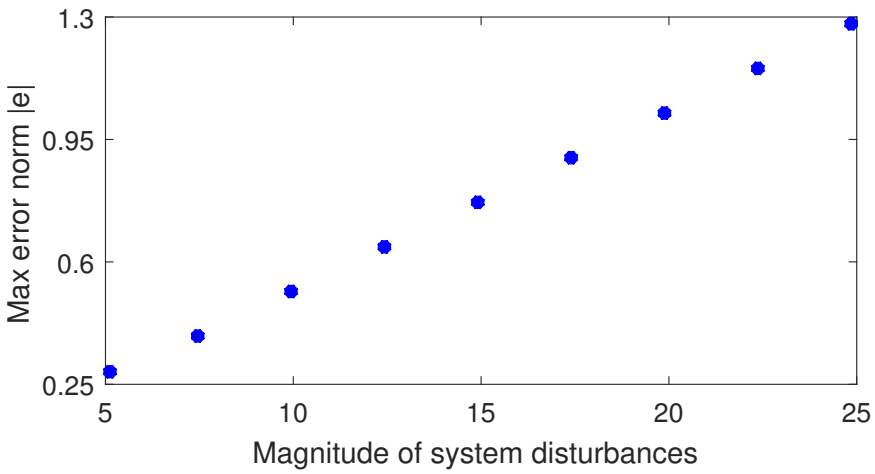

Figure 3.10: The relationship between the magnitude of system disturbances and the maximum value of the estimation error norm subject to $\sqrt{\sum_{i=1}^{5} \theta_{v_i}^2} = 2.8894$.

ror norm. The initial estimation error is set to be zero in this set of simulations. The results are given in Figure 3.9 and Figure 3.10. With the increase of the magnitude of the measurement noise/system disturbances, the upper bound of the estimation error norm increases correspondingly. The results are consistent with the analysis presented in Remark 10.

3.3.3 Application to a hybrid-tank plant

The proposed approach is also implemented to a hybrid-tank plant [58]. This plant comprises three interconnected water tanks of the same size. A schematic of the plant is shown in Figure 3.11. Each of the three tanks is equipped with a pipe at the bottom for discharge. Water enters the left tank and the right tank through two pipes by manipulating two pumps below the tanks. The left (right) tank is connected to the middle tank in three ways: one pipe at the bottom of the tanks and two additional pipes at higher levels. In this experiment, only the valves of the pipes at the bottom layer (i.e., V_5 to V_9 in Figure 3.11) are open, while all the valves of the higher-level pipes (i.e., i.e., V_1 to V_4 in Figure 3.11) are kept closed. The water levels in all three tanks can be measured online using differential pressure (DP) sensors. The DP sensors provide water level measurements every 1 sec. Taking into account the physical behaviors of the plant, a model that describes the system dynamics is established

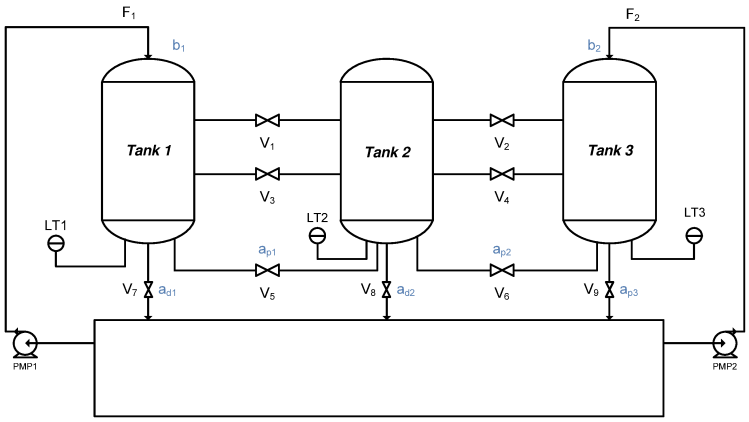

Figure 3.11: A schematic of the hybrid-tank plant.

as follows [59]:

$$
\begin{aligned}
\dot{h}_1 &= \frac{1}{S}\left(b_1 F_1 - S_{p_1} a_{p_1}\sqrt{2g(h_1-h_2)} - S_{d_1} a_{d_1}\sqrt{2gh_1}\right) \\
\dot{h}_2 &= \frac{1}{S}\Big(S_{p_1} a_{p_1}\sqrt{2g(h_1-h_2)} + S_{p_2} a_{p_2}\sqrt{2g(h_3-h_2)} \\
&\qquad - S_{d2} a_{d2}\sqrt{2gh_2}\Big) \\
\dot{h}_3 &= \frac{1}{S}\left(b_1 F_2 - S_{p_2} a_{p_2}\sqrt{2g(h_3-h_2)} - S_{d_3} a_{d_3}\sqrt{2gh_3}\right)
\end{aligned}
\tag{3.27}
$$

In Eq.(3.27), h_i denotes the water level in the i-th tank (counting from left to right), $i = 1,\ 2,\ 3$; F_1 and F_2 denote the flow rates of the water entering the left and the right tank, respectively; S_{d_i} is the cross-sectional area of the pipe for discharge of the i-th tank, $i = 1,\ 2,\ 3$; S_{p_i} is the cross-sectional area of the pipe between the i-th tank and the $i+1$-th tank, $i = 1,\ 2$; S represents the cross-sectional area of each tank; b_1 and b_2 are the tuning coefficients with respect to the inlet flow rates to the left and the right tank, respectively; a_{p_1}, a_{p_2} and a_{d_i}, $i = 1,\ 2,\ 3$, are the tuning coefficients with respect to the pipe values. The above tuning coefficients are identified to improve the accuracy of the model. These coefficients have been identified and reported in [59] as $b_1 = 4.56 \times 10^{-4}$, $b_2 = 4.16 \times 10^{-4}$, $a_{p_1} = 3.43$, $a_{p_2} = 2.15$, $a_{d_1} = 18.15$, $a_{d_2} = 7.78$ and $a_{d_3} = 18.98$.

In the experiment, we aim to estimate the water levels of the three tanks using two measurements; that is, the water levels in the left and the right tank. The plant is decomposed

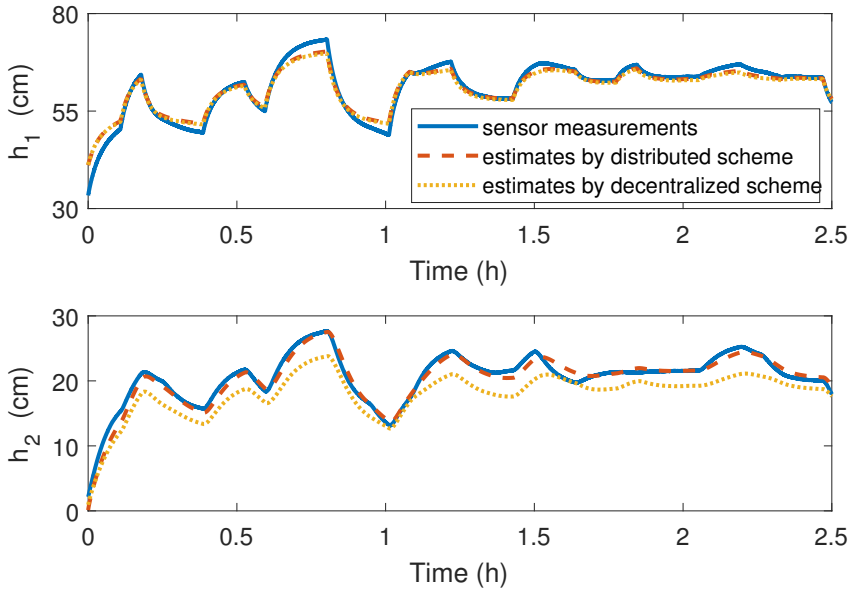

Figure 3.12: Trajectories of the state estimates based on distributed scheme (dashed red lines), the state estimates based on decentralized scheme (dotted yellow lines) and the actual sensor measurements of the water levels (solid blue lines)

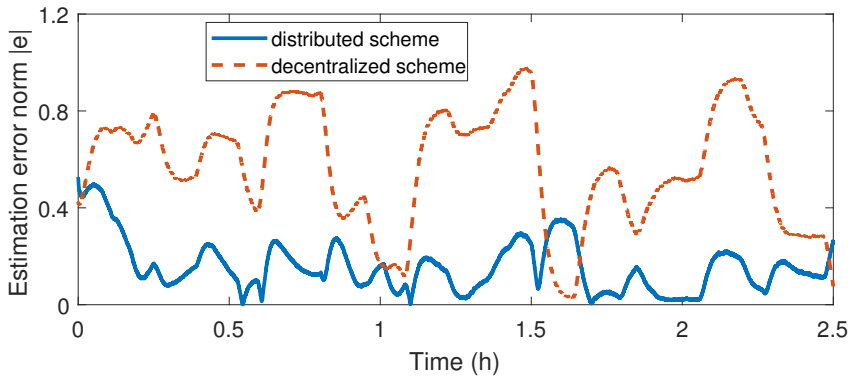

Figure 3.13: Trajectories of the norm of estimation errors given by the distributed scheme (solid blue lines) and the decentralized scheme (dashed red lines)

into two subsystems: the first subsystem consists of the left tank and the middle tank, while the right tank constitutes the second subsystem. Each subsystem has one measurement. For the first subsystem, a decentralized estimator is deigned following [55] to account for the local dynamics. Then, the proposed approach is utilized to design an augmented estimator for the first subsystem. The second subsystem sends its actual measurement to the augmented estimator for the first subsystem every 1 sec. Using the measurements from the left and the right tank, the distributed scheme is evaluated to estimate the water levels of the

left and middle tanks. The actual water levels and the estimates are shown in Figure 3.12. The results confirm that the proposed approach is applicable to this real system. The existing decentralized scheme is also implemented for performance comparison. Note that the flow rates of the water entering the process are time-varying and the process is subject to non-steady-state operation. In the decentralized scheme, constant values at one equilibrium point ($h_{1s} = 46.160$ cm, $h_{2s} = 14.891$ cm, $h_{1s} = 40.335$ cm) are used to compensate for the time-varying subsystem interaction. The state estimates given by the decentralized scheme are also given in Fig. 3.12. The trajectories of the norm of the estimation errors given by the distributed scheme and the decentralized scheme are given in Fig. 3.13. From the figures, it can be seen that the formed distributed scheme gives significantly improved state estimates (especially for h_2 which is not measured) compared to the decentralized scheme.

3.4 Summary

In the present work, we proposed a systematic method to design distributed state estimation networks by taking advantage of existing decentralized state estimation systems. Decentralized local estimators were assumed to exist for the subsystems. A compensator that compensates for the dynamics of subsystem interaction was designed and connected to the corresponding existing decentralized estimator to form an augmented estimator for the corresponding subsystem. A distributed state estimation network was developed by connecting the augmented estimators via discrete-time information exchange with each other. Sufficient conditions were given for the convergence and boundedness of the estimation error. Three application examples were used to demonstrate the proposed approach.

Appendix I

Description of the froth flotation process: The slurry consisting of coal, ash and water is fed into the first vessel. Reagents including frother and collector are added through either

the inlet or the upstream of the first vessel with the optimal reagent concentration which depends on the solid concentration as well as the type of the coal. The froth is removed from the five vessels using the paddles and is further sent to a disk filter system for drying. The underflow of each vessel enters the downstream neighboring vessel, and the tailings of the last vessel are discharged and sent to a static thickener [84]. The dynamics of the process is described by 15 system states with respect to the solids, liquids and ash in the vessels. Based on standard modeling assumptions and mass balances, a model characterizing the dynamics of the i-th unit, $\forall i \in \mathbb{I} = \{1,\ 2,\ 3,\ 4,\ 5\}$ is established as follows [84]:

$$\frac{dc_{st_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{st_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{st_i}) - r_i \tag{3.28a}$$

$$\frac{dc_{lt_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{lt_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{lt_i}) - \frac{\beta_i}{V_i} \tag{3.28b}$$

$$\frac{dc_{at_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{at_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{at_i}) - \frac{\dot{A}_i}{V_i} \tag{3.28c}$$

where c_{st_i} represents the concentration of the solids in the i-th vessel, c_{lt_i} represents the concentration of the liquid in the i-th vessel, c_{at_i} is the concentration of the ash in the i-th vessel, V_i denotes the slurry volume in the i-th vessel, while $\dot{V}_{uf_i}$ denotes the flow rate of the underflow from the i-th vessel. r_i which is specified later represents the solid removal rate from the i-th vessel, β_i denotes the mass flow rate of the liquid from the i-th vessel to overflow and $\dot{A}_i$ denotes the mass flow rate of ash from tank i to overflow. c_{st_0}, c_{lt_0} and c_{at_0} represent the concentration of the solids, the concentration of the liquid and the concentration of the ash in the feed flow, respectively. $\dot{V}_{uf_0}$ depicts the volumetric flow rate of the feed flow.

The solid removal rate r_i is calculated following $r_i = f_r k(c_{st_i} - c_\infty)$ where k denotes a constant rate, c_∞ denotes the equilibrium of the solid concentration (kg/m^3), while f_r represents the correction factor of industrial scale reactions.

Further, for the i-th vessel, $i \in \mathbb{I}$, we have:

$$\begin{aligned}
\dot{V}_{uf_i} &= \dot{V}_{uf_{i-1}} + \dot{V}_{of_i} \\
\dot{V}_{of_i} &= \frac{\beta}{\rho_l} + \frac{r_i V_i}{\rho_c} \\
\dot{A}_i &= x_{A_i} \left(\sum_{j=1}^{i} M_{sof_j} \right) - \sum_{j=1}^{i-1} \dot{A}_j \\
R_i &= \left(\frac{M_{sof_i} \dot{V}_{of_i}}{M_{sof_i} \dot{V}_{of_i} + c_{st_i} \dot{V}_{uf_i}} \right) (100 - R_{i-1}) + R_{i-1} \\
x_{A_i} &= g(R_i) \\
\dot{M}_{sof_i} &= V_i \cdot r_i
\end{aligned}$$

where ρ_c is the density of coal (kg/m^3), ρ_l denotes the density of liquid (kg/m^3), R_i is the cumulative solid recovery at stage i, x_{A_i} is the cumulative mass fraction of ash in the overflow solids at stage i, $\dot{M}_{sof_i}$ is the mass flow rate of solids in the overflow (kg/min), $\dot{V}_{of_i}$ is the volumetric flow rate of the overflow and c_{sof_i} denotes the concentration of solids in the overflow (kg/m^3), V_{feed} denotes the feed flow rate of the liquid and *sol* is a coefficient of the feed mass basis. $g(R_i)$ is an empirical function of R_i for a given frother and collector loading from [84]. A complete description of the model can be found in [84, 61].

This model is decomposed into five subsystems. $x_i = [x_{i,1}\ x_{i,2}\ x_{i,3}]^{\mathrm{T}} = [c_{st_i}\ c_{lt_i}\ c_{at_i}]^{\mathrm{T}}$ is the state vector of subsystem i, $y_i = [y_{i,1}\ y_{i,2}]^{\mathrm{T}} = [c_{lt_i}\ c_{at_i}]^{\mathrm{T}} + [v_{i_1}\ v_{i_2}]^{\mathrm{T}}$ is the vector of output measurements of subsystem i, $w_i = [w_{i,1}\ w_{i,2}\ w_{i,3}]^{\mathrm{T}}$ is the vector of additive disturbances to subsystem i, $i = 1,\ 2,\ 3,\ 4,\ 5$. The vectors of interacting dynamics for the subsystems are: $X_2 = x_1$, $X_3 = \left[x_1^{\mathrm{T}}\ x_2^{\mathrm{T}}\right]^{\mathrm{T}}$, $X_4 = \left[x_1^{\mathrm{T}}\ x_2^{\mathrm{T}}\ x_3^{\mathrm{T}}\right]^{\mathrm{T}}$ and $X_5 = \left[x_1^{\mathrm{T}}\ x_2^{\mathrm{T}}\ x_3^{\mathrm{T}}\ x_4^{\mathrm{T}}\right]^{\mathrm{T}}$.

The expressions of the functions f_i, $\tilde{f}_i$ and h_i are reported for $i = 1,\ 2$. In the first

subsystem, the vector field f_1 is with the following form:

$$f_1(x_1, w_1) = \begin{bmatrix} f_{1,1}(x_1, w_1) \\ f_{1,2}(x_1, w_1) \\ f_{1,3}(x_1, w_1) \end{bmatrix}$$

where

$$\begin{aligned} f_{1,1} =& -\left(\frac{V_{feed}}{V_1} - \frac{\beta}{\rho_l V_1} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_i} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c}\right) x_{1,1} \\ &+ f_r \cdot k \cdot (c_\infty - x_{1,1}) + \frac{V_{feed} \cdot \rho_l \cdot sol}{V_1 (1 - sol)} + w_{1,1} \\ f_{1,2} =& -\left(\frac{V_{feed}}{V_1} - \frac{\beta}{\rho_l V_1} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_1} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c}\right) x_{1,2} \\ &+ \frac{V_{feed} \cdot \rho_l}{V_1} - \frac{\beta}{V_1} + w_{1,2} \\ f_{1,3} =& -\left(\frac{V l_{feed}}{V_{tank}} - \frac{\beta}{\rho_l V_1} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_1} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c}\right) x_{1,3} \\ &+ X_{A_1} \cdot f_r k (c_\infty - x_{11}) - \frac{0.245 \cdot V_{feed} \cdot \rho_l \cdot sol}{V_1 \cdot (sol - 1)} + w_{1,3} \end{aligned}$$

Because the dynamics of the first subsystem is not affected by the other subsystems, $\tilde{f}_1$ is a zero vector field. The output measurement equations are:

$$h_1(x_1) = \begin{bmatrix} h_{1,1}(x_1) \\ h_{1,2}(x_1) \end{bmatrix} = \begin{bmatrix} c_{lt_1} \\ c_{at_1} \end{bmatrix}$$

In the second subsystem, the vector field f_2 is with the following form:

$$f_2(x_2, w_2) = \begin{bmatrix} f_{2,1}(x_2, w_2) \\ f_{2,2}(x_2, w_2) \\ f_{2,3}(x_2, w_2) \end{bmatrix}$$

where

$$
\begin{aligned}
f_{2,1} = &\left(\frac{Vl_{feed}}{V_2} - \frac{2\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_2} + \frac{f_r k (c_\infty - x_{21})}{\rho_c} \right) x_{2,1} \\
&+ f_r k (c_\infty - x_{21}) + w_{2,1} \\
f_{2,2} = &- \left(\frac{Vl_{feed}}{V_2} - \frac{2\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_2} + \frac{f_r k (c_\infty - x_{2,1})}{\rho_c} \right) x_{2,2} \\
&- \frac{\beta}{V_2} + w_{2,2} \\
f_{2,3} = &- \left(\frac{V_{feed}}{V_2} - \frac{2\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_2} + \frac{f_r k (c_\infty - x_{2,1})}{\rho_c} \right) x_{2,3} \\
&+ X_{A_2} \cdot V_2 \cdot f_r k (c_\infty - x_{21}) + w_{2,3}
\end{aligned}
$$

The interacting dynamics $\tilde{f}_2$ is given as:

$$
\tilde{f}_2(x_2, X_2) = \begin{bmatrix} \tilde{f}_{2,1}(x_2, X_2) \\ \tilde{f}_{2,2}(x_2, X_2) \\ \tilde{f}_{2,3}(x_2, X_2) \end{bmatrix}
$$

where

$$
\begin{aligned}
\tilde{f}_{2,1} = &\left(\frac{V_{feed}}{V_2} - \frac{\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c \cdot (sol-1) V_2} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{1,1} \\
&- \left(\frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{2,1} \\
\tilde{f}_{2,2} = &\left(\frac{V_{feed}}{V_2} - \frac{\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c \cdot (sol-1) V_2} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{1,2} \\
&- \left(\frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{2,2} \\
\tilde{f}_{2,3} = &\left(\frac{V_{feed}}{V_2} - \frac{\beta}{\rho_l V_2} - \frac{V_{feed} \cdot \rho_l \cdot sol}{\rho_c (sol-1) V_2} + \frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{1,3} \\
&+ X_{A_2} \cdot f_r k (c_\infty - x_{1,1}) - X_{A_1} \cdot f_r k (c_\infty - x_{1,1}) \\
&- \left(\frac{f_r k (c_\infty - x_{1,1})}{\rho_c} \right) x_{2,3}
\end{aligned}
$$

The output measurement equations are:

$$h_2(x_2) = \begin{bmatrix} h_{2,1}(x_2) \\ \\ h_{2,2}(x_2) \end{bmatrix} = \begin{bmatrix} c_{lt_2} \\ \\ c_{at_2} \end{bmatrix}$$

The explicit expressions of f_i, $\tilde{f}_i$ and h_i for each subsystem i, $i = 3,\ 4,\ 5$, are obtained using symbolic calculation in the same way, and are omitted for brevity. The Lipschitz constants with respect to the nonlinear functions in the model are obtained as $L_{\tilde{f}_1}^{X_1} = 0$ (as the dynamics of the first subsystem is not affected by the remaining subsystems), $L_{\tilde{f}_2}^{X_2} = 1.0725$, $L_{\tilde{f}_3}^{X_3} = 1.6895$, $L_{\tilde{f}_4}^{X_4} = 1.6679$, $L_{\tilde{f}_5}^{X_5} = 1.6476$.

Chapter 4

Distributed output-feedback fault detection and isolation of cascade process networks

In this chapter, we consider distributed output-feedback fault detection and isolation (FDI) of nonlinear cascade process networks that can be divided into subsystems. Based on the assumption that an exponentially convergent estimator exists for each subsystem, a distributed state estimation system is developed based on the method that is introduced in Chapter 3. It is shown in Section 4.2.5 that when there is no fault, the estimation error of the distributed estimation system converges to zero in the absence of system disturbances and measurement noise. For each subsystem, a state predictor is also designed to provide subsystem state predictions. A residual generator is designed for each subsystem based on subsystem state estimates given by the distributed state estimation system and subsystem state predictions given by the predictor. A subsystem residual generator generates two residual sequences, which act as references for FDI. A distributed FDI mechanism is proposed based on residuals as illustrated in Section 4.3. The proposed approach is able to handle both actuator faults and sensor faults by evaluating the residual signals. In Section 4.4, a

froth flotation process example is used to illustrate the effectiveness of the distributed FDI mechanism. This chapter is a revised version of [85, 97].

4.1 Preliminaries

4.1.1 Notation

The operator $|\cdot|$ denotes the Euclidean norm of a vector. A function $f(s)$ is said to satisfy the locally Lipschitz property with respect to its argument s, if there exists a positive constant L_f^s so that $|f(s_1) - f(s_2)| \leq L_f^s |s_1 - s_2|$ holds for all s_1 and s_2 that are in a local region of s and L_f^s represents the corresponding Lipschitz constant. A continuous function $\alpha : [0, a) \to [0, \infty)$ belongs to class $\mathscr{K}$ if it is strictly increasing and satisfies $\alpha(0) = 0$. A continuous function $\alpha : [0, a) \to [0, \infty)$ belongs to class $\mathscr{K}_\infty$ if it belongs to class $\mathscr{K}$ and it is subject to $a = \infty$ and $\lim_{r\to\infty} \alpha(r) = \infty$. A continuous function $\beta(r, s)$ belongs to class $\mathscr{KL}$ if for each fixed s, $\beta(r, s)$ belongs to class $\mathscr{K}$ with respect to r, and for each fixed r, it is deceasing with respect to s, and $\beta(r, s) \to 0$ as $s \to \infty$. A^+ denotes a matrix (or a vector) being the Moore-Penrose pseudoinverse of matrix (or vector) A. The symbol $\mathbb{I}$ is a set containing finite integers defined as $\mathbb{I} := \{1, \ldots, n\}$. $\mathbb{J}$ denotes a subset of $\mathbb{I}$ and is defined as $\mathbb{J} := \{2, \ldots, n\}$.

4.1.2 System description and problem formulation

We consider a class of nonlinear systems comprised of n subsystems interconnected in a cascade fashion. The diagram of a general cascade process is presented in Figure 4.1. The entire system under a fault-free condition is described in the following form:

$$\begin{aligned} \dot{x}(t) &= \mathcal{F}\left(x(t), u(t)\right) \\ y(t) &= h(x(t)) \end{aligned} \tag{4.1}$$

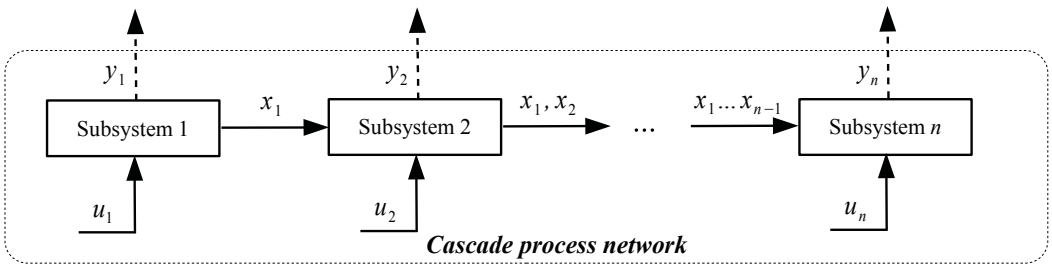

Figure 4.1: A diagram of a cascade process.

where $x(t) \in \mathbb{R}^{n_x}$ represents the state vector, $u(t) \in \mathbb{R}^{n_u}$ represents the input vector to the entire system, and $y \in \mathbb{R}^{n_y}$ denotes the measured output vector of the entire system. The function $\mathcal{F}$ characterizes the dynamics of x.

For the class of cascade nonlinear systems, the dynamics of subsystem i, $i = 1, \ldots, n-1$, are not affected by its downstream subsystems (i.e., each subsystem j for all $j = i+1, \ldots, n$). The models of the n subsystems are represented in the following form:

$$\begin{cases} \dot{x}_1(t) = f_1\left(x_1(t), u_1(t) + \tilde{u}_1(t)\right) \\ \dot{x}_2(t) = f_2\left((x_2(t), u_2(t) + \tilde{u}_2(t)\right) + \tilde{f}_2(x_1(t)) \\ \qquad\qquad\qquad \vdots \\ \dot{x}_n(t) = f_n\left(x_n(t), u_n(t) + \tilde{u}_n(t)\right) + \tilde{f}_n\left(x_1(t), x_2(t), \ldots, x_{n-1}(t)\right) \\ y_i(t) = h_i(x_i(t)) + \tilde{v}_i(t), \quad i \in \mathbb{I} \end{cases} \tag{4.2}$$

where $x_i(t) \in \mathbb{R}^{n_{x_i}}$ denotes the state vector of the i-th subsystem, $u_i \in \mathbb{R}^{n_{u_i}}$ denotes the input vector to the i-th subsystem, $y_i(t) \in \mathbb{R}^{n_{y_i}}$ is the measured output vector of the i-th subsystem, $i \in \mathbb{I}$, f_i denotes a vector function characterizing the dependence of the dynamics of x_i on itself and the disturbances w_i, the vector function $\tilde{f}_j$, $j \in \mathbb{J}$, characterizes the dependence of the dynamics of the j-th subsystem on itself as well as its upstream subsystems, and h_i is a vector function that describes the relation between y_i and x_i. We use $\tilde{u}_i$ to denote an actuator fault in the i-th subsystem, $i \in \mathbb{I}$. $\tilde{v}_i$ denotes a sensor fault in the subsystem i,

$i \in \mathbb{I}$. Without loss of generality, we assume that x_i satisfies the following constraint:

$$x_i \in \mathbb{X}_i \tag{4.3}$$

where $\mathbb{X}_i$, $\forall i \in \mathbb{I}$, denotes a convex compact set.

We note that for each subsystem j, $j \in \mathbb{J}$, in model (4.2), all the upstream subsystem states are considered in the interaction term $\tilde{f}_j(\cdot)$ to make the model more general. However, it is possible that not all the upstream subsystems states are directly involved in $\tilde{f}_j(\cdot)$. Based on this consideration, we introduce X_j, $j \in \mathbb{J}$, to denote a vector containing all the upstream subsystem states that are involved in characterizing the interaction of the j-th subsystem. In the remainder, we will use $\mathbb{L}_j$ to represent the set of the indices of the upstream subsystems whose subsystem states are explicitly involved in X_j. For instance, if x_1 and x_2 are explicitly involved in $\tilde{f}_4$, then we have $\mathbb{L}_4 = \{1, 2\}$, and $X_4(t) = [x_1(t)\ x_2(t)]^T$, $j \in \mathbb{J}$. This set of symbols will be used to describe a more efficient communication strategy in the distributed state estimation design in the next section. In this work, we consider that for the cascade process (4.2), there is only one fault present in the entire process within one prediction horizon, the definition of which will be given later.

4.2 Distributed state estimation design

In this section, we develop a distributed state estimation system for the cascade process network (4.2) following [129]. In the present work, we assume that the subsystem estimators can exchange information continuously which is different from [129], in which discrete-time communication was considered. Note that continuous communication has also been used in other distributed fault detection works (e.g., Keliris et al. [78] and Ding et al. [82]). Based on continuous communication, we will show that the estimation error converges to zero.

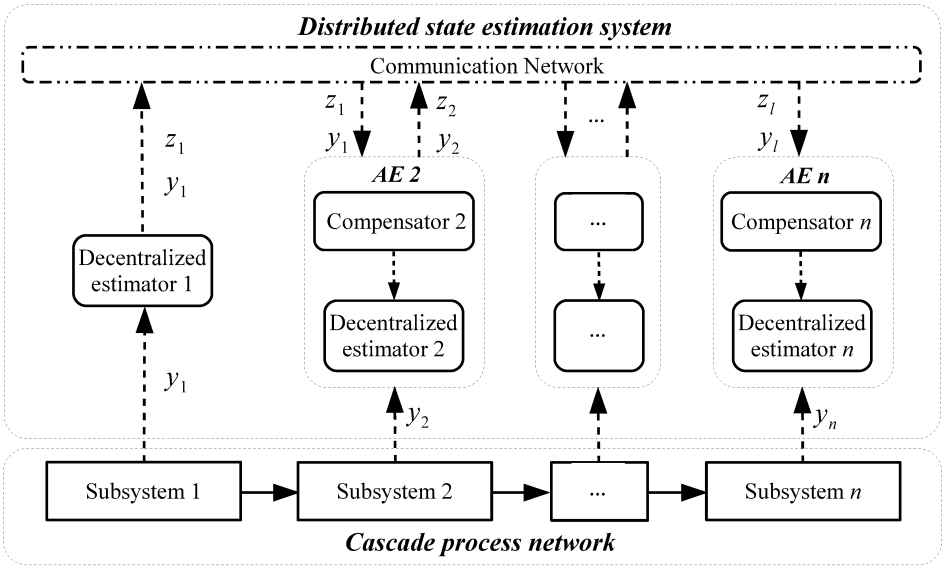

Figure 4.2: A schematic of the distributed state estimation scheme for the cascade process.

4.2.1 Description of the distributed state estimation system

A schematic of the designed distributed state estimation system is shown in Figure 4.2. It is assumed that there exists a decentralized state estimator for each subsystem. A compensator is developed for each subsystem except for the first subsystem to compensate for the dynamics of the subsystem interaction. Each existing decentralized estimator is connected to the associated compensator to form an augmented estimator (AE) for each subsystem except for the first subsystem. The estimators for the subsystems communicate and exchange information (i.e., subsystem state estimates and measurements) with the corresponding interacted subsystems continuously via a communication network. The estimator of subsystem 1 sends out its state estimate and its measurement to its downstream subsystems that are directly affected by subsystem 1. The augmented estimator for each subsystem j, $j \in \mathbb{J}$, receives subsystem state estimates and the measurements from each subsystem l, $\forall l \in \mathbb{L}_j$. In the meantime, it sends out subsystem state estimates z_j and measurement y_j continuously to its downstream subsystems of which the dynamics is directly affected by x_j, $j \in \mathbb{J}$.

4.2.2 Assumed local observers/estimators

We assume that there exists a decentralized nonlinear observer/estimator for subsystem i, $i \in \mathbb{I}$, that is described in a continuous-time framework as in the following form:

$$\dot{z}_i^o(t) = F_i(z_i^o(t), h_i(x_i(t)), u_i(t), X_{is}) \tag{4.4}$$

where z_i^o denotes the state of the decentralized nonlinear observer/estimator for the i-th subsystem in (4.4), X_{is}, $i \in \mathbb{J}$, is the steady-state vector which is corresponding to $X_i(t)$, while X_{1s} does not actually exist. We assume that the decentralized obervers/estimators in (4.4) for the subsystems satisfy the following assumption.

Assumption 4. *For each subsystem i, $i \in \mathbb{I}$, there exists a decentralized estimator in the form of (4.4), such that if the dynamics of x_i satisfy the constraint (4.3), and if the input u_i to i-th subsystem is known, and if the entire system (4.2) is operated in a fault-free condition, and all the states of the directly interacted upstream subsystems of subsystem i are at their steady states (i.e., $X_i(t) \equiv X_{is}$, $i \in \mathbb{J}$, for all t), then there exists a class $\mathscr{K}$ function ξ_i and a positive scalar λ_i, such that the estimation error of the decentralized estimator (4.4) satisfies the following inequality:*

$$|z_i^o(t) - x_i(t)| \leq \xi_i\left(|z_i^o(0) - x_i(0)|\right) e^{-\lambda_i t} \tag{4.5}$$

In (4.5), the class $\mathscr{K}$ function ξ_i describes how significantly the initial condition can affect the estimation error, and $\lambda_i > 0$ characterizes the convergence speed of the existing observer/estimator of the i-th subsystem, $i \in \mathbb{I}$. Note that for each subsystem j, $j \in \mathbb{J}$, the above assumption is satisfied only when all the states of the upstream subsystems of subsystem j that directly affect the dynamics of x_j are at steady-state values; that is, $X_j(t) \equiv X_{js}$, $j \in \mathbb{J}$, for all t.

Remark 15. *Assumption 4 implies that there exists an exponential convergent observer for*

each subsystem if all the other subsystems are kept at their steady states. This assumption can be satisfied by many existing nonlinear observer designs including the ones proposed by Gauthier et al. [55], Ciccarella et al.[30], Boizot [56], Ahmed-Ali et al. [86] as well as Atassi and Khalil [87].

4.2.3 Compensator and augmented estimator design

The dynamics of the subsystem interaction are not appropriately handled by the existing decentralized estimators. To account for the dynamics of the interaction, we design a compensator for each subsystem except for the first subsystem following [129]. By connecting each decentralized estimator to its corresponding compensator, we construct an augmented estimator for each subsystem j, $j \in \mathbb{J}$. The j-th compensator for subsystem j, $j \in \mathbb{J}$, is designed as below.

$$\dot{z}_j^c(t) = \tilde{f}_j(Z_j(t)) - \tilde{f}_j(X_{js}) + \sum_{l \in \mathbb{L}_j} K_{j,l}(z_l)(y_l(t) - h_l(z_l(t))) \tag{4.6}$$

where $z_j^c(t)$ is a compensation signal given by the j-th compensator. In the compensator (4.6) for subsystem j, $j \in \mathbb{J}$, the first term and the second term on the right-hand-side of (4.6) compensate for the deviation between the actual subsystem interaction of subsystem j and the corresponding steady-state values. The third term on the right-hand-side of (4.6) is a corrective term to further mitigate the deviation due to the use of state estimates (instead of the corresponding actual states) in the subsystem interaction. $K_{j,l}$ are the correction gains that are calculated as follows:

$$K_{j,l} = \frac{\partial \tilde{f}_j}{\partial x_l} \left(\frac{\partial h_l}{\partial x_l} \right)^+ \Bigg|_{x_l = z_l(t)} \tag{4.7}$$

for all $j \in \mathbb{J}$, $l \in \mathbb{L}_j$. We consider that the correction gains are calculated and updated continuously.

By combining the j-th compensator with the j-th decentralized estimator, an augmented estimator for the subsystem j is formed as in the following form:

$$\dot{z}_j(t) = F_j\left(z_j(t), y_j(t), u_j(t), X_{js}\right) + \tilde{f}_j(Z_j(t)) - \tilde{f}_j(X_{js}) + \sum_{l \in \mathbb{L}_j} K_{j,l}(z_l)(y_l(t) - h_l(z_l(t))) \tag{4.8}$$

where z_j represents the subsystem state estimate generated by the augmented estimator for the subsystem j, $j \in \mathbb{J}$.

Remark 16. *We note that for subsystem j, $j \in \mathbb{J}$, the developed augmented estimator only requires the state estimates and output measurements from the upstream subsystems that directly affect the dynamics of subsystem j (i.e., the estimates and the measurements of subsystem l for all $l \in \mathbb{L}_j$). The j-th augmented estimator only needs to request and receive information from each subsystem l, $l \in \mathbb{L}_j$, and the corresponding estimator via the communication network. This communication strategy saves limited communication resources and is favorable from an application point of view.*

4.2.4 Integral input-to-state stability of the augmented estimators

In this section, we carry out stability analysis individually on the subsystem estimation error dynamics of each developed augmented estimator j, $j \in \mathbb{J}$, in (4.8). First, we introduce the concept of integral input-to-state stability (iISS). Then, we state in Theorem 4 that the augmented estimator j developed for subsystem j, $j \in \mathbb{J}$, is stable in the sense of iISS when there is no fault in the system.

Definition 1. *(c.f. Sontag[88]) If there exist functions α and γ belonging to class $\mathscr{K}_\infty$, and a function β belonging to class $\mathscr{KL}$, such that for any initial condition $x(0) \in \mathbb{R}^{n_x}$ and any input $u(\cdot) \in \mathbb{R}^{n_u}$, the following estimate*

$$\alpha\left(|x(t)|\right) \leq \beta\left(|x(0)|, t\right) + \int_0^t \gamma\left(|u(s)|\right) ds \tag{4.9}$$

holds for all t, $t \geq 0$, then the system which satisfies the property in (4.9) is said to be an integral input-to-state stable (iISS) system.

Theorem 4. *Consider the augmented nonlinear estimator j, $j \in \mathbb{J}$, in the form of (4.8) with initial condition $\hat{x}_i(0)$. If there is no fault in the system, and if the output measurements are continuously available, and if each of the correction gains $K_{j,l}$ for $j \in \mathbb{J}$ and $l \in \mathbb{L}_j$ is selected following (4.7) and is bounded such that $|K_{j,l}| \leq \mathcal{K}_{j,l}$ with $\mathcal{K}_{j,l}$ being a positive constant, then the dynamics of the subsystem estimation error $e_i := z_i - x_i$ are iISS for all $x_i \in \mathbb{X}_i$, for all $i \in \mathbb{I}$ and $t \geq 0$.*

Proof: Let us consider the j-th augmented observer j, $j \in \mathbb{J}$, in the form of (4.8). In a fault-free context, the derivative of e_j is calculated as follows:

$$\begin{aligned}\dot{e}_j(t) =& F_j\left(z_j(t), y_j(t), u_j(t), X_{js}\right) - f_j\left(x_j(t), u_j(t)\right) - \tilde{f}_j(X_{js}) + \tilde{f}_j(Z_j(t)) - \tilde{f}_j\left(X_j(t)\right) \\ &+ \sum_{l \in \mathbb{L}_j} K_{j,l}(z_l)\left(h_l(x_l(t)) - h_l(z_l(t))\right).\end{aligned} \tag{4.10}$$

Performing Taylor series expansions on the term $\tilde{f}_j\left(X_j(t)\right)$ and the term $h_l\left(x_l(t)\right)$, respectively, the following equalities can be obtained:

$$\begin{aligned}\tilde{f}_j(X_j(t)) =& \tilde{f}_j(Z_j(t)) + \sum_{l \in \mathbb{L}_j} \left.\frac{\partial \tilde{f}_j}{\partial x_l}\right|_{x_l = z_l(t)} (x_l(t) - z_l(t)) + \sum_{l \in \mathbb{L}_j} o_{\tilde{f}_j}(z_l^2) \\ h_l(x_l(t)) =& h_l(z_l(t)) + \left.\frac{\partial h_l}{\partial x_l}\right|_{x_l = z_l(t)} (x_l(t) - z_l(t)) + o_{h_l}(z_l^2)\end{aligned} \tag{4.11}$$

where $o_{\tilde{f}_j}(z_l^2)$ and $o_{h_l}(z_l^2)$ represent the high-order terms of Taylor expansions of $\tilde{f}_j$ and h_l, respectively, and are subject to the following constraints:

$$o_{\tilde{f}_j}(z_l^2) \leq H_j^{\tilde{f}_j} |x_l(t) - z_l(t)|^2, \quad o_{h_l}(z_l^2) \leq H_l^{h_l} |x_l(t) - z_l(t)|^2 \tag{4.12}$$

for all $x_l \in \mathbb{X}_l$, $l \in \mathbb{L}_j$. In (4.12), $H_j^{\tilde{f}_j}$ and $H_l^{h_l}$ are finite positive constants that are associated with the Taylor expansions of function $\tilde{f}_j$, $j \in \mathbb{J}$, and function h_l, $l \in \mathbb{L}_j$, respectively.

Taking into account (4.10) and (4.11), it can be derived that:

$$\begin{aligned}\dot{e}_j(t) =& F_j\left(z_j(t), h_j(x_j(t)), u_j(t), X_{js}\right) - f_j\left(x_j(t), u_j(t)\right) - \tilde{f}_j(X_{js}) \\ &+ \sum_{l\in\mathbb{L}_j}\Big(- \left.\frac{\partial \tilde{f}_j}{\partial x_l}\right|_{x_l=z_l(t)} (x_l(t) - z_l(t)) + K_{j,l} \left.\frac{\partial h_l}{\partial x_l}\right|_{x_l=z_l(t)} (x_l(t) - z_l(t))\Big) \\ &+ \sum_{l\in\mathbb{L}_j} K_{j,l} o_{h_l}(z_l^2) - \sum_{l\in\mathbb{L}_j} o_{\tilde{f}_j}(z_l^2)\end{aligned} \tag{4.13}$$

If the correction gains $K_{j,l}$ are evaluated following (4.7), and take into account (4.12) and (4.13), we can further calculate that:

$$\begin{aligned}|\dot{e}_j(t)| \;\leq\; & |F_j(z_j(t), h_j(x_j(t)), u_j(t), X_{js}) - f_j\left(x_j(t), u_j(t)\right) - \tilde{f}_j(X_{js})| \\ & + \sum_{l\in\mathbb{L}_j}\left(H_j^{\tilde{f}_j} + K_{j,l} H_l^{h_l}\right)|e_l(t)|^2\end{aligned} \tag{4.14}$$

Taking into account Assumption 4 on the convergence property of the existing observer/estimator for subsystem j, $j \in \mathbb{J}$, if we integrate both sides of inequality (4.14) from initial time 0 to time t, the following inequality is derived:

$$\begin{aligned}|e_j(t)| \leq & \xi_j\left(|e_j(0)|\right) e^{-\lambda_j t} + \int_0^t \left(\sum_{l\in\mathbb{L}_j}\left(H_j^{\tilde{f}_j} + \mathcal{K}_{j,l} H_l^{h_l}\right)|e_l(s)|^2\right) ds \\ \leq & \xi_j\left(|e_j(0)|\right) e^{-\lambda_j t} + \int_0^t \left(\max\left\{H_j^{\tilde{f}_j} + \mathcal{K}_{j,l} H_l^{h_l} : l \in \mathbb{L}_j\right\} \sum_{l\in\mathbb{L}_j} |e_l(s)|^2\right) ds\end{aligned} \tag{4.15}$$

Let us use ϵ_j to denote the column vector containing all e_l, $l \in \mathbb{L}_j$. Then, we have

$$|\epsilon_j(s)|^2 = \sum_{l\in\mathbb{L}_j} |e_l(s)|^2 \tag{4.16}$$

Based on (4.15) and (4.16), it is derived that

$$|e_j(t)| \leq \xi_j \left(|e_j(0)|\right) e^{-\lambda_j t} + \int_0^t \left(\max \left\{ H_j^{\tilde{f}_j} + \mathcal{K}_{j,l} H_l^{h_l} : l \in \mathbb{L}_j \right\} |\epsilon_j(s)|^2 \right) ds \tag{4.17}$$

Now, let us revisit the iISS condition (4.9) in Definition 1. For each $j \in \mathbb{J}$, if the class $\mathscr{K}_\infty$ function α in (4.9) is defined as $\alpha(s) := s$, the class $\mathscr{KL}$ function $\beta(s,t)$ in (4.9) is defined as $\beta(s,t) := \gamma_j(s)\, e^{-\lambda_j t}$, while the class $\mathscr{K}_\infty$ function in (4.9) γ is defined as $\gamma(s) := cs^2$ with c being a positive scalar defined as $c := \max \left\{ H_j^{\tilde{f}_j} + \mathcal{K}_{j,l} H_l^{h_l} : l \in \mathbb{L}_j \right\}$. Then, it is seen that (4.17) is in the form of (4.9) for all j, $j \in \mathbb{J}$. This completes the proof of Theorem 4. □

In what follows, we will carry out analysis of the convergence of the estimation error of the distributed state estimation system.

Remark 17. *It is worth mentioning that in Theorem 4, the estimation errors of all the upstream subsystems that are directly involved in the dynamics of the j-th subsystem are essentially regarded as external inputs to the estimation error dynamics of the j-th augmented estimator, $j \in \mathbb{J}$. The column vector ϵ_j, $j \in \mathbb{J}$, is used to represent the external input vector to the estimation error system of the j-th augmented estimator. For instance, if one has $\mathbb{L}_5 = \{1,3,4\}$, then ϵ_5 is in the form of $\epsilon_5 = \left[e_1^T, e_3^T, e_4^T\right]^T$.*

4.2.5 Boundedness and Convergence of the estimation error

In this section, we focus on the convergence properties of the distributed state estimation scheme developed for the cascade process (4.2) in the fault-free condition. Before proceeding further, the following definition is given to characterize a specific type of class $\mathscr{KL}$ functions.

Definition 2. *A function $\bar{\beta}(r,s)$ is said to belong to class $\mathscr{I} - \mathscr{KL}$ if $\bar{\beta} \in \mathscr{KL}$ and if $\int_{s=0}^{\infty} \bar{\beta}(r,s)ds$ takes a finite value for all r, $0 \leq r < \infty$.*

Further, two lemmas that will be used for the convergence analysis are introduced.

Lemma 1. *Consider a function ψ which is defined as $\psi(r,s) := \theta(r)e^{-\sigma s}$ where $\theta(r)$ belongs to class $\mathscr{K}$ and takes a finite value. If σ is a finite positive scalar (i.e., $0 < \sigma < \infty$), then the function $\psi(r,s)$ is said to belong to class $\mathscr{I} - \mathscr{K}\mathscr{L}$ for all $r \geq 0$.*

Proof: For any function ψ described as in Lemma 1, given a fixed s, the function ψ belongs to class $\mathscr{K}$ as $\theta(r)$ is a class $\mathscr{K}$ function. Also, for each fixed r, $\psi(r,s)$ is a decreasing trajectory with respect to its second argument s and satisfies $\psi(r,\infty) = 0$. Therefore, $\psi(r,s)$ belongs to class $\mathscr{K}\mathscr{L}$.

Next, let us consider the integral of function $\psi(r,s)$ with respect to its second argument. That is,

$$\int_{s=0}^{\infty} \psi(r,s)ds = \int_{s=0}^{\infty} \theta(r)e^{-\alpha s}ds = \frac{\theta(r)}{\alpha} < \infty \tag{4.18}$$

Based on Definition 2, it is proved that $\psi(r,s)$ belongs to class $\mathscr{I} - \mathscr{K}\mathscr{L}$. This completes the proof. □

Lemma 2. *(c.f. Sontag[88]) If the dynamics of a system are iISS in the sense of Definition 1, and if for any external input $u \in \mathbb{R}^u$ such that $\int_0^\infty \gamma(|u(s)|)\,ds < \infty$ is satisfied for the class $\mathscr{K}_\infty$ function γ as defined in the condition (4.9), then the corresponding state trajectory $x(t)$ eventually approaches zero as time goes to infinity. That is, $x(t) \to 0$ as $t \to \infty$.*

In Theorem 5 below, we provide sufficient conditions that ensure the boundedness and convergence to zero of the estimation error of the distributed state estimation scheme for system (4.2) in the absence of faults.

Theorem 5. *Consider the cascade nonlinear system (2.1) with continuously available output measurements y_i and known inputs u_i for all $i \in \mathbb{I}$. For subsystem i, $i \in \mathbb{I}$, if constraint (4.3) is satisfied, and if there is no fault in the system, and if a decentralized estimator described as in (4.4) exists such that Assumption 4 holds, and if an augmented estimator is designed for each subsystem j, $j \in \mathbb{J}$, following (4.8) with correction gains $K_{j,l}$ determined following (4.7) and restricted by $|K_{j,l}| \leq \mathcal{K}_{j,l}$ for all $j \in \mathbb{J}$, $l \in \mathbb{L}_j$, then given an arbitrary scalar $\delta_i > 0$, there exists a finite positive scalar τ_i, such that $|e_i(t)| = |z_i(t) - x_i(t)| \leq \delta_i$ holds $\forall t \geq \tau_i$,*

$i \in \mathbb{I}$. *Moreover, the estimation error of the distributed state estimation system for the entire system (4.2) converges to zero as* $t \to \infty$.

Proof: In this proof, we investigate the estimation error of each subsystem individually in a recursive fashion when there is no fault. Let us use e_i to denote the estimation error for subsystem i, $i \in \mathbb{I}$; that is, $e_i(t) := z_i(t) - x_i(t)$.

First, we consider the estimation error e_1 of subsystem 1. In the absence of faults, it is derived that:

$$\dot{e}_1(t) = F_1\left(z_1(t), y_1(t), u_1(t)\right) - f_1\left(x_1(t), u_1(t)\right) \tag{4.19}$$

Subsystem 1 is not affected by other subsystems and no compensator is designed for the first subsystem. Based on Assumption 4 on the convergence of the existing observer in the form of (4.4), the following inequality can be directly obtained:

$$|e_1(t)| \leq \xi_1\left(|e_1(0)|\right) e^{-\lambda_1 t} \tag{4.20}$$

Inequality (4.20) implies that without the occurrence of any faults to system (4.2), the estimation error for subsystem 1 exponentially converges to zero as $t \to \infty$.

Next, let us consider the estimation error of subsystem 2. If we have $\tilde{f}_2 \neq 0$, e_1 can be considered as an external input to the error dynamics of subsystem 2. Based on (4.17), we can obtain the following inequality:

$$|e_2(t)| \leq \xi_2\left(|e_2(0)|\right) e^{-\lambda_2 t} + \int_0^t \left(\left(H_2^{\tilde{f}_2} + \mathcal{K}_{2,1} H_1^{h_1}\right) |\epsilon_2(s)|^2\right) ds \tag{4.21}$$

where $\epsilon_2 = e_1$. Consider the second term on the right-hand-side (RHS) of (4.21) and (4.20),

it is derived that

$$\begin{aligned}\int_0^t \left(\left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)|\epsilon_2(s)|^2\right)ds &= \left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)\int_0^t \left(|e_1(s)|^2\right)ds \\ &\leq \left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)\left(\xi_1\left(|e_1(0)|\right)\right)^2\int_0^t e^{-2\lambda_1 s}ds \\ &= \frac{1}{2\lambda_1}\left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)\left(\xi_1\left(|e_1(0)|\right)\right)^2\left(-e^{-2\lambda_1 t}+1\right)\end{aligned} \tag{4.22}$$

In (4.22), the function ξ_i belongs to class $\mathscr{K}$. Therefore, $\xi_i\left(|e(0)|\right)$ is finite given each $|e(0)|$ with a finite value. Moreover, λ_1 is a positive constant so that the term $e^{-2\lambda_1 t}$ converges to zero as $t\to\infty$. Therefore, from (4.22) we can obtain that

$$\int_0^\infty \left(\left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)|\epsilon_2(s)|^2\right)ds \leq \frac{1}{2\lambda_1}\left(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\right)\left(\xi_1\left(|e_1(0)|\right)\right)^2 < \infty \tag{4.23}$$

Based on (4.23) and the fact that (4.21) is an iISS system, $|e_2(t)|$ approaches to zero as $t\to\infty$ according to Lemma 2.

In what follows, let us focus on the estimation error of subsystem 3. From a state estimation perspective, the most complex case scenario for subsystem 3 is that its dynamics is directly affected by the dynamics of both subsystem 1 and subsystem 2. Based on Theorem 4, estimation error dynamics $e_3(t)$ are iISS. Considering the property in (4.17) that characterizes the error dynamics of each subsystem, the estimation error norm of subsystem 3 is bounded as follows:

$$|e_3(t)| \leq \xi_3\left(|e_3(0)|\right)e^{-\lambda_3 t} + \int_0^t \left(\max\left\{H_3^{\tilde{f}_3}+\mathcal{K}_{3,1}H_1^{h_1}, H_3^{\tilde{f}_3}+\mathcal{K}_{3,2}H_2^{h_2}\right\}|\epsilon_3(s)|^2\right)ds \tag{4.24}$$

where $\epsilon_3 := \left[e_1^T, e_2^T\right]^T$. For brevity, we introduce a finite scalar $c_3 > 0$ which is defined as $c_3 := \max\left\{H_3^{\tilde{f}_3}+\mathcal{K}_{3,1}H_1^{h_1}, H_3^{\tilde{f}_3}+\mathcal{K}_{3,2}H_2^{h_2}\right\}$. Then, (4.24) can be rewritten as:

$$|e_3(t)| \leq \xi_3\left(|e_3(0)|\right)e^{-\lambda_3 t} + \int_0^t c_3\left(|e_1(s)|^2+|e_2(s)|^2\right)ds \tag{4.25}$$

According to Lemma 2, if the second term on the RHS of (4.25) takes a finite value, then $|e_3(t)|\to 0$ as $t\to\infty$. Taking into account (4.22), (4.23) and c_3 with a finite value, we have that $\int_0^\infty c_3\left(|e_1(s)|^2\right)ds<\infty$. Next, we prove that $c_3\int_0^\infty\left(|e_2(s)|^2\right)ds<\infty$ also holds. By substituting (4.22) into the second term on the RHS of (4.21), one has that

$$|e_2(t)|^2\leq g_2(t):=\left(\xi_2\left(|e_2(0)|\right)e^{-\lambda_2 t}+\frac{1}{2\lambda_1}\Big(H_2^{\tilde{f}_2}+\mathcal{K}_{2,1}H_1^{h_1}\Big)\left(\xi_1\left(|e_1(0)|\right)\right)^2\left(-e^{-2\lambda_1 t}+1\right)\right)^2 \tag{4.26}$$

It can be verified that function $g_2(t)$ is a linear combination of exponential functions $e^{-2\lambda_2 t}$, $e^{-(\lambda_2+2\lambda_1)t}$ and $e^{-4\lambda_1 t}$, all of which belong to class $\mathscr{I}-\mathscr{K}\mathscr{L}$ when $\lambda_1>0$ and $\lambda_2>0$ according to Lemma 1. Therefore, $\int_0^t c_3\left(|e_2(s)|^2\right)ds<\infty$ is satisfied if $|e_2(0)|$ is also bounded. Based on the developed distributed state estimation system, the estimation error $|e_3(t)|$ for subsystem 3 goes to zero as $t\to\infty$.

By carrying out the above analysis recursively for the remaining subsystems in sequence, we can prove that $\int_0^\infty\left(\max\left\{H_j^{\tilde{f}_j}+\mathcal{K}_{j,l}H_l^{h_l}:l\in\mathbb{L}_j\right\}|\epsilon_j(s)|^2\right)ds<\infty$ holds for all $j\in\mathbb{J}$. Further, let us define $\gamma(s):=cs^2$ with $c:=\max\left\{H_j^{\tilde{f}_j}+\mathcal{K}_{j,l}H_l^{h_l}:l\in\mathbb{L}_j\right\}$ and treat the estimation errors of its upstream subsystems as external inputs to subsystem j, $j\in\mathbb{J}$ (i.e., for subsystem j, $u:=\epsilon_j$). The above facts imply that the dynamics of the estimation error e_j, $j\in\mathbb{J}$, are iISS in the form of (4.9), and that $\int_0^\infty\gamma\left(|u(s)|\right)ds<\infty$ is satisfied for each subsystem j, $j\in\mathbb{J}$.

For any positive scalar δ_j, there exists $T_j>0$ so that $\int_{T_j}^\infty\gamma\left(|\epsilon_j(s)|\right)ds\leq a\delta_j$ with $0<a<1$ due to the fact that $\int_{T_j}^\infty\gamma\left(|\epsilon_j(s)|\right)ds\to 0$ as $T_j\to\infty$, $j\in\mathbb{J}$. Define a class $\mathcal{KL}$ function $\beta_j(s,t)=\xi_j(s)\,e^{-\lambda_j t}$ for $j\in\mathbb{J}$. Taking T_j as a new initial time instant, the dynamics of the estimation error e_j can be characterized as below:

$$|e_j(t+T_j)|\leq\beta_j\left(|e_j(T_j)|,t\right)+\int_{T_j}^{t+T_j}\gamma\left(|\epsilon_j(s)|\right)ds\leq\beta_j\left(|e_j(T_j)|,t\right)+a\delta_j \tag{4.27}$$

Based on the property of the class $\mathscr{K}\mathscr{L}$ function β_j , there exists $\tau_j>T_j$ such that $\beta_j\left(|e_j(T_j)|,\tau_j\right)\leq(1-a)\delta_j$, which leads to $|e_j(t)|<\delta_j$ for all $t\geq\tau_j$, $j\in\mathbb{J}$, when there

is no fault.

Moreover, based on the obtained results and the property of Lemma 2, we can prove that the estimation error trajectory $|e_j(t)|$ of subsystem j converges to zero in the absence of faults (i.e., $|e_j(t)| \to 0$ as $t \to \infty$), $\forall j \in \mathbb{J}$. By taking into account the convergence property of subsystem 1, we can prove the convergence of the estimation error of the entire system to zero when there is no fault in system (4.2). This completes the proof of Theorem 5. □

The distributed FDI mechanism to be developed only becomes effective for FDI after the distributed state estimation system gives acceptable state estimates; that is, after time instant $\tau_{\max}$ with $\tau_{\max} := \max\{\tau_i : i \in \mathbb{I}\}$ when the estimation error of each subsystem i, $i \in \mathbb{I}$, is bounded as $|e_i| \leq \delta_i$.

4.3 Distributed fault detection and isolation design

In this section, we establish a distributed output-feedback FDI mechanism based on the distributed state estimation system described in the Section 4.2. The FDI mechanism is able to account for both actuator and sensor faults.

4.3.1 Distributed state predictors and residual generators

For each subsystem, a state predictor and a residual generator are designed. Each predictor is used to provide state predictions for the corresponding subsystem as the references. Each residual generator is designed to generate residual signals for FDI based on the associated subsystem state estimate and the reference subsystem state.

Specifically, the state predictors for the n subsystems are designed as in the following form:

$$\begin{aligned} \dot{\hat{x}}_1(t) &= f_1\left(\hat{x}_1(t), u_1(t)\right) \\ \hat{x}_j(t) &= f_j\left(\hat{x}_j(t), u_j(t)\right) + \tilde{f}_j\left(\hat{x}_1(t), \ldots, \hat{x}_{j-1}(t)\right), \ j \in \mathbb{J} \end{aligned} \tag{4.28}$$

where $\hat{x}_i(t) \in \mathbb{R}^{n_{x_i}}$ denotes the state of predictor i, $i \in \mathbb{I}$. Predictor i, $i \in \mathbb{I}$, receives the

predictions (i.e. $\hat{x}_j$, $j = 1, \ldots, i-1$) of the upstream subsystems of subsystem i, and sends out its prediction to all the downstream subsystems continuously. The predictors are re-initialized every prediction horizon T. For each subsystem, two copies of the predictor are implemented with different re-initialization time instants. Based on the predictions given by the two copies of predictors, two residual sequences (Residual 1 and Residual 2) which characterize the discrepancy between the state predictions of the predictors and the state estimates of the state estimators are generated. The selection of the prediction horizons of the two sets of residuals is shown in Figure 4.3.

After $\tau_{\max}$, the estimation error of the entire system is bounded within an acceptable region when there is no fault. The first set of the state predictors starts to be evaluated and generates the first set of predictions (i.e., $\hat{x}_i^1$, $i \in \mathbb{I}$) from t_0 with $t_0 \geq \tau_{\max}$. Specifically, starting from t_0, the state predictors in the first set are initialized (re-initialized) at the beginning of each prediction horizon by the corresponding state estimate as $\hat{x}_i^1(t_k) = z_i(t_k)$ at sampling instant t_k where $t_k = t_0 + kN$, $k \geq 0$. Then, by solving the differential equations in (4.28) within each prediction horizon subject to the initialization conditions, we can obtain $\hat{x}_i^1(t)$ for $t \in [t_k, t_{k+1})$ for $k \geq 0$. The predictions within each prediction horizon constitute the first set of piecewise-continuous reference state predictions $\hat{x}_i^1$ for subsystem i, $\forall i \in \mathbb{I}$. This set of predictions is used for the calculation of Residual 1.

The second set of the predictors is evaluated to provide another set of predictions (i.e., $\hat{x}_i^2$, $i \in \mathbb{I}$) in the same fashion as the first set but with different initialization (re-initialization) time instants. The predictors in this set start to be evaluated at $t_0 + \frac{T}{2}$. In particular, these state predictors are initialized (re-initialized) as $\hat{x}_i^2(t_k + \frac{T}{2}) = z_i(t_k + \frac{T}{2})$, $i \in \mathbb{I}$. Again, by solving the differential equations in (4.28) within each prediction horizon subject to the above initialization conditions, we obtain $\hat{x}_i^2(t)$ for $t \in [t_k + \frac{T}{2}, t_{k+1} + \frac{T}{2})$ for $k \geq 0$, such that the second set of piecewise-continuous reference state predictions $\hat{x}_i^2$ for subsystem i, $\forall i \in \mathbb{I}$, is obtained. The second set of predictions is used for the calculation of Residual 2.

Then, the residual generators calculate two residual sequences (i.e., Residual 1 and Resid-

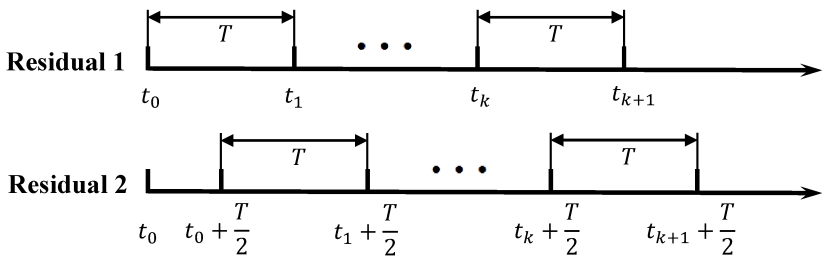

Figure 4.3: The prediction horizon selection for the two Residuals.

ual 2) based on the two sets of reference state predictions $\hat{x}_i^p$, $i \in \mathbb{I}$, $p \in \{1,\ 2\}$, and the subsystem state estimate z_i, $i \in \mathbb{I}$. Specifically, the residual generator for subsystem i, $i \in \mathbb{I}$, calculates two separate residual signals:

$$r_i^1(t) = \left|\hat{x}_i^1(t) - z_i(t)\right|, \quad t \in [t_0, t_{k+1}) \tag{4.29a}$$

$$r_i^2(t) = \left|\hat{x}_i^2(t) - z_i(t)\right|, \quad t \in [t_0 + \frac{T}{2}, t_{k+1} + \frac{T}{2}) \tag{4.29b}$$

where r_i^p represents the residual signal for subsystem i in Residual p, $p \in \{1,\ 2\}$. The two residual signals are used for fault detection and isolation.

Remark 18. *The use of two copies of the predictors with asynchronous re-initialization time instants ensures that the effect on the residuals of a fault occurring at any time can be accumulated for at least a period of $\frac{T}{2}$ (where T is the length of one prediction horizon) within the current prediction horizon of one of the two residual sequences. This design increase the probability that the residual signals become sufficiently large such that the occurred fault can be detected and isolated within the current prediction horizon of a residual sequence. Based on this design, the proposed approach can handle a wider range of faults (either sensor faults or actuator faults), and also can be used to detect and isolate an occurred fault earlier compared to the case where only one predictor is used. The advantages of using two copies of the predictors will be illustrated in Section 4.4.*

4.3.2 Fault detection and isolation

In this section, we perform analysis of the residual signals of each subsystem in both fault-free and faulty cases. Then, we present the conditions, under which the occurrence of a fault can be detected.

We introduce Lemma 3 and Lemma 4, which characterize the difference between the reference state of subsystem i, $i \in \mathbb{I}$, and the corresponding actual subsystem state in the absence of an actuator fault and in the presence of an actuator fault in subsystem i, respectively. The two lemmas will be used to determine the thresholds for the two sets of residual signals r_i^p , $i \in \mathbb{I}$, $p \in \{1,\ 2\}$.

Lemma 3. *Consider predictor (4.28) of subsystem i, $i \in \mathbb{I}$, which is re-initialized at each $t_{k\geq 0}$. If the nonlinear system (4.2) is operated in the absence of an actuator fault, and if the estimation error of subsystem i, $i \in \mathbb{I}$, is bounded as $|e_i(t_k)| \leq \delta_i$, $\forall i \in \mathbb{I}$, then the deviation of the reference state $\hat{x}_i$ of subsystem i from actual subsystem state x_i within a prediction horizon T from time instant $t_{k\geq 0}$ is bounded as:*

$$|e_i^x(t_k + \Delta)| \leq e^{L_{f_i}^{x_i}\Delta}\delta_i + \phi_i(\Delta) \tag{4.30}$$

for $i \in \mathbb{I}$, where $e_i^x := \hat{x}_i - x_i$, $\phi_1 := 0$, $\phi_j(s) := \frac{2\mathcal{N}_{\tilde{f}_j}}{L_{f_j}^{x_j}}(e^{L_{f_j}^{x_j}s} - 1)$ for $j \in \mathbb{J}$, $L_{f_i}^{x_i}$ denotes the Lipschitz constant of function f_i with respect to its first argument, $\mathcal{N}_{\tilde{f}_i}$ is a positive constant such that $|\tilde{f}_i(\cdot)| \leq \mathcal{N}_{\tilde{f}_i}$, and $\Delta := t - t_k$.

Proof: For subsystem j, the time derivative of the deviation $e_j^x(t)$ is calculated as:

$$\dot{e}_j^x(t) = f_j\left(\hat{x}_j(t), u_j(t)\right) + \tilde{f}_j\left(\hat{x}_1(t), \ldots, \hat{x}_{j-1}(t)\right) - f_j\left(x_j(t), u_j(t)\right) + \tilde{f}_j\left(x_1(t), \ldots, x_{j-1}(t)\right) \tag{4.31}$$

for $j \in \mathbb{J}$. Based on the Lipschitz properties of functions f_j, and the fact that $|\tilde{f}_j(\cdot)| \leq \mathcal{N}_{\tilde{f}_j}$,

the following inequality can be obtained:

$$|\dot{e}^x_{sj}(t)| \leq L^{x_j}_{f_j}|e^x_j(t)| + 2\mathcal{N}_{\tilde{f}_j} \tag{4.32}$$

With initial condition $|e^x_j(t_k)| = |\tilde{x}_j(t_k) - x_j(t_k)|$, based on (4.32), we can derive that:

$$|e^x_j(t)| \leq \frac{2\mathcal{N}_{\tilde{f}_j}}{L^{x_j}_{f_j}}(e^{L^{x_j}_{f_j}(t-t_k)} - 1) + |e^x_j(t_k)|e^{L^{x_j}_{f_j}(t-t_k)} \tag{4.33}$$

Since at any time instant t_k, we have that $e^x_i(t_k) = e_i(t_k)$ due to the initial condition of the i-th predictor, (4.33) can be rewritten as:

$$|e^x_j(t)| \leq \frac{2\mathcal{N}_{\tilde{f}_j}}{L^{x_j}_{f_j}}(e^{L^{x_j}_{f_j}(t-t_k)} - 1) + e^{L^{x_j}_{f_j}(t-t_k)}\delta_j \tag{4.34}$$

For subsystem 1, we have that $\mathcal{N}_{\tilde{f}_1} = 0$ and (4.34) reduces to

$$|e^x_1(t)| \leq e^{L^{x_1}_{f_1}(t-t_k)}\delta_1 \tag{4.35}$$

Taking into account the definition of Δ, Lemma 3 is proved. $\square$

Based on Lemma 3, we obtain the following lemma to describe the deviation e^x_i when an actuator fault takes place at $t_f \in [t_k, t_{k+1}]$ in subsystem i, $i \in \mathbb{I}$.

Lemma 4. *Consider predictor (4.28) of subsystem i, $i \in \mathbb{I}$, which is re-initialized at each $t_{k\geq 0}$. If the estimation error of subsystem i, $i \in \mathbb{I}$, is bounded as $|e_i(t_k)| \leq \delta_i$, $\forall i \in \mathbb{I}$, and if an actuator fault takes place at $t_f \in [t_k, t_{k+1})$ in subsystem i (i.e., $\tilde{u}_i \neq 0$), then the deviation of the reference state $\tilde{x}_i$ of subsystem i from the actual subsystem state x_i at the end of the present prediction horizon is bounded as follows:*

$$|e^x_i(t_{k+1})| \leq \psi_i(\Delta_f) + e^{L^{x_i}_{f_i}T}\delta_i \tag{4.36}$$

for $i \in \mathbb{I}$, *where* $\psi_i(s) =: \frac{L_{f_i}^{u_i}\tilde{u}_i^{\max}+2\mathcal{N}_{\tilde{f}_i}}{L_{f_i}^{x_i}}(e^{L_{f_i}^{x_i}(T-s)}-1)+e^{L_{f_i}^{x_i}(T-s)}\phi_i(s)$ *with* ϕ_i , $i \in \mathbb{I}$, *defined in Lemma 3,* $\Delta_f := t_f - t_k$, $L_{f_i}^{x_i}$ *and* $L_{f_i}^{u_i}$ *denote the Lipschitz constants of function* f_i *with respect to its first argument and its second argument, respectively,* $\mathcal{N}_{\tilde{f}_i}$ *is a positive constant such that* $|\tilde{f}_i(\cdot)| \leq \mathcal{N}_{\tilde{f}_i}$ *with* $\mathcal{N}_{\tilde{f}_1} = 0$ *and* $\tilde{u}_i^{\max} := \max_{\tau \in [t_f, t_{k+1}]} |\tilde{u}_i(\tau)|$, $i \in \mathbb{I}$.

Proof: The time derivative of the deviation $e_j^x(t)$ (for $t \in [t_f, t_{k+1}]$) can be calculated as:

$$\begin{aligned}\dot{e}_j^x(t) =& f_j\left(\hat{x}_j(t), u_j(t)\right) + \tilde{f}_j\left(\hat{x}_1(t), \ldots, \hat{x}_{j-1}(t)\right) \\ &- f_j\left(x_j(t), u_j(t) + \tilde{u}_j(t)\right) + \tilde{f}_j\left(x_1(t), \ldots, x_{j-1}(t)\right)\end{aligned} \tag{4.37}$$

for $j \in \mathbb{J}$. Based on the Lipschitz properties of the function f_j, and the fact that $|\tilde{f}_j(\cdot)| \leq \mathcal{N}_{\tilde{f}_j}$, it is calculated that:

$$|\dot{e}_j^x(t)| \leq L_{f_j}^{x_j}|e_j^x(t)| + L_{f_j}^{u_j}\tilde{u}_j^{\max} + 2\mathcal{N}_{\tilde{f}_j} \tag{4.38}$$

for $t \in [t_f, t_{k+1}]$. Based on (4.38), we can further derive that:

$$|e_j^x(t)| \leq \frac{L_{f_j}^{u_j}\tilde{u}_j^{\max} + 2\mathcal{N}_{\tilde{f}_j}}{L_{f_j}^{x_j}}(e^{L_{f_j}^{x_j}(t-t_f)}-1) + |e_j^x(t_f)|e^{L_{f_j}^{x_j}(t-t_f)} \tag{4.39}$$

Taking into account the initial condition $|e_j^x(t_f)| \leq e^{L_{f_j}^{x_j}(t_f-t_k)}\delta_j + \phi_j(t_f - t_k)$ based on (4.30), $|e_j^x(t_{k+1})|$ is bounded as:

$$|e_j^x(t_{k+1})| \leq \frac{L_{f_j}^{u_j}\tilde{u}_j^{\max} + 2\mathcal{N}_{\tilde{f}_j}}{L_{f_j}^{x_j}}(e^{L_{f_j}^{x_j}(t_{k+1}-t_f)}-1) + e^{L_{f_j}^{x_j}(t_{k+1}-t_f)}\phi_j(t_f - t_k) + e^{L_{f_j}^{x_j}T}\delta_j \tag{4.40}$$

for $j \in \mathbb{J}$. In terms of subsystem 1, we have that $\mathcal{N}_{\tilde{f}_1} = 0$, thus (4.40) can be rewritten as:

$$|e_1^x(t_{k+1})| \leq \frac{L_{f_1}^{u_1}\tilde{u}_1^{\max}}{L_{f_1}^{x_1}}(e^{L_{f_1}^{x_1}(t_{k+1}-t_f)}-1) + e^{L_{f_1}^{x_1}T}\delta_1 \tag{4.41}$$

Considering the definition of Δ_f, Lemma 4 is proved. □

Theorem 6 below characterizes the upper bounds on the residual signals for each subsys-

tem when there is no fault, and presents sufficient conditions, which can be used to detect the occurrence of an actuator/sensor fault.

Theorem 6. ***(Fault detection)*** *Consider the cascade nonlinear system (4.2), for which a distributed state estimation system with each augmented estimator designed following (4.8) is utilized. Moreover, we consider that at time instant t_0, the estimation error for each subsystem i has been bounded as $|e_i(t_0)| \leq \delta_i$, $\forall i \in \mathbb{I}$. If in Residual 1, any signal $r_i^1(t)$, $i \in \mathbb{I}$, at $t \in [t_k, t_{k+1})$ breaches a detection threshold $\theta_i^1(\Delta_1) := e^{L_{f_i}^{x_i}\Delta_1}\delta_i + \phi_i(\Delta_1) + \delta_i$ with $\Delta_1 := t - t_k$, or if in Residual 2, any signal $r_i^2(t)$, $i \in \mathbb{I}$, at $t \in [t_k + \frac{T}{2}, t_{k+1} + \frac{T}{2})$ breaches an detection threshold $\theta_i^2(\Delta_2) := e^{L_{f_i}^{x_i}\Delta_2}\delta_i + \phi_i(\Delta_2) + \delta_i$ with $\Delta_2 := t - t_k - \frac{T}{2}$, then it indicates that a fault (either an actuator fault or a sensor fault) has taken place in the system.*

Proof: If no fault (including actuator fault and sensor fault) occurs in any of the subsystems, then based on Lemma 3, the residual signal for subsystem i, $i \in \mathbb{I}$, of Residual 1 should be constrained as follows:

$$r_i^1(t) = \left|\hat{x}_i^1(t) - z_i(t)\right| \leq \left|\hat{x}_i^1(t) - x_i(t)\right| + |x_i(t) - z_i(t)| \leq e^{L_{f_i}^{x_i}(t-t_k)}\delta_i + \phi_i(t - t_k) + \delta_i \tag{4.42}$$

Therefore, within one prediction horizon of Residual 1 from t_k to t_{k+1}, if the signal $r_i^1(t)$ breaches the detection threshold $\theta_i^1(\Delta_1)$ for any $i \in \mathbb{I}$, then a fault (an actuator fault or a sensor fault) has taken place.

Similarly, we can prove that within one prediction horizon of Residual 2, if $r_i^2(t)$ breaches the corresponding detection threshold $\theta_i^2(\Delta_2)$, then a fault has taken place. □

Remark 19. *In the nonlinear context, the thresholds for the residual signals may be difficult to determine a priori. In implementations, the thresholds can be selected as (conservative) upper bounds of the residual signals after the estimation error given by the distributed state estimation system has converged into a small neighborhood to account for disturbances and measurement noise.*

Remark 20. *The condition presented in (4.30) provides a conservative upper bound on the deviation $|e_i^x|$, $i \in \mathbb{I}$, between the reference state $\hat{x}_i$ and the actual subsystem state x_i within a prediction horizon, which depends on the estimation error at the current time instant and the size of the prediction horizon. This upper bound can be made sufficiently small by appropriately choosing the size T and ensuring the estimation error to be sufficiently small, which can lead to less conservative thresholds.*

Remark 21. *In the present work, model-plant mismatch is not considered. In the presence of model-plant mismatch, the estimation error may be bounded within a region instead of exponentially converging to zero. Accordingly, the thresholds on the residual signals selected for FDI should be increased to avoid false alarms.*

4.3.3 Detectable faults

In an output-feedback context, not all types of faults can be detected by FDI mechanisms (see, e.g., Keliris et al. [78] and Du and Mhaskar [62]). In this subsection, we present sufficient conditions that characterize detectable actuator faults and sensor faults within the current prediction horizon of a residual sequence based on the proposed FDI scheme.

To facilitate the analysis, the increment in the estimation error for each subsystem after the occurrence of a fault (either an actuator fault or a sensor fault) is calculated. First, suppose that an actuator fault occurs in the i-th subsystem, $i \in \mathbb{I}$, at t_f (i.e., $\tilde{u}_i(\tau) \neq 0$, $\tau \geq t_f$). Based on (4.13), if the correction gains $K_{j,l}$, $j \in \mathbb{J}$, $l \in \mathbb{L}_j$, are determined following (4.7), then the derivative of the estimation error of subsystem i, $i \in \mathbb{I}$, given by the distributed

state estimation scheme is calculated as:

$$\dot{e}_1(t) = F_j\left(z_1(t), h_1(x_1(t)), u_1(t)\right) - f_j\left(x_1(t), u_1(t) + \tilde{u}_1(t)\right) \tag{4.43a}$$

$$\begin{aligned}\dot{e}_j(t) =& F_j\left(z_j(t), h_j(x_j(t)), u_j(t), X_{js}\right) - f_j\left(x_j(t), u_j(t) + \tilde{u}_j(t)\right) \\ & - \tilde{f}_j(X_{js}) + \sum_{l\in\mathbb{L}_j} K_{j,l} o_{h_l}(z_l^2) - \sum_{l\in\mathbb{L}_j} o_{\tilde{f}_j}(z_l^2),\ j \in \mathbb{J}\end{aligned} \tag{4.43b}$$

Therefore, the increment in the estimation error for subsystem i, $i \in \mathbb{I}$, from t_f (when the actuator fault takes place) to a time instant s (which is within the current prediction horizon) is given as follows:

$$\mathcal{M}_{\tilde{u}_1}(s) := \int_{t_f}^{s} \Big(F_1(z_1(\tau), h_1(x_1(\tau)), u_1(\tau)) - f_1(x_1(\tau), u_1(\tau) + \tilde{u}_1(\tau))\Big) d\tau \tag{4.44a}$$

$$\begin{aligned}\mathcal{M}_{\tilde{u}_j}(s) := \int_{t_f}^{s} \Big(& F_j(z_j(\tau), h_j(x_j(\tau)), u_j(\tau), X_{js}) - f_j(x_j(\tau), u_j(\tau) + \tilde{u}_j(\tau)) - \tilde{f}_j(X_{js}) \\ & + \sum_{l\in\mathbb{L}_j} K_{j,l} o_{h_l}(z_l^2) - \sum_{l\in\mathbb{L}_j} o_{\tilde{f}_j}(z_l^2)\Big) d\tau,\ j \in \mathbb{J}\end{aligned} \tag{4.44b}$$

Next, we consider the case where a sensor fault takes place in subsystem i, $i \in \mathbb{I}$, at time instant t_f. Similarly, the increment in the estimation error for each subsystem (denoted by $\mathcal{M}_{\tilde{v}_i}$ for subsystem i, $i \in \mathbb{I}$) from t_f to a time instant s within the current prediction horizon of a residual sequence is given as below.

$$\mathcal{M}_{\tilde{v}_1}(s) := \int_{t_f}^{s} \Big(F_1(z_1(\tau), h_1(x_1(\tau)) + \tilde{v}_1(\tau), u_1(\tau)) - f_1\left(x_1(\tau), u_1(\tau)\right)\Big) d\tau \tag{4.45a}$$

$$\begin{aligned}\mathcal{M}_{\tilde{v}_j}(s) := \int_{t_f}^{s} \Big(& F_j(z_j(\tau), h_j(x_j(\tau)) + \tilde{v}_j(\tau), u_j(\tau), X_{js}) - f_j\left(x_j(\tau), u_j(\tau)\right) - \tilde{f}_j(X_{js}) \\ & + \sum_{l\in\mathbb{L}_j} K_{j,l} o_{h_l}(z_l^2) - \sum_{l\in\mathbb{L}_j} o_{\tilde{f}_j}(z_l^2)\Big) d\tau,\ j \in \mathbb{J}\end{aligned} \tag{4.45b}$$

Then, the following two theorems summarize the findings.

Theorem 7. ***(Detectable actuator faults)*** *Consider the nonlinear cascade system (4.2), for which a distributed state estimation system is designed following Section 4.2, such that the estimation error for each subsystem i has been bounded as $|e_i(t_0)| \leq \delta_i$, $\forall i \in \mathbb{I}$. Then, an actuator fault in subsystem i (denoted as $\tilde{u}_i$, $i \in \mathbb{I}$) is detectable within the current prediction horizon of a residual sequence by the proposed FDI approach if $\tilde{u}_i$ takes place at $t_f \in [t_k, t_k + \frac{T}{2})$, and satisfies the following condition:*

$$|\bar{\varepsilon}_i + \mathcal{M}_{\tilde{u}_i}(t_{k+1})| > \theta_i^1(T) + e^{L_{f_i}^{x_i} T}\delta_i + \psi_i(t_f - t_k) \tag{4.46}$$

for all $|\bar{\varepsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$; or if the actuator fault takes place at $t_f \in [t_k + \frac{T}{2}, t_{k+1})$, and satisfies the condition:

$$|\bar{\varepsilon}_i + \mathcal{M}_{\tilde{u}_i}(q)| > \theta_i^2(T) + e^{L_{f_i}^{x_i} T}\delta_i + \psi_i(t_f - t_k - \frac{T}{2}) \tag{4.47}$$

for all $|\bar{\varepsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$, with $q = t_{k+1} + \frac{T}{2}$.

Proof: First, let us characterize the class of actuator faults that are detectable by the proposed FDI scheme. Suppose that an actuator fault occurs in the i-th subsystem, $i \in \mathbb{I}$, at $t_f \in [t_k, t_k + \frac{T}{2})$ (i.e., $\tilde{u}_i(t) \neq 0$, $t \geq t_f$). Taking into account the increment in the estimation error for subsystem i from t_f to t_{k+1} (i.e., the end of the current prediction horizon of Residual 1), the state estimation error for subsystem i at time instant t_{k+1} can be expressed as:

$$e_i(t_{k+1}) = e_i(t_f) + \mathcal{M}_{\tilde{u}_i}(t_{k+1}) \tag{4.48}$$

for $i \in \mathbb{I}$, where the function $\mathcal{M}_{\tilde{u}_i}$ denotes the increment in the subsystem estimation error as specified in (4.45).

On the other hand, based on the triangle inequality, the residual signal for subsystem i of Residual 1, $i \in \mathbb{I}$, satisfies:

$$r_i^1(t_{k+1}) = \left|\hat{x}_i^1(t_{k+1}) - z_i(t_{k+1})\right| \geq \left|\, |e_i(t_{k+1})| - |e_i^x(t_{k+1})| \,\right| \tag{4.49}$$

If an actuator fault takes place in the i-th subsystem, $i \in \mathbb{I}$, at $t \in [t_k, t_k + \frac{T}{2})$ and if the condition in (4.46) is satisfied for all $|\bar{\varepsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$, by considering the property (4.36) given in Lemma 4, we have that $r_i^1(t_{k+1}) \geq \theta_i^1(T)$ so that the fault can be detected within the current prediction horizon of Residual 1. Similarly, if an actuator fault takes place at $t \in [t_k + \frac{T}{2}, t_{k+1})$ and if inequality (4.47) holds, then we have $r_i^2(t_{k+1} + \frac{T}{2}) \geq \theta_i^2(T)$ and the fault can be detected within the current prediction horizon of Residual 2. This completes the proof of Theorem 7. $\square$

Theorem 8. ***(Detectable sensor faults)*** *Consider the nonlinear cascade system (4.2), for which a distributed state estimation system is designed following Section 4.2, such that the estimation error for each subsystem i has been bounded as $|e_i(t_0)| \leq \delta_i$, $\forall i \in \mathbb{I}$. Then, a sensor fault in subsystem i (denoted as $\tilde{v}_i$, $i \in \mathbb{I}$) is detectable if it takes place at $t_f \in [t_k, t_k + \frac{T}{2})$, and the following condition:*

$$|\bar{\varepsilon}_i + \mathcal{M}_{\tilde{v}_i}(t_{k+1})| > \theta_i^1(T) + e^{L_{f_i}^{x_i} T}\delta_i + \phi_i(T) \tag{4.50}$$

holds for all $|\bar{\varepsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$; or if $\tilde{v}_i$ occurs at $t_f \in [t_k + \frac{T}{2}, t_{k+1})$ and the following condition:

$$|\bar{\varepsilon}_i + \mathcal{M}_{\tilde{v}_i}(q)| > \theta_i^2(T) + e^{L_{f_i}^{x_i} T}\delta_i + \phi_i(T) \tag{4.51}$$

is satisfied for all $|\bar{\varepsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$, where $q = t_{k+1} + \frac{T}{2}$.

A sketch of the proof: Let us consider the case when a sensor fault takes place in the system (4.2). Following a similar procedure in the proof of Theorem 7, we can prove that if a sensor fault takes place in the i-th subsystem, $i \in \mathbb{I}$, at time instant $t_f \in [t_k, t_k + \frac{T}{2})$ and if the condition (4.50) is satisfied for all $|\bar{\epsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$; or if a sensor fault takes place at $t_f \in [t_k + \frac{T}{2}, t_{k+1})$ and if (4.51) holds for all $|\bar{\epsilon}_i| \leq \delta_i$, $i \in \mathbb{I}$, then a residual signal r_i^1 (or r_i^2) will breach the corresponding threshold θ_i^1 (or θ_i^2) within the prediction horizon of the Residual 1 (or Residual 2). $\square$

4.3.4 Fault isolation

In this section, we propose an approach to further isolate the subsystem in which the fault occurs.

Theorem 9. ***(Fault isolation)*** *Consider the cascade system (4.2), for which a distributed state estimation system described in Section 4.2 is utilized. Consider that at time instant t_0, the estimation error for each subsystem i has been bounded as $|e_i(t)| \leq \delta_i$, $\forall i \in \mathbb{I}$. Further, consider that in a faulty condition, only one fault occurs, and satisfies at least one of the conditions from (4.46) to (4.51) on detectable faults. If either r_1^1 or r_1^2 breaches the corresponding threshold, then the fault has taken place in subsystem 1; or if r_j^p, $j \in \mathbb{J}$, breaches the corresponding threshold for at least one $p \in \{1,\ 2\}$, while $r_l^p \leq \theta_l^p$ holds for all $p \in \{1,\ 2\}$, $l \in \{1, \ldots, j-1\}$, then the fault takes place at subsystem j.*

Proof: First, let us consider the case the fault satisfies either the condition (4.46) or the condition (4.50), such that Residual 1 can be used for FDI. Subsystem 1 can be regarded as an integral system itself as its dynamics is not affected by its downstream subsystems. Also, the subsystem estimator (the subsystem predictor) for subsystem 1 is not affected by the estimators (the predictors) of its downstream subsystems at all. The above fact implies that if $r_1^1(t) > \theta_i^1(\Delta_1)$ takes place, then there is a fault occurred in this subsystem.

The remainder of Theorem 9 is proved by contradiction. We consider the case that r_j^1, $j \in \mathbb{J}$, breaches its corresponding threshold between t_k and t_{k+1}, while $r_l^1(t) \leq \theta_l^1(\Delta_1)$ holds for all $l \in \{1, \ldots, j-1\}$ and $t \in [t_k, t_{k+1}]$. Then, let us consider subsystem 1 to subsystem j, $j \in \mathbb{J}$, as an entire system, which is not affected by the remained downstream subsystems (i.e., subsystem $j+1$ to subsystem n). According to Theorem 6, r_j^1, $j \in \mathbb{J}$, breaching the threshold between t_k and t_{k+1} implies that one fault takes place in one of the first j subsystems. If the fault takes place in subsystem l, $l \in \{1, \ldots, j-1\}$, but not in subsystem j, and if it satisfies either the condition (4.46) or the condition (4.50), based on Theorem 7 and Theorem 8, it is guaranteed that $r_l(t_{k+1}) \geq \theta_l(t_{k+1})$. The above argument contradicts

the condition that $r_l(t) \leq \theta_l(t)$ $\forall l \in \{1, \ldots, j-1\}$ and $t \in [t_k, t_{k+1}]$. As a result, the fault does not take place in any subsystem l, $l \in \{1, \ldots, j-1\}$, which demonstrates that the fault must occur in subsystem j.

Then, let us consider that the fault satisfies either the condition (4.47) or the condition (4.51). Similarly, we can prove that if $r_1^2(t) > \theta_i^2(\Delta_2)$ is satisfied, then the fault takes place in the first subsystem. Also, if r_j^2, $j \in \mathbb{J}$, breaches the corresponding threshold between t_k and t_{k+1}, while $r_l^2(t) \leq \theta_l^2(\Delta_2)$ holds for all $l \in \{1, \ldots, j-1\}$ and $t \in \left[t_k + \frac{T}{2}, t_{k+1} + \frac{T}{2}\right]$, then the fault takes place in subsystem j, $j \in \mathbb{J}$. This completes the proof of Theorem 9. □

Remark 22. *The proposed approach can be used to isolate the subsystem in which a fault has taken place. Currently, one cannot identify the type of the occurred fault by examining the residual signals of the subsystems following the proposed approach. One approach to augment the proposed approach so that the type of the fault can be identified based on the residuals (i.e., to determine whether it is an actuator fault or a sensor fault) is to have redundant sensors to measure each output of the entire system.*

Remark 23. *If a fault takes place but does not satisfy any of the fault detectability conditions from (4.46) to (4.51), then it may not be detected or isolated by the proposed FDI mechanism in the horizon during which a fault takes place. Note that it is still possible that the fault can be detected after one or several prediction horizons after its occurrence.*

4.4 Application to a froth flotation process example

In this section, the proposed FDI mechanism is applied to to a chemical process example to illustrate its effectiveness.

4.4.1 Process description

First, we introduce a process of froth flotation units that is utilized for the clean and recovery of fine coal produced by Coal Handling and Preparation Plants (CHPPs). The froth-flotation-unit process is comprised of five interconnected vessels. A schematic diagram of the process is presented in Figure 4.4. The slurry containing coal, ash and water is fed into the first vessel. Reagents consisting of 'frother' and 'collector' are added through the upstream of the each vessel to facilitate the formation of froth and the attachment of air bubbles to coal particles. The concentrations of the two reagents to each vessel are controlled to achieve improved productivity and profitability based on the solid concentration and the type of the coal. The froth is removed from the five vessels using the paddles and is further sent to a disk filter system for drying. The tailings of each of the first four vessels are fed into its neighboring downstream vessel, while the tailings of the fifth vessel are discharged and sent to a static thickener [84]. The dynamics of the process can be characterized by the concentration dynamics of the solids, liquids and ash in the vessels. Based on standard modeling assumptions and mass balances, the equations that describe the dynamics of the i-th vessel, $i \in \{1,\ 2,\ 3,\ 4,\ 5\}$, are given as follows [84]:

$$\frac{dc_{st_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{st_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{st_i}) - s_{r_i} \tag{4.52a}$$

$$\frac{dc_{lt_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{lt_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{lt_i}) - \frac{\beta_i}{V_i} \tag{4.52b}$$

$$\frac{dc_{at_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{at_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{at_i}) - \frac{\dot{A}_i}{V_i} \tag{4.52c}$$

where c_{st_i} represents the concentration of the solids in the i-th vessel (kg/m^3), c_{lt_i} represents the concentration of the liquid in the i-th vessel (kg/m^3), c_{at_i} is the concentration of the ash in the i-th vessel (kg/m^3), V_i denotes the slurry volume in the i-th vessel (m^3), $\dot{V}_{uf_i}$ denotes the flow rate of the underflow from the i-th vessel (m^3/min), r_i is the rate of the solid removal from the i-th vessel, β_i denotes the mass flow rate of the liquid from the i-th

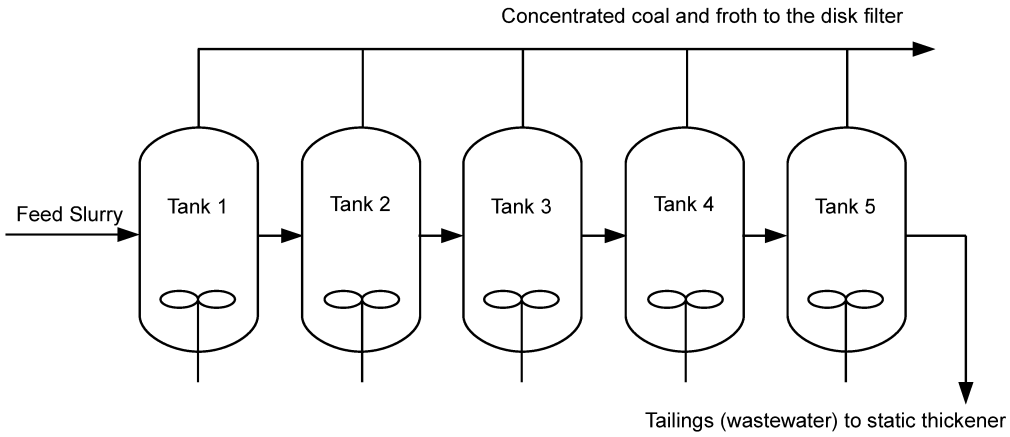

Figure 4.4: A schematic of the forth-flotation-unit process.

vessel to overflow (kg/min), $\dot{A}_i$ denotes the mass flow rate of ash from tank i to overflow (kg/min), c_{st_0}, c_{lt_0} and c_{at_0} represent the concentration of the solids, the concentration of the liquid and the concentration of the ash in the feed flow, respectively, and $\dot{V}_{uf_0}$ represents the volumetric flow rate of the feed flow.

The solid removal rate s_{r_i} is calculated following the equation below:

$$s_{r_i} = f_r k(c_{st_i} - c_\infty) \tag{4.53}$$

where k denotes a constant rate, c_∞ denotes the equilibrium of the solid concentration (kg/m^3), while f_r represents the correction factor of industrial scale reactions. A more detailed description of the model can be found in Canright et al. [84] and Yin et al. [61].

4.4.2 Simulation settings and results

We assume that the states c_{lt_i} and c_{at_i} are the output measurements. We use u_i^1 and u_i^2 to denote the input of frother and the input of collector to vessel i, $i \in \{1,\ 2,\ 3,\ 4,\ 5\}$. The entire process is decomposed into five subsystems according to its physical topology. For each vessel, a decentralized exponentially convergent estimator is designed based on the high-gain observer design proposed by Ciccarella et al. [30]. For each of the second to

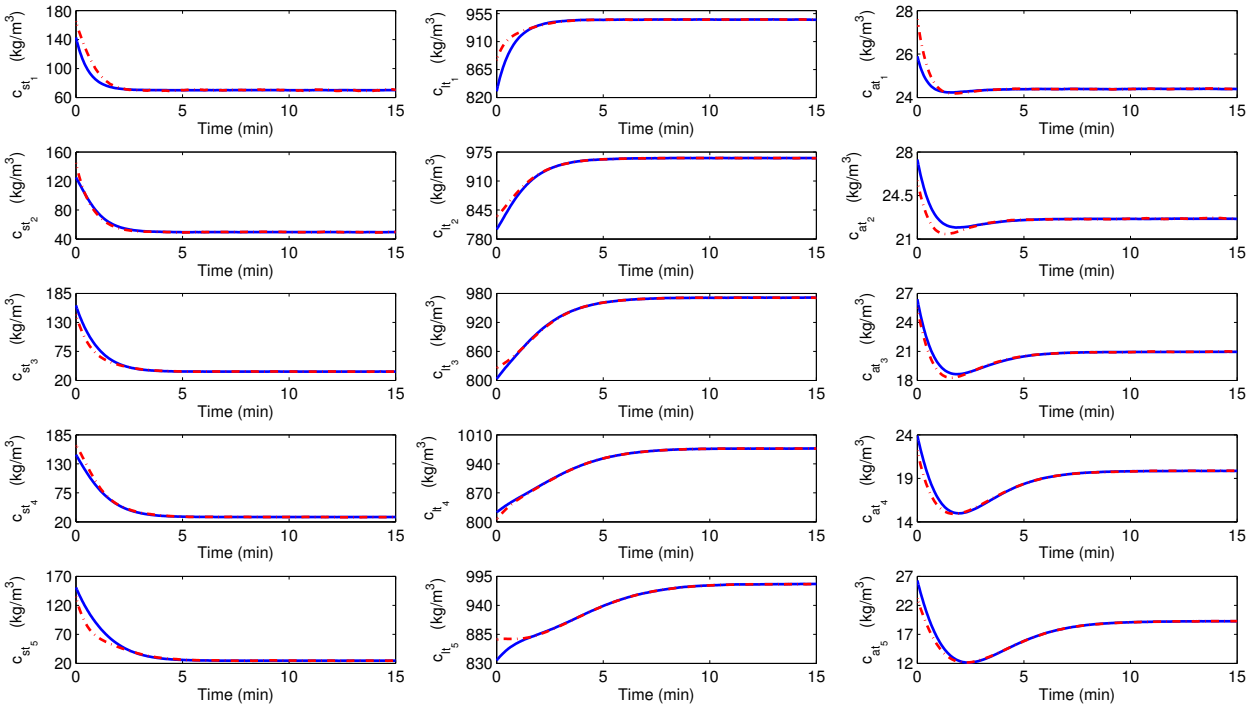

Figure 4.5: The state estimates given be the distributed state estimation system (red lines) and the actual system states (blue lines) in a fault-free condition.

the fifth vessels, an augmented estimator is designed by connecting each compensator to the associated existing decentralized estimator. The estimators of the five vessels communicate to exchange subsystem state estimates and measurements continuously, such that a distributed state estimation network is formed for the cascade process.

In the simulations, random measurement noise and system disturbances following normal distribution are added. We assume the initial state of the process is: $x_0 = [143\ \text{kg/m}^3,\ 830\ \text{kg/m}^3,\ 25.9\ \text{kg/m}^3,\ 125\ \text{kg/m}^3,\ 801.5\ \text{kg/m}^3,\ 27.4\ \text{kg/m}^3,\ 162.3\ \text{kg/m}^3,\ 803\ \text{kg/m}^3,\ 26.4\ \text{kg/m}^3,\ 147.5\ \text{kg/m}^3,\ 823\ \text{kg/m}^3,\ 23.9\ \text{kg/m}^3,\ 151\ \text{kg/m}^3,\ 835.5\ \text{kg/m}^3,\ 26.3\ \text{kg/m}^3]^T$, while the initial state estimate for the distributed state estimation system is set to be: $z_0 = [165\ \text{kg/m}^3,\ 880\ \text{kg/m}^3,\ 27.6\ \text{kg/m}^3,\ 145\ \text{kg/m}^3,\ 830\ \text{kg/m}^3,\ 25.7\ \text{kg/m}^3,\ 145\ \text{kg/m}^3,\ 825.5\ \text{kg/m}^3,\ 25.5\ \text{kg/m}^3,\ 165\ \text{kg/m}^3,\ 805.7\ \text{kg/m}^3,\ 22.2\ \text{kg/m}^3,\ 135\ \text{kg/m}^3,\ 875\ \text{kg/m}^3,\ 23.2\ \text{kg/m}^3]^T$. We first select an appropriate time instant t_0 by examining the state estimation errors when the process is operated in a fault-free condition. The actual system states and the corresponding state estimates given by the developed distributed state estimation systems are

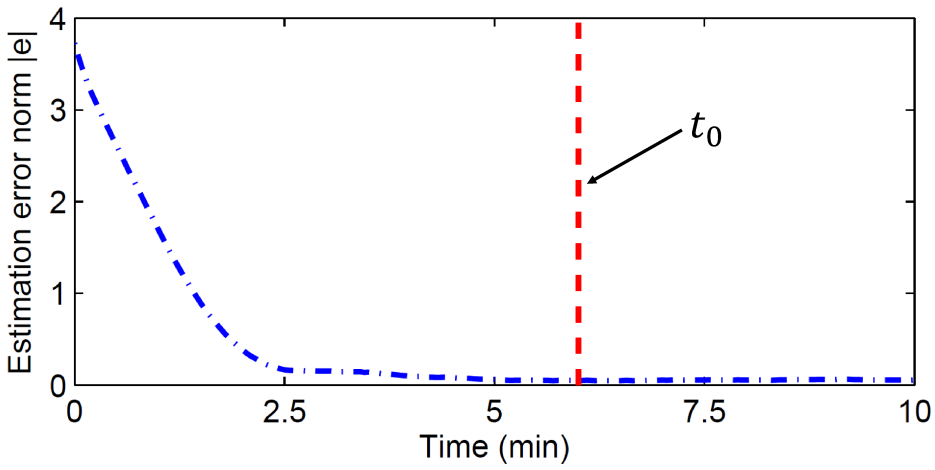

Figure 4.6: The trajectory of the normalized estimation error norm in a fault-free condition.

presented in Figure 4.5. The trajectory of the normalized estimation error norm is shown in Figure 4.6. We see that the state estimates are able to converge to the actual system states. Moreover, in the presence of system disturbances and measurement noise, the estimation error does not further decrease significantly after around 5 min. We set $t_0 = 6$ min, after which the FDI mechanism starts to be evaluated. The size of the prediction horizons for the two sets of residuals is chosen as $T = 2$ min.

Next, we take into account two faulty cases to illustrate the developed FDI mechanism: a single actuator fault case and a single sensor fault case. In both the two cases, we calculate normalized residual signals for each subsystem. Specifically, we denote $r_i^p(t) = \hat{x}_i^p - z_i = \left[r_{i,1}^p,\ r_{i,2}^p,\ r_{i,3}^p\right]^T$, and use $r_{i,j}^{p,\max}$ to denote a pre-determined value which should be no less than the maximum value of $r_{i,j}^p$ in the time span in the fault-free context. Then, the normalized residual signals for each subsystem i, $i \in \{1,\ 2,\ 3,\ 4,\ 5\}$, is calculated as follows:

$$|\bar{r}_i^p(t)| = \sqrt{\sum_{j\in\{1,2,3\}} \left(\frac{r_{i,j}^p(t)}{r_{i,j}^{p,\max}}\right)^2} \tag{4.54}$$

For this process, the thresholds for the residuals are selected by observing the magnitudes of the residual signals in the fault-free context. In particular, we first find a time instant t_0 as shown in Figure 4.6, after which the estimation error given by the distributed state

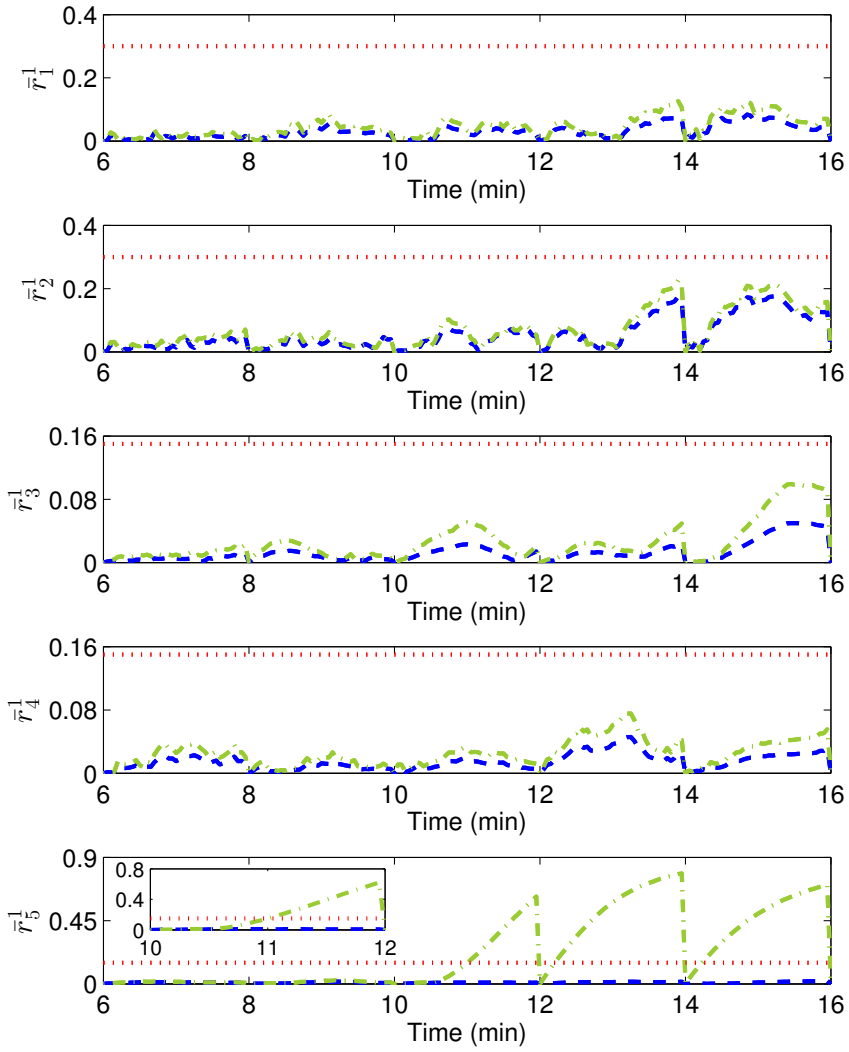

Figure 4.7: The residual signals of Residual 1 for the subsystems in a fault-free condition (blue dashed lines); the residuals of units with an actuator fault (green dash-dotted lines); selected thresholds (red dotted lines).

estimation system has converged into a small region. Then, the residual signals in the fault-free context after t_0 are generated and normalized following (4.54), and the upper bounds of the normalized residual signals for the subsystems are found. The thresholds are selected a little conservatively based on these upper bound values to account for disturbances and measurement noise. For distributed FDI, we choose the thresholds for the residual signals for the five subsystems as 0.3, 0.35, 0.15, 0.15, and 0.15.

First, we consider that there is an abrupt additive fault occurring at $t_f = 10.4$ min in the

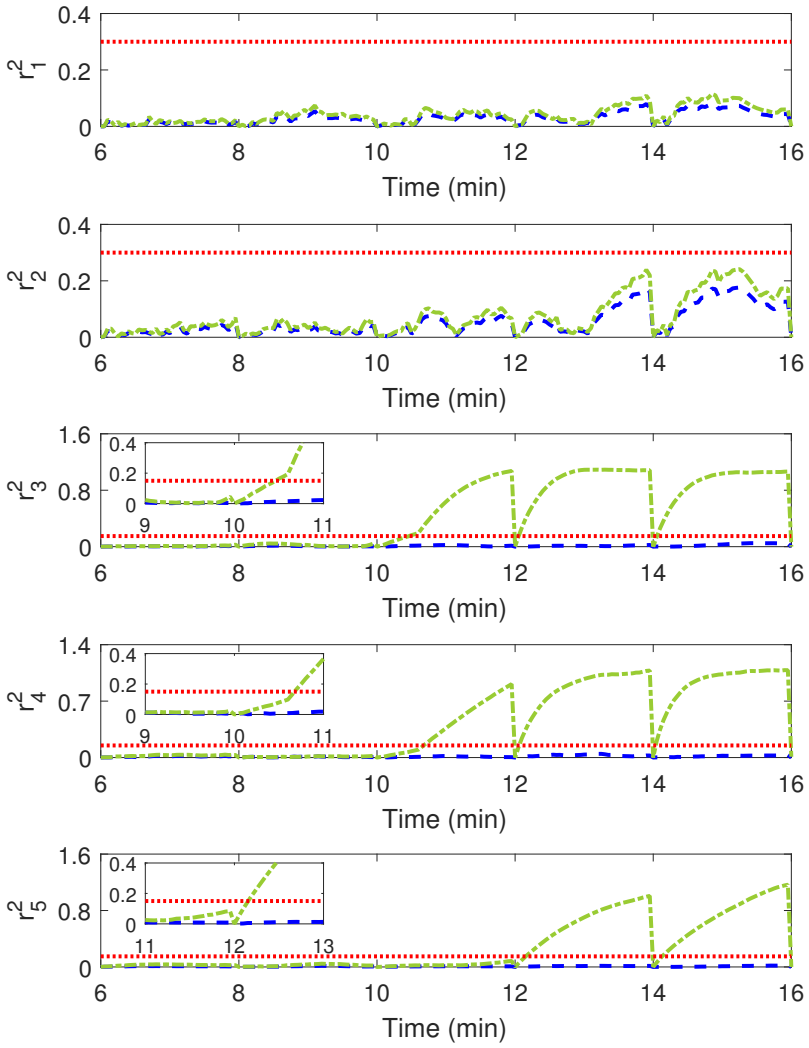

Figure 4.8: The residual signals of Residual 1 for the subsystems in a fault-free condition (blue dashed lines); the residuals of units with a sensor fault (green dash-dotted lines); selected thresholds (red dotted lines).

actuator for the reagent 'collector' fed into the 5-th vessel. The actuator fault is modeled as a constant bias $\tilde{u}_5^2(t) = 0.31 \times s(t - t_f)$ where $s(t - t_f) = \begin{cases} 0, t < t_f \\ 1, t \geq t_f \end{cases}$. The trajectories of the residual signals of Residual 1 for the five vessels are given in Figure 4.7. We see that all the residual signals for the first four vessels remain below the corresponding thresholds, while the residual signal for the last vessel breaches the corresponding threshold at around 11 min. Based on Theorem 6 and Theorem 9, the above observations result in the detection and isolation of a fault in the 5-th vessel.

Further, we consider that there is an additive sensor fault taking place in the sensor of

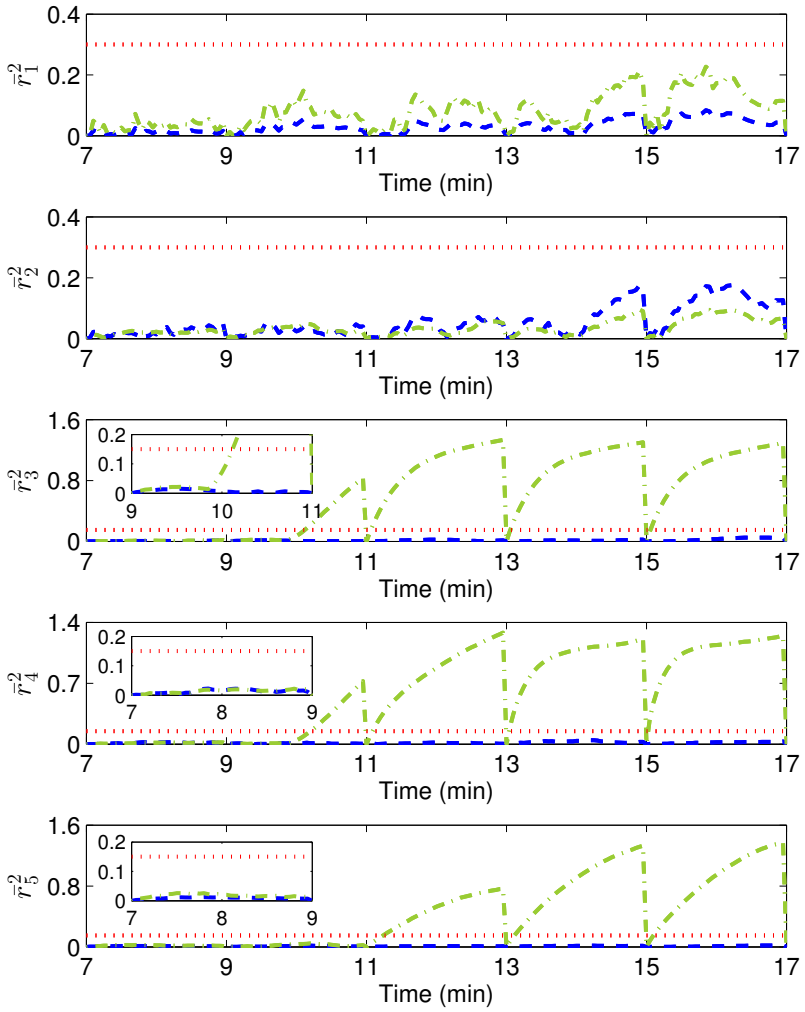

Figure 4.9: The residual signals of Residual 2 for the subsystems in a fault-free condition (blue dashed lines); the residuals of units with a sensor fault (green dash-dotted lines); selected thresholds (red dotted lines).

c_{lt_3}. A non-abrupt fault in the sensor for the measurement of c_{lt_3} from time $t_f = 9.8$min described by $\tilde{v}_3^l(t) = 40 \times \left(1 - e^{-(t-t_f)}\right) \times s(t - t_f)$kg/m^3 is added. The residual signals from both Residual 1 and Residual 2 are checked. The trajectories of the residual signals for the five subsystems given by Residual 1 and Residual 2 are shown in Figure 4.8 and Figure 4.9, respectively. The residual signals from both Residual 1 and Residual 2 for the first two vessels remain below the corresponding thresholds, while the residual signals for the rest three vessels all breach their corresponding thresholds after the occurrence of the fault. By examining the residual signals and taking advantage of Theorem 6 as well as Theorem 9,

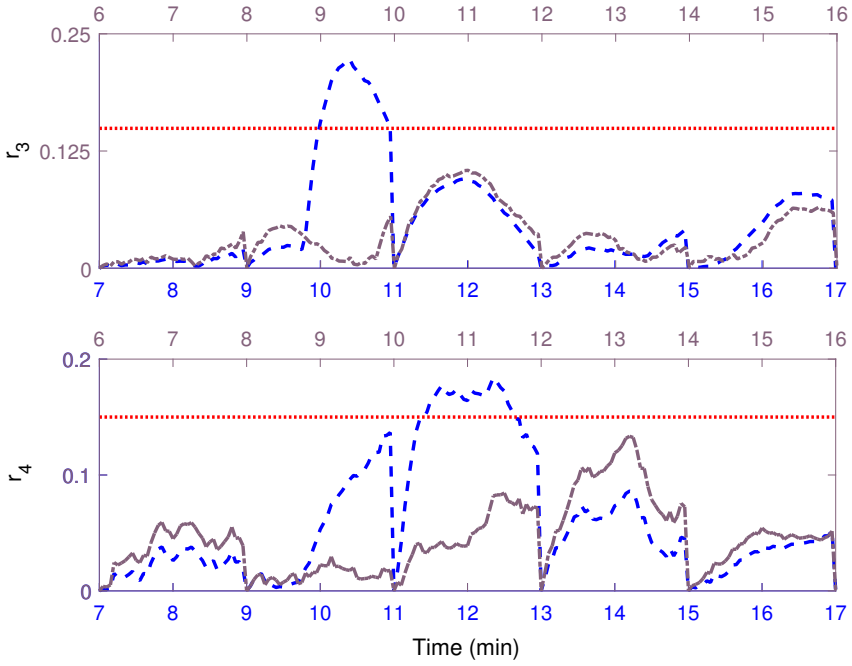

Figure 4.10: The residual signals for the third and fourth subsystems from Residual 1 with a sensor fault (black-solid lines); the residual signals for the third and fourth subsystems from Residual 2 with a sensor fault (blue-dashed lines); selected thresholds (red dotted lines).

a fault is detected by both Residual 1 and Residual 2. Moreover, the results show that by using two copies of subsystem predictors, the fault can be detected and isolated earlier. In particular, if we use only one set of subsystem predictors and only Residual 1 is available, the fault is detected at the time instant around 10.5 min, which is in the next prediction horizon of the horizon in which the fault occurred as seen from Figure 4.8. Based on Residual 2 in Figure 4.9, the occurred fault is detected earlier at the time instant around 10.1 min, and is identified as a fault in the third vessel.

Another set of simulations is carried out to show that the design with two copies of predictors can handle certain faults that cannot be detected when only one residual sequence is available. Suppose that a sensor fault occurs in the sensor of c_{lt_3} at $t_f = 9.8$min, which is modeled as $\tilde{v}_4^l(t) = 6.6e^{-(t-t_f)} \times s(t-t_f)\text{kg/m}^3$. The residual signals generated by both Residual 1 and Residual 2 are given in Figure 4.10. For brevity, only the residual signals for subsystem 3 and subsystem 4 are given. From Figure 4.10, it can be seen that this sensor

fault is not detected by Residual 1. This is mainly because the occurrence of this sensor fault is very close to the end of a prediction horizon (time instant 10min). On the other hand, the residual signal from Residual 2 for the third vessel breaches the threshold and is captured by Residual 2. This indicates that a fault has occurred in the third vessel based on Theorem 6 and Theorem 9.

Remark 24. *The purpose of using the normalized residual signals in (5.20) instead of the original residual signals is to mitigate the effects caused by the large difference in the magnitudes of the estimation errors of different states. Note that the use of the normalized residual signals in FDI does not violate the analysis on the proposed FDI mechanism.*

4.5 Summary

A systematic approach on distributed output-feedback FDI was proposed for nonlinear cascade processes. By assuming the existence of decentralized observers for the subsystems, a distributed state estimation was designed. In the absence of faults, the convergence to zero of the estimation error of the distributed estimation system was proved. A state predictor was designed for each subsystem to calculate subsystem state predictions. A residual generator was developed for each subsystem to generate residual signals for FDI. A distributed FDI mechanism applicable to both actuator faults and sensor faults was developed by evaluating the subsystem residual signals. Potential faults that can be detected and isolated by the developed mechanism were characterized. The proposed approach was shown to be effective via the application to a froth flotation process example.

Chapter 5

Subsystem decomposition and configuration for distributed state estimation

Distributed state estimation plays a very important role in process control. Improper subsystem decomposition for distributed state estimation may increase the computational burdens, degrade the estimation performance or even deteriorate the observability of the entire system. In this chapter, we investigate the subsystem decomposition problem for distributed state estimation of nonlinear systems. In Section 5.2, a systematic procedure for subsystem decomposition for distributed state estimation is proposed. Key steps in the procedure include observability test of the entire system, observable states identification for each output measurement, relative degree analysis and sensitivity analysis between measured outputs and states. Considerations with respect to time-scale multiplicity and direct graph are discussed. A few examples are used to illustrate the applicability of the methods used in different steps. Finally, in Section 5.3, the effectiveness of the entire distributed state estimation configuration procedure is also demonstrated via an application to a chemical process example used in coal handling and preparation plants. This chapter is a revised version of [72, 61].

5.1 System description

In this chapter, we consider a class of nonlinear systems described as the following state-space form:

$$\dot{x}(t) = f(x(t)) \tag{5.1a}$$

$$y(t) = h(x(t)) \tag{5.1b}$$

where $x(t) \in \mathbb{R}^{n_x}$ denotes the n-dimensional state vector of nonlinear system, $y(t)$ denotes the m-dimensional vector of measured outputs. It is assumed that system state vector x evolves within an operating region X, i.e., $x \in X$. We introduce $x_i(t)$, $i = 1, \ldots, n$ and $y_j(t)$, $j = 1, \ldots, m$ to denote the ith element of state vector x and jth element of output vector y, respectively.

The objective of this work is to develop a systematic procedure that can be used to decompose system (1) into a few subsystems for the purpose of distributed state estimation. We assume that we know the number of subsystems, p, that the system will be divided into and the number of subsystems is no greater than the number of measured outputs (i.e., $p \leq m$).

5.2 Proposed decomposition procedure

In this section, the proposed subsystem decomposition and configuration procedure for distributed state estimation is described. A flow diagram of the proposed procedure is shown in Figure 5.1.

5.2.1 Observability consideration

To design a state estimation system, a prerequisite is that the entire system is observable in the operating region. The first step in the proposed procedure is to check if system (5.1) is

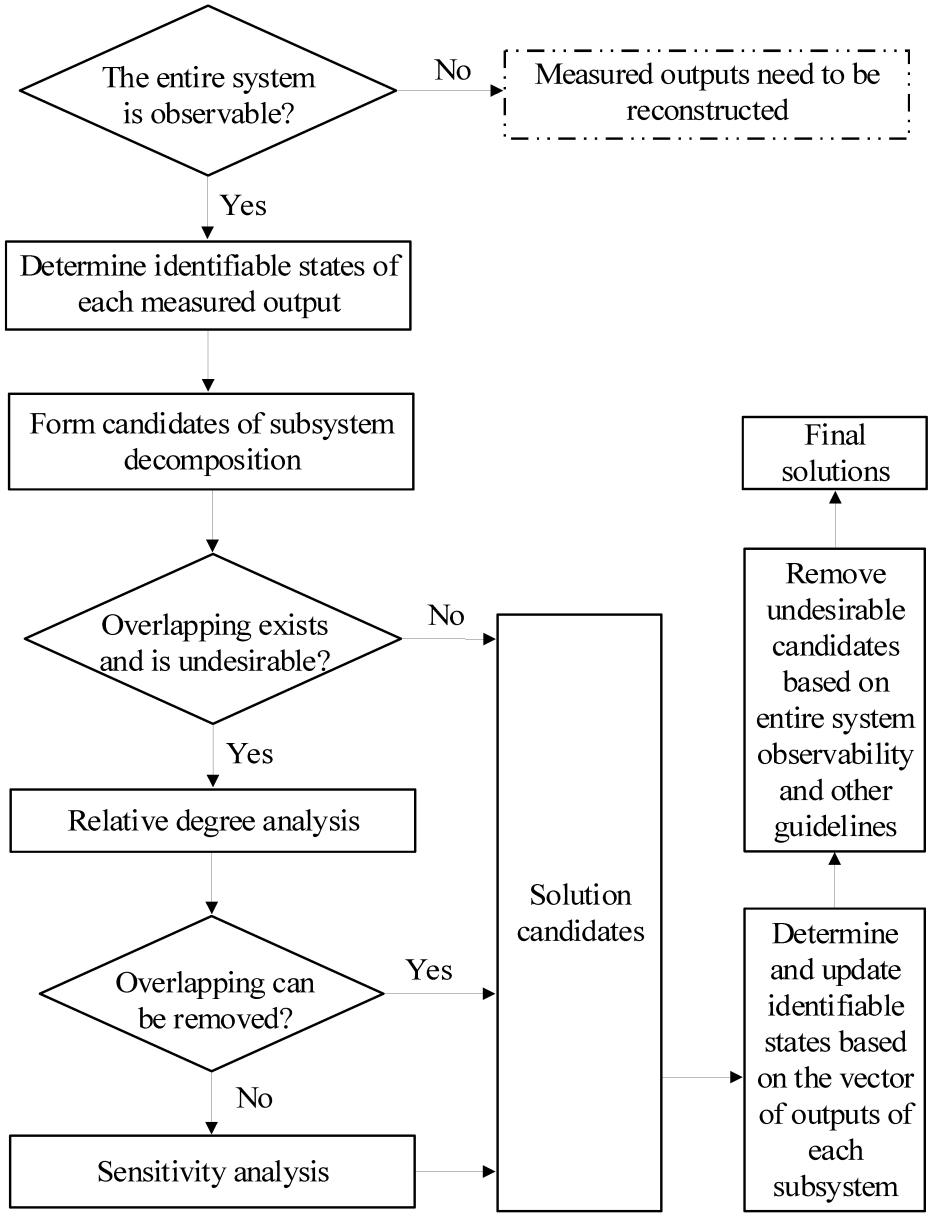

Figure 5.1: The flow diagram of the proposed decomposition procedure

observable in X. To check the observability of system (5.1), we can check if the observability matrix of the nonlinear system (5.1) is full rank. The observability matrix is defined as follows [30]:

$$Q(x) = \frac{d}{dx} \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{bmatrix} \tag{5.2}$$

where the symbol $L_f h$ denotes the Lie derivative of function h with respect to function f, defined as $L_f h(x) = \frac{\partial h}{\partial x} f(x)$, and $L_f^r h$ denotes the rth order Lie derivative, denoted as $L_f^r h(x) = L_f L_f^{r-1} h(x)$. If the matrix $Q(x)$ is full rank for all $x \in X$, then system (5.1) is locally observable in X. Otherwise, the system is not observable. Further steps in the proposed procedure cannot be performed. It is necessary to reconstruct the vector of measured outputs to ensure that system (5.1) is observable.

Remark 25. *Note that it may be computationally expensive to calculate the observability matrix and to check its rank in the entire region X. One approach is to linearize the nonlinear system along typical operating trajectories of the system and check the observability of the linearized systems obtained at different points along the trajectories.*

5.2.2 Identifiable states of each output

If the entire system is observable based on the set of measured outputs, the next step is to determine the identifiable states corresponding to each measured output. That is to determine the states can be estimated based on each measured output which can be done via Kalman canonical decomposition for system (5.1). However, we note that it may be computationally demanding to implement Kalman observable canonical decomposition based on the nonlinear systems. A practical approach is to perform Kalman decomposition of the linearized model at points along typical operating trajectories in region X.

Suppose that the linearized model obtained at a point x_s is as follows:

$$\begin{aligned}\dot{\tilde{x}} =& A\tilde{x} \\ \tilde{y}_j =& C_j\tilde{x}, j = 1, \ldots, m\end{aligned} \tag{5.3}$$

where $A = \frac{\partial f(x)}{\partial x}\Big|_{x=x_s}$ and $C_j = \frac{\partial h_j(x)}{\partial x}\Big|_{x=x_s}$. The local observability matrix $\mathcal{O}_j$ with respect to output variable y_j is constructed as follows:

$$\mathcal{O}_j = \begin{bmatrix} C_j \\ C_jA \\ \vdots \\ C_jA^{n-1} \end{bmatrix} \tag{5.4}$$

If $rank(\mathcal{O}_j) = n$, then all the variables of the state vector x are observable around the point x_s based on measurements of y_j. If $rank(\mathcal{O}_j) = n_1 < n$, then, we need to determine which states can be estimated based on y_j. Specifically, we first randomly select n_1 linearly independent rows from $\mathcal{O}_j$ to form the first n_1 rows of a square matrix P_j, then we construct the rest $n - n_1$ rows such that P_j is nonsingular. Once P_j is constructed, we perform a projection transformation as $z_j = P_j\tilde{x}$ and obtain the observability canonical form of the linearized model:

$$\dot{z}_j(t) = P_jAP_j^{-1}z_j(t) \tag{5.5a}$$

$$y_j = C_jP_j^{-1}z_j(t) \tag{5.5b}$$

We denote $\bar{A}_j = P_jAP_j^{-1}$ and $\bar{C}_j = C_jP_j^{-1}$, then the matrices are with the following forms:

$$\bar{A}_j = \begin{bmatrix} \bar{A}_j^o & 0 \\ \bar{A}_j^{21} & \bar{A}_j^{\bar{o}} \end{bmatrix}, \ \bar{C}_j = \begin{bmatrix} \bar{C}_j^o & 0 \end{bmatrix} \tag{5.6}$$

which indicates that the decomposed subsystem $(\bar{A}_j^o, \bar{C}_j^o)$ is observable whereas the subsystem $(\bar{A}_j^{\bar{o}}, 0)$ is unobservable. This result implies that the first n_1 state variables of z_j are observable given y_j while the rest $n - n_1$ state variables of z_j associated with $(\bar{A}_j^{\bar{o}}, 0)$ are unobservable.

For any $i = 1, \cdots, n$, if $\tilde{x}_i$ is a linear combination of the first n_1 elements of z_j, then $\tilde{x}_i$ is observable. In other words, if the last $n - n_1$ elements of ith row of matrix P_j^{-1}, $\forall i \in n$, are zero, then $\tilde{x}_i$ is observable. This implies x_i is locally observable at point x_s. Note that the above procedure should be performed for different points along the typical trajectories to ensure the reliability of the above conclusion.

Remark 26. *We note that linearization of the nonlinear system is required to be carried out along typical trajectories for this method. A linearized system model represents the dynamics of the original nonlinear system in a neighborhood of the reference point at which it is linearized. The size of the neighborhood depends on the nature of the original system. If an unexpected behavior is encountered such that the system operates at a region that is not characterized by the selected typical trajectories, it is possible that the observability analysis results may not hold any more. One practical recommendation is to consider different typical trajectories that cover the primary dynamics of the system to avoid the occurrence of the above situation. The recommendation is also applicable for the sensitivity analysis introduced later in Section 5.2.5.*

Remark 27. *The above method provides a natural way to identify the observable states for each output measurement. This method can be easily extended to determine the identifiable states for a given output vector containing more than one output variable. This will be used in Section 3.6, where we update the identifiable states for each local estimator we construct.*

5.2.3 Subsystem decomposition candidates

Once the identifiable states of an output are determined, the candidates of subsystem decomposition are formed by grouping the m outputs (and their identifiable states) into p groups

(i.e., subsystems). Obviously, there should be no empty groups. It can be calculated that there are $\frac{1}{p!}\sum_{j=1}^{p}(-1)^{p-j}C_p^j j^m$ combinations of output grouping.

Example 1. *Consider the following model of a four-water-tank system [114]:*

$$\begin{aligned}
\dot{H}_1 =& \frac{S_{o3}\sqrt{2gH_3} - S_{o1}\sqrt{2gH_1}}{S_1} + \frac{r_L k_L}{S_1} v_L \\
\dot{H}_2 =& \frac{S_{o4}\sqrt{2gH_4} - S_{o2}\sqrt{2gH_2}}{S_2} + \frac{r_R k_R}{S_2} v_R \\
\dot{H}_3 =& \frac{-S_{o3}\sqrt{2gH_3}}{S_3} + \frac{(1-r_R)k_R}{S_3} v_R \\
\dot{H}_4 =& \frac{-S_{o4}\sqrt{2gH_4}}{S_4} + \frac{(1-r_L)k_L}{S_4} v_L \\
y =& \begin{bmatrix} H_1 & H_2 \end{bmatrix}^T
\end{aligned} \tag{5.7}$$

where H_l, $l = 1, 2, 3, 4$, denotes the water level in each tank, S_{ol} and S_l, $l = 1, 2, 3, 4$, are sectional area constants, v_L and v_R are constant system inlets.

Suppose that we would like to decompose the system into two subsystems (i.e., $p = 2$). First, we verify that based on y, the entire system is observable. Then, we determine the identifiable states of each output. Following the approach discussed in Section 5.2.2, it is found that based on y_1 (i.e., H_1), states H_1 and H_3 can be estimated while based on y_2 (i.e., H_2), states H_2 and H_4 can be estimated. The only feasible decomposition is as follows:

Subsystem 1

$$\begin{aligned}
\dot{H}_1 =& \frac{S_{o3}\sqrt{2gH_3} - S_{o1}\sqrt{2gH_1}}{S_1} + \frac{r_L k_L}{S_1} v_L \\
\dot{H}_3 =& \frac{-S_{o3}\sqrt{2gH_3}}{S_3} + \frac{(1-r_R)k_R}{S_3} v_R \\
y_1 =& H_1
\end{aligned} \tag{5.8}$$

Subsystem 2

$$\begin{aligned}\dot{H}_2 =& \frac{S_{o4}\sqrt{2gH_4} - S_{o2}\sqrt{2gH_2}}{S_2} + \frac{r_R k_R}{S_2} v_R \\ \dot{H}_4 =& \frac{-S_{o4}\sqrt{2gH_4}}{S_4} + \frac{(1-r_L)k_L}{S_4} v_L \\ y_2 =& H_2\end{aligned} \tag{5.9}$$

5.2.4 Relative degree analysis

Each state should be associated with at least one subsystem (i.e., each state should be estimated by at least one subsystem estimator). If the subsystems in a solution candidate generated in Section 5.2.3 do not have overlaps in terms of observable states, then the subsystem decomposition forms a solution candidate that can be further checked by entire system observability and other guidelines that will be discussed in Section 3.6. If the subsystems have overlaps on states and overlapping decomposition is acceptable, the candidate can also be checked against the guidelines discussed in Section 3.6. However, if the subsystems have overlaps on states and this is not desirable, further analysis should be performed.

To address state overlapping in subsystems, we propose an effective index based on relative degree to determine the closeness between a system state variable x_i, $i = 1, \ldots, n$, and an output variable y_j, $j = 1, \ldots, m$.

We first define a matrix F as:

$$F = \frac{\partial f(x)}{\partial x} = \begin{bmatrix} \mathcal{F}_1 & \mathcal{F}_2 & \ldots & \mathcal{F}_n \end{bmatrix} \tag{5.10}$$

where $\mathcal{F}_1,\ \mathcal{F}_2, \cdots,\ \mathcal{F}_n$ are n-dimensional column vectors of the matrix F.

Similar to relative degree in output feedback control design [135], we define the relative

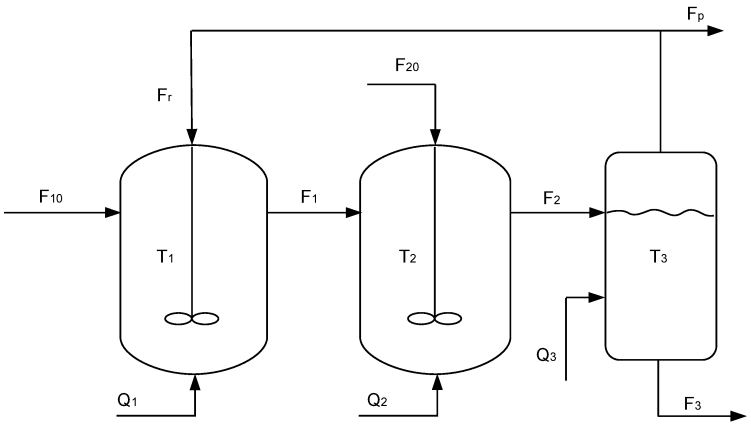

Figure 5.2: Schematic of the reactor-separator chemical process

degree $\mathcal{D}_{ij}$ between a state/output pair (x_i, y_j) as follows:

$$\begin{cases} \mathcal{D}_{ij} = 0, \ if \ \frac{\partial h_j(x)}{\partial x_i} \neq 0 \\ \mathcal{D}_{ij} = 1, \ if \frac{\partial h_j(x)}{\partial x_i} \equiv 0 \ \ and \ L_{\mathcal{F}_i} h_j(x) \neq 0 \\ \mathcal{D}_{ij} = d_{ij}, \ if \ \frac{\partial h_j(x)}{\partial x_i} \equiv 0 \ and \ L_{\mathcal{F}_i} L_f^{k-1} h_j(x) \equiv 0 \ and \ L_{\mathcal{F}_i} L_f^{d_{ij}-1} h_j(x) \neq 0 \end{cases} \tag{5.11}$$

where $k = 1, 2, \ldots, d_{ij} - 1$. The relative degree is $\mathcal{D}_{ij} = \infty$ if one fails to find a finite integer d_{ij}.

With the increase of the relative degree, the closeness of a state/output pair reduces correspondingly. Based on this developed criterion, we can determine the closeness between any state variable x_i and output variable y_j that we are concerned with and address the overlapping problem we may encounter. To be specific, if y_j and y_l are used in different local estimators and both can be used to estimate state variable x_i, then we can exploit the proposed concept to calculate $\mathcal{D}_{ij}$ and $\mathcal{D}_{il}$. If $\mathcal{D}_{ij} < \mathcal{D}_{il}$, then it is better to use the local estimator directly associated with output y_j to estimate the state x_j.

Example 2. *We consider a reactor-separator process that consists of two continuous stirred tank reactors (CSTR) and one flash tank separator. The schematic is presented in Figure 5.2*

and the system model is as follows [15]:

$$
\begin{aligned}
\frac{dx_{A_1}}{dt} =& \frac{F_{10}}{V_1}(x_{A10} - x_{A_1}) + \frac{F_r}{V_1}(x_{Ar} - x_{A1}) - k_1 e^{\frac{-E_1}{RT_1}} x_{A_1} \\
\frac{dx_{B_1}}{dt} =& \frac{F_{10}}{V_1}(x_{B10} - x_{B_1}) + \frac{F_r}{V_1}(x_{Br} - x_{B1}) + k_1 e^{\frac{-E_1}{RT_1}} x_{A_1} - k_2 e^{\frac{-E_2}{RT_1}} x_{B_1} \\
\frac{dT_1}{dt} =& \frac{F_{10}}{V_1}(T_{10} - T_1) + \frac{F_r}{V_1}(T_3 - T_1) + \frac{-\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_1}} x_{A_1} + \frac{-\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_1}} x_{B_1} + \frac{Q_1}{\rho c_p V_1} \\
\frac{dx_{A_2}}{dt} =& \frac{F_1}{V_2}(x_{A_1} - x_{A_2}) + \frac{F_{20}}{V_2}(x_{A20} - x_{A_2}) - k_1 e^{\frac{-E_1}{RT_2}} x_{A_2} \\
\frac{dx_{B_2}}{dt} =& \frac{F_1}{V_2}(x_{B_1} - x_{B_2}) + \frac{F_{20}}{V_2}(x_{B20} - x_{B_2}) + k_1 e^{\frac{-E_1}{RT_2}} x_{A_2} - k_2 e^{\frac{-E_2}{RT_2}} x_{B_2} \\
\frac{dT_2}{dt} =& \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_{20}}{V_2}(T_{20} - T_2) + \frac{Q_2}{\rho c_p V_2} + \frac{-\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_2}} x_{A_2} + \frac{-\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_2}} x_{B_2} \\
\frac{dx_{A_3}}{dt} =& \frac{F_2}{V_3}(x_{A_2} - x_{A_3}) - \frac{(F_r + F_p)}{V_3}(x_{Ar} - x_{A_3}) \\
\frac{dx_{B_3}}{dt} =& \frac{F_2}{V_3}(x_{B_2} - x_{B_3}) - \frac{(F_r + F_p)}{V_3}(x_{Br} - x_{B_3}) \\
\frac{dT_3}{dt} =& \frac{F_2}{V_3}(T_2 - T_3) + \frac{Q_3}{\rho c_p V_3} + \frac{(F_r + F_p)}{\rho c_p V_3}(x_{Ar}\Delta H_{vap1} + x_{Br}\Delta H_{vap2} + x_{Cr}\Delta H_{vap3})
\end{aligned}
\tag{5.12}
$$

It is assumed that the amount of reaction occurring in the separator is negligible. The algebraic equations modeling the composition of the overhead stream relative to the composition of the liquid holdup in the flash tank are presented as follows:

$$x_{Ar} = \frac{\alpha_A x_{s2}^A}{\alpha_A x_{s2}^A + \alpha_B x_{s2}^B + \alpha_C x_{s2}^C} \tag{5.13a}$$

$$x_{Br} = \frac{\alpha_B x_{B3}}{\alpha_A x_{s2}^A + \alpha_B x_{s2}^B + \alpha_C x_{s2}^C} \tag{5.13b}$$

$$x_{Cr} = \frac{\alpha_C x_{s2}^C}{\alpha_A x_{s2}^A + \alpha_B x_{s2}^B + \alpha_C x_{s2}^C} \tag{5.13c}$$

In this process, the state vector is denoted as $x = [x_{A_1}\ x_{B_1}\ T_1\ x_{A_2}\ x_{B_2}\ T_2\ x_{A_3}\ x_{B_3}\ T_3]^T$ *where* x_{A_i}, x_{B_i} $i = 1,\ 2,\ 3$, *represent the concentration of materials A and B, respectively,* T_i *denotes the temperature in tank* i, $i = 1,\ 2,\ 3$. *The definition of the variables can be found in [15]. We assume that* $y = [T_1\ \ T_2\ \ T_3]^T$ *is the measured output vector and the objective is*

Table 5.1: Observable states with each given output measurement for the reactor-separator process

Output	Observable states					
y_1	x_{A_1}	x_{B_1}	T_1	x_{A_3}	x_{B_3}	T_3
y_2	x_{A_1}	x_{B_1}	T_1	x_{A_2}	x_{B_2}	T_2
y_3	x_{A_2}	x_{B_2}	T_2	x_{A_3}	x_{B_3}	T_3

Table 5.2: Relative degree between each state and each output measurement

	x_{A_1}	x_{B_1}	T_1	x_{A_2}	x_{B_2}	T_2	x_{A_3}	x_{B_3}	T_3
y_1	1	1	0	n/a	n/a	n/a	2	2	1
y_2	2	2	1	1	1	0	n/a	n/a	n/a
y_3	n/a	n/a	n/a	2	2	1	1	1	0

to design three (i.e., $p = 3$) subsystems for distributed state estimation.

Firstly, we verify that the entire system is locally observable within a neighborhood around the equilibrium point $x = x_s$. Secondly, we aim at identifying the observable states for each specific output measurement y_i, $i = 1,\ 2,\ 3$. The results are presented in Table 5.1. From the results we can know that with known reactant temperature T_1 in the first CSTR, we are able to estimate all the state variables in the first tank and the separator. Likewise, if we know temperature T_2 in the second CSTR, then we can estimate all system states in the two CSTRs. Moreover, state variables in the second CSTR and the separator can be estimated given temperature T_3.

One decomposition candidate can be formed with state overlaps according to the three outputs and their identifiable states. Specifically, subsystems are configured as Subsystem 1: $\{y_1,\ x_{A_1},\ x_{B_1},\ T_1,\ x_{A_3},\ x_{B_3},\ T_3\}$; Subsystem 2: $\{y_2,\ x_{A_2},\ x_{B_2},\ T_2,\ x_{A_1},\ x_{B_1},\ T_1\}$; Subsystem 3: $\{y_3,\ x_{A_3},\ x_{B_3},\ T_3,\ x_{A_2},\ x_{B_2},\ T_2\}$.

If overlapping is not desirable, it can be addressed by the proposed relative degree analysis. Following (5.11), we calculate the relative degree $\mathcal{D}_{ij}$ between the output measurement y_j, $j = 1,\ 2,\ 3$ and overlapped states. The results are given in Table 5.2.

For each state, we choose the output with a smaller relative degree to estimate the state.

Based on the results in Table 5.2, we use subsystem i (containing output y_i), $i = 1,\ 2,\ 3$ to estimate states x_{A_i}, x_{B_i} and T_i, $i = 1,\ 2,\ 3$, that is, Subsystem 1: $\{y_1,\ x_{A_1},\ x_{B_1},\ T_1\}$; Subsystem 2: $\{y_2,\ x_{A_2},\ x_{B_2},\ T_2\}$; Subsystem 3: $\{y_3,\ x_{A_3},\ x_{B_3},\ T_3\}$.

Remark 28. *In many existing distributed state estimation results (e.g., [15]), a system is decomposed into subsystems based on operating units which is a natural and intuitive approach. The results of Example 2 shows that the relative degree based subsystem decomposition can be consistent with (and provide theoretical support for) the decomposition method according to operating units.*

5.2.5 Sensitivity analysis

It is possible that a state that can be estimated by different subsystems that have the same relative degree with respect to the outputs in the two subsystems. In this case, it is not sufficient to avoid state overlapping by only performing relative degree analysis. To address this issue and inspired by [137], sensitivity analysis should be carried out.

Performing Taylor expansion on the jth output variable $y_j(t)$, it is derived that

$$y_j(t) = \sum_{k=0}^{\infty} \frac{(t - t_s)^k}{k!} \frac{d^k y(t_s)}{dt^k} \tag{5.14}$$

From (5.1), the kth derivative of y with respect to t is denoted as

$$\frac{d^k y_j(t_s)}{dt^k} = L_f^k h_j(x(t_s)) \tag{5.15}$$

Substituting (5.15) into (5.14), we have

$$y_j(t) = \sum_{k=0}^{\infty} \frac{(t - t_s)^k}{k!} L_f^k h_j(x(t_s)) \tag{5.16}$$

Ignoring higher orders of (5.16) with $k > \mathcal{D}_{ij}$, it is re-written as:

$$y_j(t) \approx \sum_{k=0}^{D_{ij}} \frac{(t-t_s)^k}{k!} L_f^k h_j(x(t_s)) \tag{5.17}$$

Then we define $\mathcal{R}_{ij}$ as follows:

$$\mathcal{R}_{ij} = \left| \frac{\partial y_j}{\partial x_i} \right| \tag{5.18}$$

We noted that this method is designed for the case that the relative degree satisfies $\mathcal{D}_{ij} > 0$. From (5.18), we have

$$\begin{aligned} \mathcal{R}_{ij} =& \left| \frac{\partial y_j}{\partial x_i} \right| = \left| \frac{\partial h_j(x(t_s))}{\partial x_i} + \sum_{k=1}^{\mathcal{D}_{ij}} \frac{(t-t_s)^k}{k!} L_{\mathcal{F}_i} L_f^{k-1} h_j(x(t_s)) \right| \\ =& \left| \frac{(t-t_s)^{\mathcal{D}_{ij}}}{\mathcal{D}_{ij}!} L_{\mathcal{F}_i} L_f^{\mathcal{D}_{ij}-1} h_j(x(t_s)) \right| \end{aligned} \tag{5.19}$$

Since $\frac{(t-t_s)^{d_i^*}}{d_i^*!}$ is a constant coefficient for a fixed relative degree, the sensitivity index between state variable x_i and output variable y_j is defined as follows:

$$\bar{\mathcal{R}}_{ij} = \left| L_{\mathcal{F}_i} L_f^{d_i^*-1} h_j(x(t_s)) \right| \tag{5.20}$$

where $d_i^* = \mathcal{D}_{ij} = \mathcal{D}_{il}$. With (5.20), we calculate the sensitivity indices $\bar{\mathcal{R}}_{ij}$ and $\bar{\mathcal{R}}_{il}$ for state/output pairs (x_i, y_j) and (x_i, y_l), respectively. If $\bar{\mathcal{R}}_{ij} > \bar{\mathcal{R}}_{il}$, then it indicates that x_i is more sensitive to y_j compared to y_l. As a result, we should use y_j to estimate the state x_i, and vice versa. It is recommended to calculate the sensitivity indices at different points along typical trajectories to get a more complete picture of the sensitivity.

Remark 29. *It is recommended to conduct relative degree analysis first since it requires less knowledge on the system model compared to sensitivity analysis, which makes the calculation much easier and more efficient. Sensitivity analysis needs to be carried out only when relative degrees for different state/output pairs are equal. We note that sensitivity analysis actually*

serves as a supplementary tool of relative degree analysis and they do not conflict with each other on analysis results.

5.2.6 Update of identifiable states

Potential solutions to subsystem decomposition are generated via above analysis. Before forming the final solutions, we verify the observability of the entire system again by identifying observable states of each subsystem for each candidate. In this step, the observable states of the outputs within a subsystem should be determined based on the entire set of the outputs. This can be done following similar approach as described in Section 5.2.2. In Section 5.2.2, the observable states are identified based on each individual output. A combination of a few outputs may be able to estimate states that can not be estimated by either of the individual outputs. The subsystem should be updated with all the states it can estimate.

The necessity of this step is illustrated via the following example.

Example 3. *Consider a distillation tower consisting of a number of trays. For analytical convenience, we consider the distillation tower has three trays, and the system model is given as follows [117]:*

$$\begin{aligned} \dot{x}_{i,j} =& \frac{1}{M_j}\left[\bar{V}(y_{i,j+1} - y_{i,j}) + R(x_{i,j-1} - x_{i,j})\right] \\ \dot{H}_{L,j} =& \frac{\bar{V}}{M_j}(H_{V,j+1} - H_{V,j}) + \frac{R}{M_j}(H_{L,j-1} - H_{L,j}) \end{aligned} \tag{5.21}$$

where $x_{i,j}$ represents the species composition of product i, $i = A,\ B,\ C$, in tray j, $j = 1,\ 2,\ 3$, of the distillation tower; $H_{L,j}$ is the enthalpy of liquid mixture in corresponding tray j, $j = 1,\ 2,\ 3$; M_j, $j = 1,\ 2,\ 3$, is the liquid hold-up in tray j; $\bar{V}$ and R denote the effluent flow rates to the distillation tower; $y_i = \alpha_i x_i / \left(\sum_k^{A,B,C} \alpha_k x_k\right)$ with α_i, $i = A,\ B,\ C$, being a weight coefficient for vapor phase molar composition. We note that $x_{i,0}$ and $H_{L,0}$ denote the species compositions and enthalpy from the condenser, and are assumed to be known in this example. The enthalpy $H_{L,j}$, $j = 1,\ 2,\ 3$, in each tray is assumed to be available as output measurements. Following the method proposed in Section 5.2.2, it is

calculated that based on each individual output measurement $H_{L,j}$, $j = 1,\ 2,\ 3$, all states with respect to species composition are unobservable. However, based on combined output vector $y = [H_{L,1}\ H_{L,2}\ H_{L,3}]^T$, system states $x_{i,j}$, $i = A,\ B,\ C$, $j = 1,\ 2$, (i.e. species compositions in the 1st and 2nd trays) become observable. This example demonstrates that a combination of different outputs may be used to estimate additional states that cannot be estimated using each individual output.

Remark 30. *In terms of Example 3, one may intuitively decompose the distillation tower system into three subsystems based on the topology (i.e., the three physical trays) without carrying out the proposed procedure. However, as illustrated in Example 3, such decomposition structure will render the system unobservable. This phenomenon implies that intuitive decomposition approaches are not always reliable, it also illustrates the usefulness and necessity of the proposed subsystem decomposition procedure.*

5.2.7 Reduction of feasible solutions

After the above step in Section 5.2.6, we need to check again each candidate to see if the union of the states in its subsystems is the entire system state vector. Specifically, we obtain the set of observable system states $\mathcal{X}_j$, $j = 1, \cdots, p$ for each local estimator, and denote $\mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2 \ldots \cup \mathcal{X}_p$. If $x_i \in \mathcal{X}$ holds for all i, $i = 1, \ldots, n$, then the designed candidate is feasible. If the union $\mathcal{X}$ does not contain all the states in the entire system state vector, the candidate should be removed.

All the remaining candidates are feasible solutions for subsystem decomposition. Further guidelines can be taken into account to reduce the number of feasible solutions. For instance, a candidate with subsystems of similar number of states is more favorable compared with a candidate has unbalanced states in subsystems from a computational and communicational point of view. Also, subsystem decomposition aligns with the system's physical topology is more favorable than other decompositions.

By performing the above steps, final solutions to subsystem decomposition for distributed

state estimation are obtained. If there is no feasible solution, we should reduce the number of decomposed subsystems (i.e., p) and then carry out the above steps again until a feasible solution is found.

5.2.8 Further considerations

5.2.8.1 Time-scale multiplicity

Time-scale multiplicity is a common feature of many chemical processes including catalytic crackers, distillation columns, biochemical reactors, etc. For chemical processes, time-scale multiplicity arises typically due to the strong coupling of physicochemical phenomena. When time-scale multiplicity presents, direct application of state estimation methods on the original system may lead to ill-conditioning or even deteriorate the stability of the entire system. Therefore, time-scale multiplicity should be also taken into account in subsystem decomposition. One example of distributed state estimation based on time-scale separation is [42], in which a distributed MHE scheme is designed for a class of systems that can be described in the standard singularly perturbed framework. However, we note that there are many open problems in subsystem decomposition with respect to time-scale separation including how to handle systems that can not be described in the standard singularly perturbed framework.

5.2.8.2 Direct graph

Although relative degree analysis is effective in terms of determining the closeness between a given system state and an output measurement, the relative degrees for specific state/output pairs in large-scale processes may be relatively large, and the computation becomes increasingly complex because the calculation for high order Lie derivatives is challenging. Moreover, exact system parameters may not be available and relative degree cannot be calculated. Graph theory based methods like the direct graph method used in [118] can be adopted to overcome the above difficulties in the calculation of the relative degrees.

Denoting $f(x) = [f_1(x)\ f_2(x)\ \ldots\ f_n(x)]^T$, The construction of a direct graph for a typical nonlinear system model (5.1) is conducted using the following rules:

1. If $\partial f_k(x)/\partial x_i \neq 0$, $k, i = 1,\ 2,\ \ldots,\ n$, then there is an edge from x_i to x_k.
2. If $\partial h_j(x)/\partial x_i \neq 0$, $j = 1,\ 2,\ \ldots,\ m$, $i = 1,\ 2,\ \ldots,\ n$, then there is an edge from x_i to y_j.

The edges that connect the state and output variables are unidirectional. We define the total number of edges involved in a path as the length of the path based on graph theory [119]. Between x_i, $i = 1,\ 2,\ \ldots,\ n$ and y_j, $j = 1,\ 2,\ \ldots,\ m$, there may be more than one path, and the length of the shortest path that connects x_i and y_j is denoted as $\mathcal{L}_{ij}$. It is found that there is a fixed relationship between $\mathcal{L}_{ij}$ and the relative distance $\mathcal{D}_{ij}$ for most systems, i.e., $\mathcal{L}_{ij} - 1 = \mathcal{D}_{ij}$. This proposition is inspired by similar results with respect to controller configuration presented in [137, 118] and holds for most theoretical and practical systems.

Remark 31. *In this work, we assume that we know the number of subsystems that we would like to have. An important future research topic is to design a measure that can be used to optimally determine the number of subsystems, which is a challenging and practical issue. Basically, the number of subsystems should be no larger than the number of measured outputs. The measure should take into account the costs associated with communication, computation, estimation performance as well as process topology. In addition, if state estimates obtained by distributed state estimation systems are used for distributed output feedback control, the number of subsystems should be determined by taking both estimation and control into account.*

5.3 Application to a chemical process example

In this section, the developed distributed state estimation configuration methodology for nonlinear systems is applied to froth flotation units in Coal Handling and Preparation Plants (CHPPs).

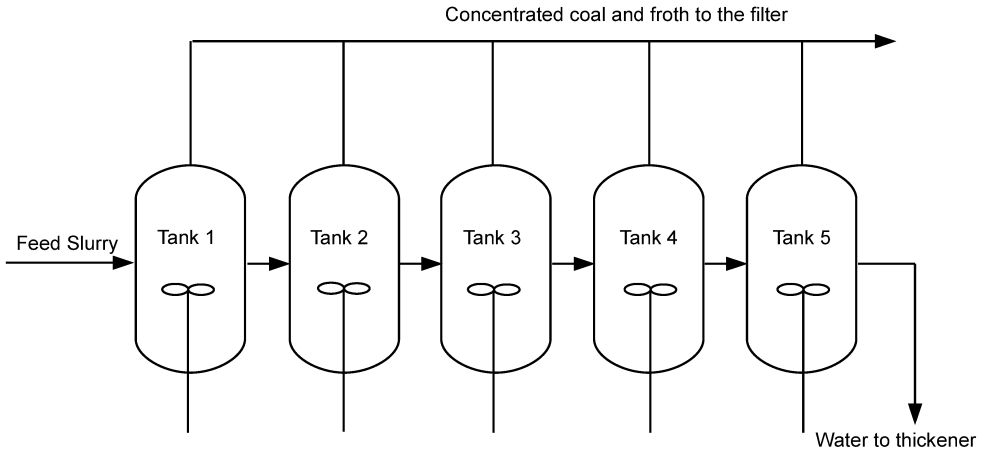

Figure 5.3: Schematic of the froth flotation units

The floatation units consist of five continuous stirred tank reactors (CSTR) with slurry as the input and froth and tailings as the outputs. The schematic of the process is shown in Figure 5.3.

The dynamics of the process can be characterized by the dynamics of the Solid (both coal and no-coal solids), Liquid (mainly water) and Ash (non-coal solids). It is note that the Ash is an individual component while it also constitutes a part of the Solid. Mass balances are completed for the 5 tanks for a total of 15 states. The equations describing the dynamics of stage (tank) i, where $i \in \mathbb{I} = \{1, \cdots, 5\}$ are as follows [84]:

$$\frac{dc_{st_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{st_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{st_i}) - r_i \tag{5.22a}$$

$$\frac{dc_{lt_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{lt_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{lt_i}) - \frac{\beta_i}{V_i} \tag{5.22b}$$

$$\frac{dc_{at_i}}{dt} = \frac{\dot{V}_{uf_{i-1}}}{V_i}(c_{at_{i-1}}) - \frac{\dot{V}_{uf_i}}{V_i}(c_{at_i}) - \frac{\dot{A}_i}{V_i} \tag{5.22c}$$

where c_{st_i} denotes the Solid concentration in tank i (kg/m^3), c_{lt_i} denotes the Liquid concentration (kg/m^3), c_{at_i} denotes the Ash concentration (kg/m^3), $\dot{V}_{uf_i}$ is the volumetric flow rate of the underflow from tank i (m^3/min) and V_i is the slurry volume in tank i (m^3). r_i is the rate of Solid removal from tank i as defined below, β_i is the mass flow rate of Liquid to

overflow from tank i (kg/min) and $\dot{A}_i$ is the mass flow rate of Ash to overflow from tank i (kg/min). c_{st_0}, c_{lt_0},c_{at_0} and $\dot{V}_{uf_0}$ denote the Solid concentration, Liquid concentration, Ash concentration and volumetric flow rate of the feed flow. On the right hand side of the above model, the first term of each equation represents the amount entering from the previous stage, the second term represents the amount leaving in the underflow and the third term represents the amount leaving in the overflow.

The rate of Solid removal r_i is given by [84]:

$$r_i = f_r k(c_{st_i} - c_\infty) \tag{5.23}$$

where k is the rate constant l/min, c_∞ is the equilibrium Solid concentration (kg/m^3) and f_r is the correction factor for industrial scale reactions.

The following equations are also defined for stage i, $i \in \mathbb{I}$:

$$\dot{V}_{uf_{i-1}} = \dot{V}_{uf_i} + \dot{V}_{of_i} \tag{5.24}$$

$$\dot{V}_{of_i} = \frac{\beta}{\rho_l} + \frac{r_i V_i}{\rho_c} \tag{5.25}$$

$$\dot{A}_i = x_{A_i}\left(\sum_{j=1}^{i} \dot{M}_{sof_j}\right) - \sum_{j=1}^{i-1} \dot{A}_j \tag{5.26}$$

$$R_i = \left(\frac{c_{sof_i}\dot{V}_{of_i}}{c_{sof_i}\dot{V}_{of_i} + c_{st_i}\dot{V}_{uf_i}}\right)(100 - R_{i-1}) + R_{i-1} \tag{5.27}$$

$$x_{A_i} = f(R_i) \tag{5.28}$$

where ρ_c is the density of coal (kg/m^3), ρ_l denotes the density of Liquid (kg/m^3), R_i is the cumulative Solid recovery at stage i, x_{A_i} is the cumulative mass fraction of Ash in the overflow Solid at stage i, $\dot{M}_{sof_i}$ is the mass flow rate of Solid in the overflow (kg/min), $\dot{V}_{of_i}$ is the volumetric flow rate of the overflow and c_{sof_i} denotes the concentration of Solid in the overflow (kg/m^3). $g(R_i)$ is an empirical function of R_i for a given frother and collector

loading from [84].

We assume that the Liquid concentration c_{lt_5} and Ash concentration c_{at_5} are the two measured outputs (which are the smallest set of measurements that ensures the observability of the entire system). The output vector of the process can be represented as

$$y = Cx \tag{5.29}$$

where

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.30}$$

We assume that the objective is to decompose the nonlinear process into two subsystems (i.e. $p = 2$) for distributed state estimation. Since there are two measured outputs, one output will be associated with one subsystem. Following the proposed procedure, we first determine the observability of the entire process. Since the observability matrix (7.3) is challenging to be calculated due to the high order of this system, we determine the observability by linearizing the model around the operating trajectories starting from a few different initial conditions in the operating range of this process that is determined following [84]. We verify that the entire system is locally observable.

Second, we identify the observable states of each measured output, i.e., $y_1 = c_{lt_5}$ and $y_2 = c_{at_5}$. It can be found that given $y_1 = c_{lt_5}$, states c_{st_i}, c_{lt_i}, $i \in \mathbb{I}$, can be estimated; given output $y_2 = c_{at_5}$, states c_{st_i}, c_{at_i}, $i \in \mathbb{I}$, can be estimated. Therefore, c_{lt_i}, $i \in \mathbb{I}$, should be estimated by the subsystem corresponding to y_1 and c_{at_i}, $i \in \mathbb{I}$, should be estimated by the subsystem corresponding to y_2. Note that c_{st_i}, $i \in \mathbb{I}$, can be estimated based either on y_1 or y_2. One feasible solution is to have two subsystems with state overlaps. Specifically, the set of states contained in subsystem 1 is $\{c_{st_i}, c_{lt_i}, i \in \mathbb{I}\}$ and the output is y_1; the set of states contained in subsystem 2 is $\{c_{st_i}, c_{at_i}, i \in \mathbb{I}\}$ and the output is y_2.

If we would like to avoid state overlapping in the two subsystems, then we should select an appropriate local subsystem to estimate c_{st_i}, $i \in \mathbb{I}$. To achieve this goal, we compare

Table 5.3: Relative degree between Solid concentration of each tank and each output measurement

	c_{st_1}	c_{st_2}	c_{st_3}	c_{st_4}	c_{st_5}
$j=1$	5	4	3	2	1
$j=2$	5	4	3	2	1

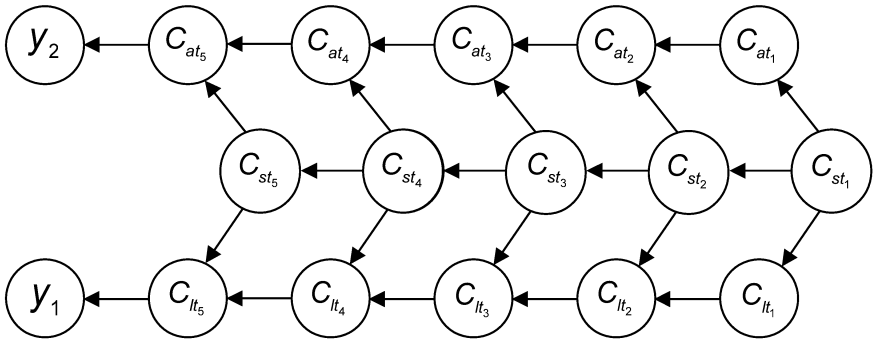

Figure 5.4: Direct graph for the froth flotation units

the closeness between each state c_{st_i}, $i \in \mathbb{I}$, and y_j, $j = \{1,\ 2\}$, which can be evaluated by comparing relative degrees of different state/ouput pairs. Direct graph is used to estimate the relative degrees and the graph is presented as in Figure 5.4. Based on Figure 5.4 and the property $\mathcal{D}_{ij} = \mathcal{L}_{ij} - 1$, the relative degrees are calculated and given in Table 5.3. We see that for any i, $i \in \mathbb{I}$, the relative degrees D_{i1} and D_{i2} are equal. Therefore, the optimal subsystem to estimate c_{st_i}, $i \in \mathbb{I}$, cannot be determined by comparing the relative degrees.

To further compare the closeness of state/output pairs, it is desirable to perform sensitivity analysis, with which we can further determine the closeness of any state/output pair by calculating the sensitivity indices $\bar{\mathcal{R}}_{i1}$ and $\bar{\mathcal{R}}_{i2}$ for state/output pairs (c_{st_i}, y_1) and (c_{st_i}, y_2), $i \in \mathbb{I}$, following (5.20).

We use the steady states of the system to calculate the sensitivity indices of different state/ouput pairs. Specifically, The steady state vector used for calculation is $x(t_s) = [70.1230; 945.5981; 24.3929; 49.5183; 961.4517; 21.9073; 36.7930; 971.2429; 20.9207; 29.1359; 977.1344; 20.6919; 24.6183; 980.6103; 18.9031]^T$. The obtained results are listed in Table 5.4. The table shows that for any i, $i \in \mathbb{I}$, $\bar{\mathcal{R}}_{i1} > \bar{\mathcal{R}}_{i2}$, which means state c_{st_i}, $i \in \mathbb{I}$, is more sensitive to the change in y_1 compared to y_2. Therefore, it is favourable to estimate state

Table 5.4: Sensitivity index between solid concentration of each tank and each output measurement

	$\bar{\mathcal{R}}_{5j}$	$\bar{\mathcal{R}}_{4j}$	$\bar{\mathcal{R}}_{3j}$	$\bar{\mathcal{R}}_{2j}$	$\bar{\mathcal{R}}_{1j}$
$j=1$	0.5577	1.0390	1.5019	1.9817	2.5435
$j=2$	0.2856	0.1717	0.4057	0.2490	0.4447

c_{st_i}, $i \in \mathbb{I}$, based on y_1.

The subsystems without state overlap are configured as follows: the set of states contained in subsystem 1 is $\{c_{st_i}, c_{lt_i}, i \in \mathbb{I}\}$ and the output is y_1; the set of states contained in subsystem 2 is $\{c_{at_i}, i \in \mathbb{I}\}$ and the output is y_2.

5.4 Summary

In this chapter, a subsystem decomposition procedure for distributed state estimation of nonlinear processes was proposed. In this procedure, techniques including observability test of the entire system, identification of observable states given each measured output, relative degree analysis and sensitivity analysis between output measurements and system states were adopted. Discussions on some considerations for special cases were carried out. Several examples were used to demonstrate the applicability of the proposed techniques/procedure.

Chapter 6

Subsystem decomposition and distributed moving horizon estimation of wastewater treatment plants

In this chapter, we propose a subsystem decomposition approach and a distributed moving horizon estimation (MHE) method for wastewater treatment plants. While the proposed approach is general, Benchmark Simulation Model No.1 for wastewater treatment plant is considered. The plant is decomposed into smaller subsystems based on the approach proposed in Chapter 5. Three subsystems are formed considering subsystem interaction and nonlinearity of the subsystems as shown in Section 6.2. In Section 6.3, an iterative distributed MHE scheme is proposed for the wastewater treatment plant. Innovation triggered evaluation of the local MHEs is used to reduce the computational complexity of the estimation scheme. In Section 6.4, extensive simulations are performed to illustrate the effectiveness and applicability of the proposed subsystem decomposition and distributed estimation methods. This chapter is a revised version of [131].

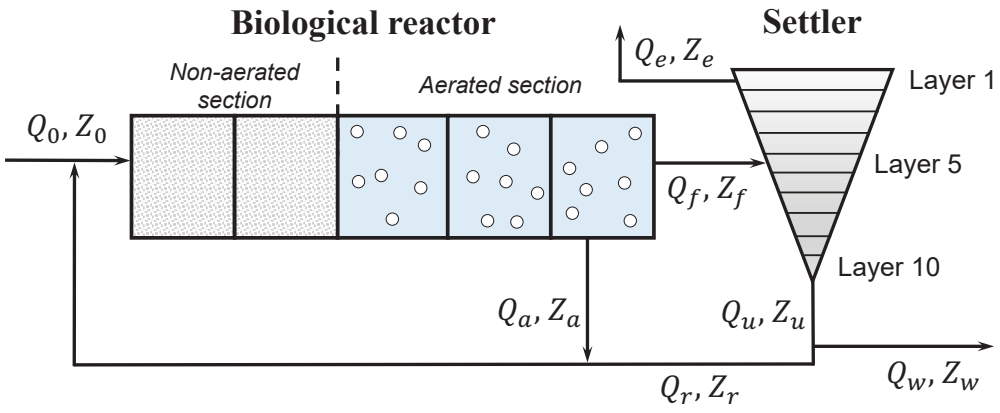

Figure 6.1: A schematic of the wastewater treatment plant

6.1 Preliminaries

6.1.1 Notation

The operator $L_f h$ represents the Lie derivative of function h with respect to function f, which is calculated following $L_f h(x) = \frac{\partial h}{\partial x} f(x)$. $L_f^r h$ represents the r-th order Lie derivative of function f, denoted by $L_f^r h(x) = L_f L_f^{r-1} h(x)$. $\{x(t_s)\}_{s=j}^{k}$ denotes a discrete-time sequence of x from time instant t_j to t_k, i.e., $x(t_j),\ x(t_{j+1}),\ \ldots,\ x(t_k)$. $x(t_s|t_k)$ denotes the prediction of $x(t_s)$ made at t_k. $\mathbb{K}_+$ represents a set which comprises all non-negative integers. The operator $\mathrm{diag}\,(v)$ represents a diagonal matrix where the elements of the vector v are on its main diagonal. Considering two matrices (or vectors) A and B of the same dimension, the operator $A \circ B$ represents the Hadamard product calculated element by element as $(A \circ B)_{i,j} = A_{i,j} \times B_{i,j}$. When A and B are identical, the product $A \circ A$ is called the Hadamard power of A and is denoted by $A^{\circ 2}$.

6.1.2 Model description of wastewater treatment plants

The WWTP based on Benchmark Simulation Model No.1 (BSM1) consists of a multi-chamber biological activated sludge reactor and a secondary settler. A schematic of the WWTP process is shown in Figure 7.3 [130]. The biological reactor comprises five inter-connected chambers: the first two anoxic chambers accounting for the non-aerated section

and the remaining three aerobic chambers constituting the aerated section. In the non-aerated section, pre-denitrification reactions where nitrate is converted into nitrogen take place; while in the aerated section, nitrification reactions where ammonium is oxidized into nitrate take place. In this way, biological nitrogen is removed based on pre-denitrification and nitrification in the reactor [130].

In this process, wastewater enters the first chamber of the biological reactor at concentration Z_0 and flow rate Q_0. A portion of the effluent of the last aerobic chamber of the reactor is fed into the settler at flow rate Q_f and concentration Z_f, while the rest portion (which is the inner recycle) flows back to the first chamber at flow rate Q_a. The secondary settler comprises 10 nonreactive layers, the 5-th of which is the feed layer. The outlets of the settler are discharged in three ways: (a) the overflow of the settler which contains purified water with concentration Z_e is removed continuously via the top layer of the settler at flow rate Q_e; (b) a portion of the underflow of the settler is recycled to the first chamber at flow rate Q_r and concentration Z_r; (c) the rest portion of the underflow is discharged through the 10-th layer at flow rate Q_w. In the process model, eight biological reactions are taken into account, and 13 major compounds are considered in these reaction processes. The concentrations of the 13 compounds in the five chambers constitute the state variables which are used for modelling the biological reactor. The definitions of the 13 state variables for each chamber are given in Table 6.1. The dynamic model of the secondary settler is established based on mass balances of the sludge considering solid flux due to gravity [133]. Specifically, the model of each layer of the settler contains 8 states; that is, S_O, S_{ALK}, S_{NH}, S_{NO}, S_S, S_I, S_{ND} and X. X denotes the concentration of suspended solids in a layer and is calculated as the summation of X_S, X_I, X_{B_A}, X_{B_H}, X_P and X_{ND} in the corresponding layer. The established process model is given in the Appendix II in this chapter. A more detailed description of the WWTP can be found in [130, 131].

The effluent quality (EQ) and the overall cost index (OCI) are two commonly used criteria in performance assessment of WWTPs. Detailed explanations of EQ and OCI can be found

Table 6.1: Process variables of the biological reactor of the WWTP

Definition	Notation	Unit
inert particulate organic matter	X_I	g COD $\cdot$ m^{-3}
inert soluble organic matter	S_I	g COD $\cdot$ m^{-3}
slowly biodegradable and soluble substrate	X_s	g COD $\cdot$ m^{-3}
readily biodegradable and soluble substrate	S_s	g COD $\cdot$ m^{-3}
particulate generated from decay of organisms	X_P	g COD $\cdot$ m^{-3}
particulate biodegradable organic nitrogen	X_{ND}	g N $\cdot$ m^{-3}
biomass of active autotrophs	X_{B_A}	g COD $\cdot$ m^{-3}
biomass of active heterotrophs	X_{B_H}	g COD $\cdot$ m^{-3}
nitrite nitrogen and nitrate	S_{NO}	g N $\cdot$ m^{-3}
free and saline ammonia	S_{NH}	g N $\cdot$ m^{-3}
dissolved oxygen	S_O	g (-COD) $\cdot$ m^{-3}
alkalinity	S_{ALK}	mol $\cdot$ m^{-3}
biodegradable and soluble organic nitrogen	S_{ND}	g N $\cdot$ m^{-3}

in [130, 149]. In implementation, it is very challenging to obtain the values of the two indices online due to the difficulty in obtaining real-time measurements of certain state that directly affect EQ and OCI. Alternatively, the two indices can be estimated using state estimates of the actual process states.

6.1.3 Measurement selection for state estimation

For WWTP described by the BSM No.1 model, we consider that there are 56 output variables that are measured online and can be used for state estimation. Specifically, in each chamber of the biological reactor, the concentrations of dissolved oxygen, free and saline ammonia (i.e., NH_3 and NH_4^+), nitrate and nitrate nitrogen, alkalinity, chemical oxygen demand (COD), filtered chemical oxygen demand (COD_f), biological oxygen demand (BOD) and the concentration of suspended solids can be measured online [123, 130]. In the settler, the states of the top layer and the bottom layer are measured. The measured output variables in each

Table 6.2: Measured output variables in the i-th chamber ($i = 1, \ldots, 5$) of the biological reactor

Measured output variable	Expression in the form of process states	Symbol
concentration of dissolved oxygen	S_O	$y_{c_{i,1}}$
concentration of nitrate and nitrite nitrogen	S_{NO}	$y_{c_{i,2}}$
concentration of NH_3 and NH_4^+	S_{NH}	$y_{c_{i,3}}$
concentration of alkalinity	S_{ALK}	$y_{c_{i,4}}$
COD	$S_S + S_I + X_S + X_I + X_{B_A} + X_{B_H}$	$y_{c_{i,5}}$
concentration of suspended solids	$X_S + X_I + X_{B_A} + X_{B_H} + X_P + X_{ND}$	$y_{c_{i,6}}$
COD_f	$S_S + S_I$	$y_{c_{i,7}}$
BOD	$S_S + X_S$	$y_{c_{i,8}}$

chamber of the biological reactor are described in Table 7.5, where $y_{c_{i,l}}$ is used to denote the l-th measurement of chamber i, $i = 1, \ldots, 5$, $l = 1, \ldots, 8$. The measured output variables in the top/bottom layer of the reactor are shown in Table 7.6. In this table, $s_{p,q}$, $p = 1, \ldots, 8$, denote the eight state variables S_O, S_{ALK}, S_{NH}, S_{NO}, S_S, S_I, S_{ND} and X in the q-th layer ($q = 1, \ldots, 10$) of the settler, respectively. Accordingly, $y_{s_{p,q}}$ denotes the measurement of state $s_{p,q}$ ($p = 1, \ldots, 8$) in the top layer (when $q = 1$) or the bottom layer (when $q = 10$).

6.1.4 Compact form of the WWTP model

Taking into account process and measurement noise, the WWTP model can be described by a compact nonlinear state-space model as follows:

$$\dot{x}(t) = f\left(x(t), u(t)\right) + w(t) \tag{6.1a}$$

$$y(t) = h\left(x(t)\right) + v(t) \tag{6.1b}$$

Table 6.3: Measured output variables in the top layer ($q = 1$) and bottom layer ($q = 10$) of the settler

Measured output variables	Expression in the form of process states	Symbol
concentration of dissolved oxygen	S_O	$y_{s_{1,q}}$
concentration of alkalinity	S_{ALK}	$y_{s_{2,q}}$
concentration of NH_3 and NH_4^+	S_{NH}	$y_{s_{3,q}}$
concentration of nitrate and nitrite nitrogen	S_{NO}	$y_{s_{4,q}}$
concentration of readily biodegradable and soluble substrate	S_S	$y_{s_{5,q}}$
concentration of inert soluble organic matter	S_I	$y_{s_{6,q}}$
concentration of biodegradable and soluble organic nitrogen	S_{ND}	$y_{s_{7,q}}$
concentration of suspended solids	$X_S + X_I + X_{B_A} + X_{B_H} + X_P + X_{ND}$	$y_{s_{8,q}}$

where $x \in \mathbb{R}^{145}$ is the state vector containing all the state variables of the process, $y \in \mathbb{R}^{56}$ represents the vector consisting of all the measured output variables, $u \in \mathbb{R}^3$ denotes the input vector consisting of both the manipulated inputs and the uncontrolled input to the WWTP plant, $w \in \mathbb{R}^{145}$ denotes the vector of additive disturbances to the process, and $v \in \mathbb{R}^{56}$ is the vector of measurement noise. Note that all the 56 measured output variables are linearly dependent on the system states, thus the output can be expressed as $h(x) = Cx$ where C is a 56 by 145 matrix.

6.1.5 Relative degree analysis

In this section, the definition of relative degree in the state estimation context is reviewed. The relative degree is a measure that will be used for the decomposition of the entire process into subsystems for state estimation. Relative degree analysis was first used for manipulated

input-controlled output pairs [135]. The relative degree (for an input-output pair) represents the number of integrations needed for an input to affect an output [136]. A smaller relative degree implies a more direct effect of the input on the considered output. Therefore, relative degree is a measure of topological (physical) closeness [135, 134, 137, 94]. The relative degree defined for state estimation is an analog of the relative degree for an input-output pair. Specifically, let us denote the i-th element in vector x by ξ_i, $i = 1, \ldots, n_x$ and denote the j-th element in vector y by η_j, $j = 1, \ldots, n_y$. Relative degree analysis evaluates physical closeness between a state variable ξ_i, $i = 1, \ldots, n_\xi$, and an output measurement η_j, $j = 1, \ldots, n_y$.

Define a square matrix F which is calculated as follows:

$$F = \frac{\partial f(x)}{\partial x} = \begin{bmatrix} \mathcal{F}_1 & \mathcal{F}_2 & \ldots & \mathcal{F}_{n_x} \end{bmatrix} \tag{6.2}$$

where $\mathcal{F}_1,\ \mathcal{F}_2, \cdots,\ \mathcal{F}_{n_x}$ are column vectors of the matrix F with dimension n_x.

As an analog of the relative degree in control [134, 135, 137], the relative degree $\mathcal{D}_{ij}$ for a state-output pair (ξ_i, η_j), $i = 1, \ldots, n_x$, $j = 1, \ldots, n_y$, is defined as below [61]:

$$\begin{cases} \mathcal{D}_{ij} = 0,\ if\ \frac{\partial h_j(x)}{\partial \xi_i} \neq 0 \\ \mathcal{D}_{ij} = 1,\ if \frac{\partial h_j(x)}{\partial \xi_i} \equiv 0\ \text{ and } L_{\mathcal{F}_i} h_j(x) \neq 0 \\ \mathcal{D}_{ij} = d_{ij},\ if\ \frac{\partial h_j(x)}{\partial \xi_i} \equiv 0 \text{ and } L_{\mathcal{F}_i} L_f^{k-1} h_j(x) \equiv 0 \text{ and } L_{\mathcal{F}_i} L_f^{d_{ij}-1} h_j(x) \neq 0 \end{cases} \tag{6.3}$$

where $h_j(x)$ represents the output function associated with the j-th measurement. d_{ij} is a positive integer, $k = 1, 2, \ldots, d_{ij} - 1$. The relative degree becomes $\mathcal{D}_{ij} = \infty$ if a finite integer d_{ij} does not exist for (6.3).

A smaller relative degree implies a stronger structural closeness of the considered state-measured output pair. Based on this measure, we can determine the closeness between each state variable x_i and each output variable y_j that are under consideration and then group the state and output variables which have comparatively small relative degrees.

6.2 Process decomposition and subsystem configuration

In this section, we develop an approach based on the method in [61] to decompose the entire WWTP process into subsystems. For each configured subsystem, a local MHE-based estimator will be designed within a distributed framework to estimate the states of the corresponding subsystem. In particular, we aim to decompose the original process model (6.1) into m subsystems in the following form:

$$\dot{x}_i(t) = f_i\left(x_i(t), u_i(t)\right) + \tilde{f}_i\left(X_i(t), u_i(t)\right) + w_i(t) \tag{6.4a}$$

$$y_i(t) = C_i x_i(t) + v_i(t) \tag{6.4b}$$

where $x_i \in \mathbb{R}^{n_{x_i}}$ is the vector of the i-th subsystem state, $u_i \in \mathbb{R}^{n_{u_i}}$ represents the input vector to the i-th subsystem, $w_i \in \mathbb{R}^{n_{w_i}}$ is the vector of disturbances of the i-th subsystem, $y_i \in \mathbb{R}^{n_{y_i}}$ is the vector of the output measurements of subsystem i, $v_i \in \mathbb{R}^{n_{v_i}}$ is the measurement noise vector of the i-th subsystem, and $X_i \in \mathbb{R}^{n_{X_i}}$ represents a vector containing the subsystem states which have direct effects on the dynamics of x_i, $i \in \{1, \ldots, m\}$. Accordingly, we introduce $\mathbb{I}_i$, $i \in \{1, \ldots, m\}$, to denote the set containing the indices of the subsystems, of which the states are explicitly involved in X_i. For instance, if the dynamics of subsystem i is affected by the dynamics of subsystem 1 and subsystem 3, then we have $X_i = \left[x_1^{\mathrm{T}}\ x_3^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\mathbb{I}_i = \{1,\ 3\}$. Furthermore, if the set $\mathbb{I}_i$ does not contain i, then the interaction of subsystem i is considered as "additive interaction". Otherwise if $\mathbb{I}_i$ also contains i, we call this type of interaction "multiplicative interaction". These two types of interaction also apply to the characterization of interaction between groups which will be discussed later.

6.2.1 Observability of the WWTP

First, we need to verify the observability of the entire process. For nonlinear processes, one way to test the observability is to check whether the full-rank condition is satisfied for the model (6.1) by the nonlinear observability matrix $Q(x)$ defined as follows [30]:

$$Q(x) = \frac{d}{dx} \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{bmatrix} \tag{6.5}$$

However, due to the high dimension of the process, higher order Lie derivatives are very challenging to calculate and the observability matrix can be ill-conditioned. Instead, we take advantage of the measure of linear observability gramian to check the local observability of the process. Specifically, the entire process (6.1) is linearized along a typical state trajectory and the observability gramian is calculated at each sampled point along this trajectory. By examining the rank of the gramian at each point, it can be confirmed that the entire process (6.1) is locally observable along the considered trajectory given the 56 measured outputs.

6.2.2 Process decomposition and subsystem configuration

In this section, we consider the process decomposition and subsystem configuration task from a distributed estimation point of view. A flowchart of the procedure used to fulfill this task is given in Figure 6.2.

First, we decompose the entire process into small groups (not the finally configured subsystems for distributed estimation) by considering the topology of the process. This subtask is accomplished via two key steps: (1) to divide the measurements of the entire process into groups by considering the physical topology; (2) to perform relative degree analysis to assign each state to a measurement group by pairing the state with appropriate

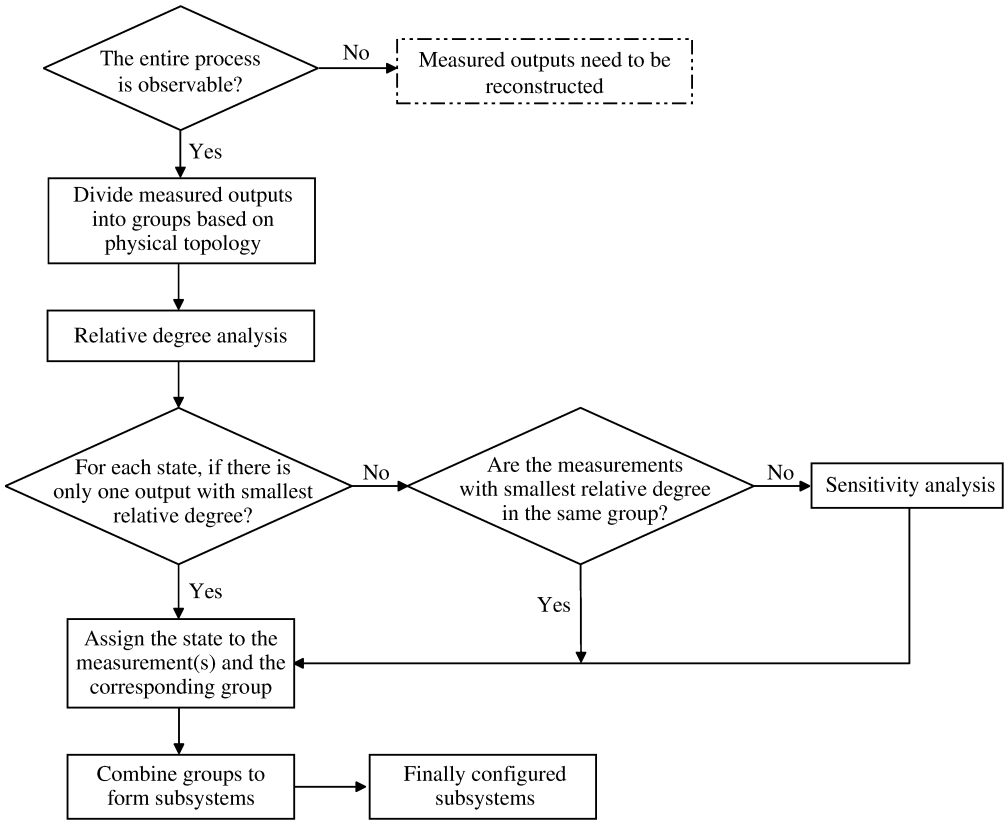

Figure 6.2: A flowchart of the procedure for process decomposition and subsystem configuration

measured output(s).

Within a distributed framework, the measurements of a subsystem need to be sent to the corresponding estimator every sampling time. If the measurements of one subsystem are from distant physical operating units, it will lead to additional communication costs. Based on this consideration, we first group the 56 measurements into small groups according to their physical location. Specifically, we divide the measurements of the reactor into five groups, and each group contains the eight available measurements in a chamber in the reactor. In terms of the settler, one may consider decomposing its measurements into groups according to the ten layers. However, this is infeasible, since no measurement is available in each of the middle layers (i.e., 2nd layer to 9th layer). Alternatively, the measurements of the settler are grouped into eight groups according to the eight types of measurements.

Now, we are in a position to allocate the process states to the 13 measurement groups.

Table 6.4: Relative degree analysis for the states of chamber i, $i = 1, \ldots, 5$

States	Smallest relative degree to all measurements	Candidate measurement(s) to be paired with
$S_{I,i}$	0	$y_{c_{i,5}}$, $y_{c_{i,7}}$
$S_{S,i}$	0	$y_{c_{i,5}}$, $y_{c_{i,7}}$, $y_{c_{i,8}}$
$X_{I,i}$	0	$y_{c_{i,5}}$, $y_{c_{i,6}}$
$X_{S,i}$	0	$y_{c_{i,5}}$, $y_{c_{i,6}}$, $y_{c_{i,8}}$
$X_{B_{H,i}}$	0	$y_{c_{i,5}}$, $y_{c_{i,6}}$
$X_{B_{A,i}}$	0	$y_{c_{i,5}}$, $y_{c_{i,6}}$
$X_{P,i}$	0	$y_{c_{i,6}}$
$S_{O,i}$	0	$y_{c_{i,1}}$
$S_{NO,i}$	0	$y_{c_{i,2}}$
$S_{NH,i}$	0	$y_{c_{i,3}}$
$S_{ND,i}$	1	$y_{c_{i,3}}$, $y_{c_{i,4}}$
$X_{ND,i}$	0	$y_{c_{i,6}}$
$S_{ALK,i}$	0	$y_{c_{i,4}}$

We explore the relative degrees between states and measured outputs following Eq.(6.3). Every state is paired with a measured output that has the smallest relative degree from the state and is assigned to the corresponding group of the measured output. The corresponding results are given in Table 6.4 and Table 6.5. From the results in Table 6.4 and Table 6.5, the state-output pairing for certain system states can have more than one options. We may further carry out sensitivity analysis [61] to identify the most appropriate measurement for each state to pair with. However, this is not necessary for the WWTP. Specifically, it is seen that for each state, the candidate measurements for pairing (if there is more than one) are from the same measurement group. This implies that we are already able to assign each state to a measurement group based on the results in Table 6.4 and Table 6.5.

Consequently, we obtain 13 groups of state and measurement variables. Each of the first five groups contains 13 system states and 8 measurements that are from one chamber

Table 6.5: Relative degree analysis for the states of the settler

States	Smallest relative degree to all measurements	Candidate measurement to be paired with
$s_{p,1}$	0	$y_{s_{p,1}}$
$s_{p,2}$	1	$y_{s_{p,1}}$
$s_{p,3}$	2	$y_{s_{p,1}}$
$s_{p,4}$	3	$y_{s_{p,1}}$
$s_{p,5}$	4	$y_{s_{p,1}}$
$s_{p,6}$	4	$y_{s_{p,10}}$
$s_{p,7}$	3	$y_{s_{p,10}}$
$s_{p,8}$	2	$y_{s_{p,10}}$
$s_{p,9}$	1	$y_{s_{p,10}}$
$s_{p,10}$	0	$y_{s_{p,10}}$

of the reactor. Each remaining group contains 10 states and 2 measured outputs which are with respect to one material concentration in the different layers of the settler. We use the measure of linear observability gramian to check the observability of each group along a typical trajectory. The states involved in each group are observable based on the measurements from the same group by treating interacting dynamics as known inputs. The 13 groups serve as the basis for the subsystem configuration for distributed state estimation.

Subsequently, we combine groups to configure subsystems. The number of subsystems (equivalently the number of local estimators in distributed state estimation) is user-specified, and may vary based on different requirements on computation complexity, communication capability, maintenance cost, etc. In this design, the number of subsystems for distributed state estimation is selected to be three. For subsystem configuration based on the formed groups, the following aspects are considered: (I) for subsystem i in the form of (6.4a), it is favorable to make X_i be free of x_i; that is, on the right-hand-side of Eq.(6.4a), $i \in \{1, \ldots, m\}$, x_i only exists in f_i; (II) the groups of which the states have more significant coupling are expected to be combined; (III) it is better to combine the groups from the same physical

unit since this may significantly reduce information exchange; (IV) the numbers of states in the configured subsystems with nonlinear local dynamics are expected to be similar, such that the computation complexity of the corresponding estimators will not be much different; (V) it is favorable to isolate the linear dynamics (if there exist) from the nonlinear dynamics in one subsystem such that linear approaches can be applied to this subsystem; (VI) it is expected that the configured subsystems do not have overlapping states which can lead to additional computational cost of local estimators.

We first explore how the state from different groups interact with each other. Based on (I) and (II), groups where the state dynamics have multiplicative interaction terms or have obviously strong interaction (qualitatively) will be arranged in the same subsystem. We examine the ordinary differential equations (ODEs) describing the state dynamics in each group. It is found that there exist multiplicative interaction terms in the ODEs describing the dynamics of suspended solids (i.e., X) in the ten layers of the settler. Specifically, these multiplicative interaction terms contain X in the ten layers (which are in one group) and the states of the fifth chamber (which are from another group) due to the interconnection between the reactor and the settler via material flows. Therefore, these two groups are first grouped together. Also, because of the inner recirculation from the fifth to the first chamber of the reactor, the state dynamics of the first chamber (accounted for by one group) are significantly affected by the dynamics of the fifth chamber. Therefore, the group containing the states of the first chamber is further combined with the above two groups. The first subsystem containing 36 states and 18 output measurements is configured.

Next, let us consider criterion (III) for subsystem configuration. The second to the fourth chambers of the reactor are within the same physical structure (i.e., the biological reactor) and are coupled in a cascade manner. Consequently, we decide to combine the three groups accounting for these three chambers to form subsystem 2. This configured subsystem also has nonlinear local dynamics of 39 state variables, which is similar to subsystem 1, such that the other criteria for subsystem configuration are not violated.

Table 6.6: The states in each configured subsystem of the process

Subsystem 1	Chamber 1, Chamber 5, Material concentration of X in the ten layers
Subsystem 2	Chamber 2, Chamber 3, Chamber 4
Subsystem 3	Concentrations of soluble materials S_I, S_S, S_O, S_{NO}, S_{NH}, S_{ND} and S_{ALK} in the ten layers

Now, there are seven groups remaining for subsystem configuration. Each of them incorporates 10 states with respect to one type of soluble material in the settler. We check the ODEs describing the dynamics of the seven units. For each of the seven groups, its dynamics are only affected by the dynamics of the fifth chamber (already arranged in subsystem 1) through one state which is the corresponding concentration in the feed layer. It is found that this interaction is additive interaction. Moreover, by treating the interacting dynamics from the fifth chamber as external inputs, the local dynamics of each group are linear. This fact makes it possible to configure a subsystem with only linear local dynamics. Therefore, the seven groups with respect to the seven soluble concentrations in the ten layers of the settler are made in one configuration as subsystem 3, which has linear local dynamics. The state variables of the WWTP involved in each subsystem are described in Table 6.6.

Remark 32. *It is also possible to design a distributed state estimation scheme which consists of 13 local estimators by directly treating each of the 13 groups as one subsystem. However, despite substantially increased computational efficiency, the implementation and maintenance of a distributed scheme containing 13 local estimators could be costly. Therefore, a smaller number of subsystems is determined (which is 3 in the proposed design) such that a balance is achieved among computation, communication, maintenance, etc.*

Remark 33. *Configuring a subsystem containing only linear dynamics can simplify the design procedure and reduce the computational complexity of the corresponding subsystem estimator. Within a distributed MHE framework, a linear MHE approach can be used to*

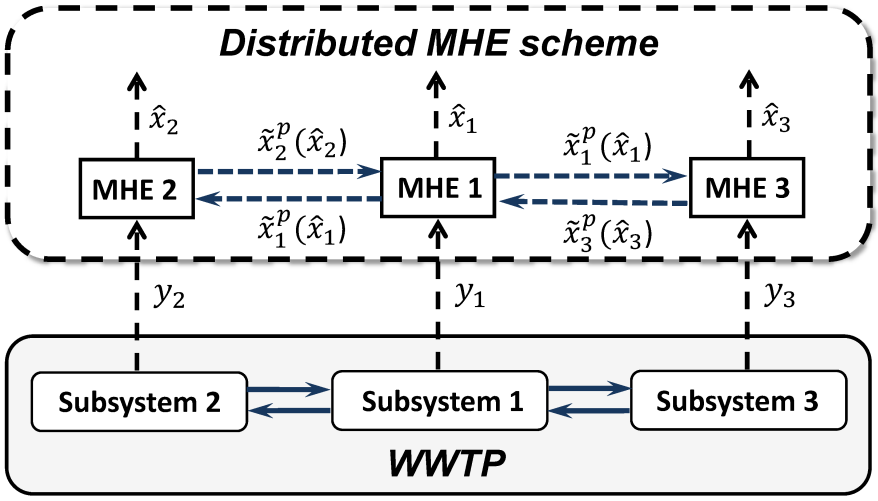

Figure 6.3: A diagram of the distributed MHE scheme

design a local estimator for the configured subsystem with linear local dynamics to reduce the computational complexity. Note that to configure a subsystem of linear local dynamics, only the local dynamics (i.e., $f_i(x_i(t), u_i(t))$ in (6.4)) are required to be linear in order to design a linear MHE local estimator. This point will be made clear by the MHE-based estimator designs in the next section.

Remark 34. *There exist several other solutions to the subsystem configuration. For example, the seven decomposed groups in subsystem 3 could be grouped to form two subsystems, such that in total four subsystems can be configured. However, considering the linear local dynamics of the seven units, a linear MHE algorithm can be applied to properly handle the computational complexity, such that one-subsystem configuration adopted in this work could be more favorable from an application perspective.*

6.3 Distributed MHE design for WWTPs

In this section, we take advantage of the configured subsystems and propose an iterative distributed state estimation scheme for the WWTP. A schematic diagram that describes the distributed MHE scheme is given in Figure 6.3. Within the proposed distributed framework,

three local estimators are designed for the three subsystems: a nonlinear MHE-based estimator is designed for each of subsystem 1 and subsystem 2, while an estimator is developed based on linear MHE for subsystem 3. Each MHE estimator retrieves the output measurements of the associated subsystem at every sampling instant $t_k := t_0 + k\Delta$ with k being a non-negative integer, t_0 denoting the initial time instant and Δ being the size of the interval between two consecutive sampling instants. In the meantime, the MHE estimators are required to exchange information in terms of subsystem state estimates through networked communication at each $t_{k\geq 0}$. Considering the interaction among the configured subsystems, we have $\mathbb{I}_2 = \mathbb{I}_3 = \{1\}$ and $\mathbb{I}_1 = \{2, 3\}$. Therefore, the MHE estimator for subsystem 1 is required to communicate with the MHE estimators for the other two subsystems bidirectionally, while the estimators designed for subsystem 1 and subsystem 3 do not exchange information with each other. At every sampling instant, each local estimator is evaluated to provide subsystem state estimates based on the information collected from its associated subsystem as well as the subsystem estimators for its interacting subsystems. To achieve faster convergence of the estimation error dynamics, we require that each local estimator is evaluated iteratively several times at each sampling instant until the innovation for each subsystem enters a certain sufficiently small region.

6.3.1 Iterative distributed state estimation algorithm

In this section, we present an algorithm to demonstrate the key implementation steps for the proposed distributed state estimation scheme. We first introduce the following definition that will be used in determining the triggering conditions for iterative evaluation of the distributed MHE scheme.

Definition 3. *The normalized Euclidean norm of the deviation between two arbitrary vectors $a \in \mathbb{R}^{n_a}$ and $\hat{a} \in \mathbb{R}^{n_a}$ at each time t_k (which is denoted by $|\varepsilon(a(t_k), \hat{a}(t_k))|$) is defined as*

follows:

$$\left|\varepsilon(a(t_k),\hat{a}(t_k))\right| = \sqrt{\sum_{i\in n_a}\left(\frac{a_i(t_k)-\hat{a}_i(t_k)}{\max\limits_{t_n\in[t_0,t_\infty)}|a_i(t_n)-\hat{a}_i(t_n)|}\right)^2} \tag{6.6}$$

where a_i ($\hat{a}_i$) denotes the i-th element of the vector a ($\hat{a}$), $i = 1,\ldots,n_a$.

Algorithm 3. *Innovation-triggered iterative distributed MHE for WWTPs*

1. *At time instant t_0, each MHE estimator i, $i \in \mathbb{I}$, is initialized with an initial guess $\tilde{x}_i(t_0)$, and also requests the initial guess $\tilde{x}_l(t_0)$ from each MHE estimator l, $l \in \mathbb{I}_i$.*

2. *At each sampling time $t_{k\geq 0}$, set $p = 1$, and do the following steps:*

   2.1. *MHE estimator i receives the output measurement $y_i(t_k)$ from the corresponding subsystem i, $i \in \mathbb{I}$.*

   2.2. *Based on $y_i(t_k)$ and the latest state estimate sequences for each interacting subsystem l, $l \in \mathbb{I}_i$, each MHE estimator i is evaluated and gives a state estimate sequence for subsystem i, $i \in \mathbb{I}$.*

   2.3. **If** $\left|\varepsilon(y_i(t_k), C_i\tilde{x}_i^p(t_k|t_k))\right| \leq T_i$ *holds* $\forall i$, $i \in \mathbb{I}$, **then**

   go to step 2.4.

   Else, *do:*

   if $p > 1$ *and* $\left|\varepsilon\left(\tilde{x}_i^p(t_k|t_k), \tilde{x}_i^{p-1}(t_k|t_k)\right)\right| \leq \delta_i$, $\forall i \in \mathbb{I}$, **then**

   go to step 2.4.

   else, *do:*

   - *MHE estimator i, $i \in \mathbb{I}$, requests and receives the latest state estimate sequence*

     generated at t_k from each MHE estimator l, and updates the latest state estimate

     sequence for subsystem l, $l \in \mathbb{I}_i$.

   - *Set $p = p + 1$, and go to Step 2.2.*

2.4. Set $\hat{x}_i(t_k) = \tilde{x}_i^p(t_k|t_k)$ as the estimate of each subsystem i, $i \in \mathbb{I}$, for time instant t_k.

2.5. MHE estimator i, $i \in \mathbb{I}$, requests and receives the state estimate sequence of each subsystem l updated in the p-th iteration at time instant t_k by the l-th estimator, $l \in \mathbb{I}_i$. Set $k = k + 1$ and go to step 2.1.

In Algorithm 3, T_i represents the threshold on the normalized Euclidean norm of the innovation for each MHE estimator i, $i \in \mathbb{I}$. δ_i denotes the threshold on the normalized Euclidean norm of the deviation between two subsystem estimates obtained at two consecutive iteration steps given by the i-th MHE estimator, $i \in \mathbb{I}$. At each sampling time, if the condition on the subsystem innovations (in Step 2.3) is not satisfied, the MHE estimators for the subsystems will be evaluated several times iteratively. The number of iterations is not fixed at each sampling time. It is dependent on the innovations given by the subsystem MHE estimators as well as the deviation between the state estimates obtained at the most recent two iteration steps for each subsystem.

The proposed algorithm helps achieve a balance between fast convergence of the estimation error dynamics and efficient computation. Specifically, after the estimation error converges and the innovation for the subsystems becomes sufficiently small, the distributed MHE system can switch to a non-iterative mode where the estimators for the subsystems communicate with each other and are evaluated only once at one sampling time.

Remark 35. *The calculation of $|\varepsilon(\tilde{x}_i^p(t_k|t_k), \tilde{x}_i^{p-1}(t_k|t_k))|$ in Step 2.3 of Algorithm 3 requires the values of $\max_{t_n \in [t_0, t_\infty)} \left|\tilde{x}_{i,l}^p(t_n|t_n) - \tilde{x}_{i,l}^{p-1}(t_n|t_n)\right|$ where $\tilde{x}_{i,l}$ represents the l-th element of vector $\tilde{x}_i$, for $l = 1, \ldots, n_{x_i}$. In implementation, these values can be calculated based on the evaluation of the proposed distributed MHE with a fixed number of iteration steps $p = 2$ for a sufficiently long operation time. We choose to calculate $|\varepsilon(y_i(t_k), C_i\tilde{x}_i^p(t_k|t_k))|$, $i \in \mathbb{I}$, in a similar way with $p = 1$.*

Remark 36. *Note that the values of T_i and δ_i, $i \in \mathbb{I}$, are user-specified values and should*

be carefully selected through extensive simulations and analysis before real implementation. While smaller values on T_i and δ_i may accelerate the convergence speed of the estimation error dynamics for the entire process, they may lead to much increased computational load at the initial stage when the state estimates have not well tracked the actual dynamics.

Remark 37. *Note that the implementation algorithm can be reduced to a non-iterative distributed MHE strategy by disabling Step 2.3. In implementation, we may also set an upper bound on the iteration step p to avoid endless loops, although this is not encountered in the simulations for the WWTP.*

6.3.2 Design of the distributed MHE estimators

In this section, we present the design of the MHE estimators for the three configured subsystems of the WWTP. The estimators for subsystem 1 and subsystem 2 are based on nonlinear MHE, while the estimator for subsystem 3 is designed using a linear MHE approach. Note that the distributed MHE scheme proposed in this work is adapted from the distributed MHE designs for nonlinear processes reported in [15, 42].

6.3.2.1 MHE estimators for subsystem 1 and subsystem 2

At the p-th iteration step of time instant t_k, a nonlinear MHE estimator is designed for each subsystem i, $i = 1, 2$, as follows:

$$\min_{\left\{\tilde{x}_i^p(t_d|t_k)\right\}_{d=k-N_i}^{k}} \sum_{d=k-N_i}^{k-1} |\tilde{w}_i^p(t_d)|^2_{Q_i^{-1}} + \sum_{d=k-N_i}^{k} |\tilde{v}_i^p(t_d)|^2_{R_i^{-1}} \tag{6.7a}$$

$$\textbf{s.t.} \quad \dot{\tilde{x}}_i^p(\tau|t_k) = f_i\left(\tilde{x}_i^p(\tau|t_k), u_i(t_d)\right) + \tilde{f}_i\left(\tilde{X}_i^{p-1}(t_d), u_i(t_d)\right) + \tilde{w}_i^p(t_d),$$

$$\tau \in [t_d, t_{d+1}],\ d = k - N_i, \ldots, k-1 \tag{6.7b}$$

$$\tilde{v}_i^p(t_d) = y_i(t_d) - C_i\tilde{x}_i^p(t_d), \qquad d = k - N_i, \ldots, k \tag{6.7c}$$

$$\tilde{x}_i^p(t_d) \in \mathbb{X}_i,\ \tilde{v}_i^p(t_d) \in \mathbb{V}_i, \qquad d = k - N_i, \ldots, k \tag{6.7d}$$

$$\tilde{w}_i^p(t_d) \in \mathbb{W}_i, \qquad d = k - N_i, \ldots, k-1 \tag{6.7e}$$

where $\tilde{x}_i^p$, $\tilde{w}_i^p$ and $\tilde{v}_i^p$ represent the estimates of x_i, w_i, and v_i within the current horizon obtained in the p-th iteration of time t_k, respectively, Q_i and R_i are positive definite matrices representing the covariances of subsystem disturbances and noise, N_i denotes the length of the estimation horizon of MHE estimator i, $i \in \mathbb{I}$, $\mathbb{X}_i$, $\mathbb{W}_i$ and $\mathbb{V}_i$ are compact sets that bound the subsystem state, disturbance and measurement noise of subsystem i such that $x_i \in \mathbb{X}_i$, $w_i \in \mathbb{W}_i$, and $v_i \in \mathbb{V}_i$. In Eq.(6.7b), vector $\tilde{X}_i^{p-1}(t_d)$ is an estimate of $X_i(t_d)$ and contains the latest estimates of subsystem l, $l \in \mathbb{I}_i$, available to MHE estimator i, $i \in \mathbb{I}$, for $d = k - N_i, \ldots, k-1$ according to Algorithm 3.

In the optimization problem (6.7), Eq.(6.7a) represents the cost function for the MHE estimator i ($i = 1, 2$) to be minimized treating $\{\tilde{x}_i^p(t_d|t_k)\}_{d=k-N_i}^{k}$ as the decision variables, Eq.(6.7b) and Eq.(6.7c) are the subsystem model constraints, while Eq.(6.7d) and (6.7e) take into account the constraints on system states and disturbances. The last point $\tilde{x}_i^p(t_k|t_k)$ in the optimal state estimate sequence given by the last iteration at t_k is adopted as the estimate of $x_i(t_k)$ (denoted by $\hat{x}_i(t_k)$).

6.3.2.2 MHE estimator for subsystem 3

Subsystem 3 has linear local dynamics. The local dynamics described by the vector field f_3 in Eq.(6.4) can be expressed by a linear form as $f_3\left(\tilde{x}_3, u_3\right) = A_3 x_3$ where A_3 is a real-valued square matrix of dimension 70. Considering the linear local dynamics, we design an estimator to estimate the states of subsystem 3 based on a linear MHE approach [138]. At the p-th iteration step of time instant t_k, the MHE estimator for each subsystem 3 is designed as follows:

$$\min_{\left\{\tilde{x}_3^p(t_d|t_k)\right\}_{d=k-N_3}^{k}} \left\{ \begin{array}{l} \sum\limits_{d=k-N_3}^{k-1} |\tilde{w}_3^p(t_d)|^2_{Q_3^{-1}} + \sum\limits_{d=k-N_3}^{k} |\tilde{v}_3^p(t_d)|^2_{R_3^{-1}} \\ \qquad +|(\hat{x}_3(t_{k-N_3}) - \tilde{x}_3^p(t_{k-N_3}|t_k))|^2_{\Pi_3^{-1}(t_k|t_k)} \end{array} \right\} \tag{6.8a}$$

$$\textbf{s.t.} \quad \dot{\tilde{x}}_3^p(t|t_k) = A_3\tilde{x}_3^p(t|t_k) + \tilde{f}_3\left(\tilde{X}_3^{p-1}(t_d), u_3(t_d)\right) + \tilde{w}_3^p(t_d),$$

$$t \in [t_d, t_{d+1}],\ d = k - N_3, \ldots, k-1 \tag{6.8b}$$

$$\tilde{v}_3^p(t_d) = y_3(t_d) - C_3\tilde{x}_i^p(t_d),\ d = k - N_3, \ldots, k \tag{6.8c}$$

$$\tilde{x}_3^p(t_d) \in \mathbb{X}_3,\ \tilde{v}_3^p(t_d) \in \mathbb{V}_3, \qquad d = k - N_3, \ldots, k \tag{6.8d}$$

$$\tilde{w}_3^p(t_d) \in \mathbb{W}_3, \qquad d = k - N_3, \ldots, k-1 \tag{6.8e}$$

where $\tilde{x}_3^p$, $\tilde{w}_3^p$ and $\tilde{v}_3^p$, respectively, are the estimate sequences of x_3, w_3, and v_3 within the current horizon obtained at the p-th iteration of each sampling time, Q_3 and R_3 are the covariance matrices of the subsystem disturbances and noise, N_3 is the length of the estimation horizon for the linear MHE, $\mathbb{X}_3$, $\mathbb{W}_3$ and $\mathbb{V}_3$ are compact sets that bound the subsystem state, disturbance and measurement noise of subsystem i such that $x_i \in \mathbb{X}_i$, $w_i \in \mathbb{W}_i$, and $v_i \in \mathbb{V}_i$., $\Pi_3(t_k|t_k)$ is an invertible matrix accounting for the arrival cost of the linear MHE. The design of Π_3 will be clarified later. In Eq.(6.8b), $\tilde{X}_3^p$ is an estimate of X_3 containing the latest state estimate of subsystem 1 which is available to MHE estimator 3.

In the optimization problem (6.8), Eq.(6.8a) is the cost function for the MHE estimator 3

to be minimized with $\{\tilde{x}_3^p(t_d|t_k)\}_{d=k-N_3}^{k}$ being the decision variables, Eq.(6.8b) and Eq.(6.8c) account for the model of subsystem 3. Eq.(6.8d) and Eq.(6.8e) are the constraints on the subsystem states, disturbances and noise. At each sampling time $t_{k\geq 0}$, the optimization problem (6.7) is also solved p time(s) iteratively following Algorithm 3. The last point in the optimal estimate sequence $\{\tilde{x}_3^p(t_d|t_k)\}_{d=k-N_3}^{k}$ given in the last iteration is adopted as the estimate of the dynamics of subsystem 3 at t_k.

The matrix Π_3 involved in the cost function (6.8a) is designed as the *a posteriori* error covariance matrix of continuous-discrete time Kalman filtering. Specifically, $\Pi_3(t_k|t_k)$ is updated for each sampling time following:

$$\Pi_3(t_k|t_{k-1}) = \phi\left(t_k, t_{k-1}\right)\Pi_3(t_{k-1}|t_{k-1})\phi^{\mathrm{T}}\left(t_k, t_{k-1}\right) + \int_{t_{k-1}}^{t_k} \phi\left(t_k, \tau\right) Q_3 \phi^{\mathrm{T}}\left(t_k, \tau\right) d\tau \quad (6.9a)$$

$$\Pi_3\left(t_k|t_k\right) = \Pi_3\left(t_k|t_{k-1}\right) - \Pi_3\left(t_k|t_{k-1}\right) C_3^{\mathrm{T}} (C_3 \Pi_3\left(t_k|t_{k-1}\right) C_3^{\mathrm{T}} + R_3)^{-1} C_3 \Pi_3\left(t_k|t_{k-1}\right) \quad (6.9b)$$

s.t. $\Pi_3(t_0|t_0) = \bar{\Pi}_3$

where $\bar{\Pi}_3$ is an initial guess of the error covariance matrix, and $\phi\left(t, t_{k-1}\right)$ denotes the state transition matrix of subsystem 3 and is calculated as $\phi(t|t_{k-1}) = e^{A_3(t-t_{k-1})}$.

It is worth mentioning that at t_k, $0 \leq k < N_i$, each linear/nonlinear MHE estimator for subsystem i, $i \in \mathbb{I}$, is essentially a full-information estimator which gives an optimal state estimate sequence from the initial time t_0 to current time t_k instead of an estimate sequence of size N_i. In this case, the optimization problem (6.7) is solved iteratively using the information about subsystem state estimates and measurements from the t_0 to t_k. At each sampling time $t_{k\geq N_i}$, for MHE estimator i, $i \in \mathbb{I}$, the optimization problem (6.7) is solved p time(s) iteratively within a moving window of a fixed horizon N_i following Algorithm 3.

Remark 38. *We note that the calculation of (6.9) does not require any information from the process. Therefore, $\bar{\Pi}_3(t_k|t_k)$, $k \geq 0$ can be obtained in an off-line manner such that the computational load of the linear MHE for subsystem 3 can be reduced.*

6.4 Simulation results

In this section, we apply the proposed distributed MHE scheme to the WWTP considering different weather conditions.

6.4.1 Simulation settings

The values of the process parameters are determined following [130]. The data containing the information about the wastewater feed to the process for dry, rainy and stormy weather conditions were obtained from the International Water Association website [139]. In particular, the flow rate of the wastewater feed Q_0 and the corresponding concentration Z_0 vary and are dependent on the current weather condition.

The measurements are sampled synchronously at each sampling instant $\{t_{k\geq 0}\}$ where $t_k = t_0 + k\Delta_s$ with $t_0 = 0$ being the initial time instant, $\Delta_s = 15$min being the sampling period and $k \in \mathbb{K}_+$. We consider that at each sampling instant t_k, each measurement is instantly available to the corresponding subsystem estimator. Let x_{is} denote the vector of steady-state values corresponding to x_i, $i = 1, 2, 3$, which were calculated based on a 100-day open-loop operation [149]. The initial conditions of the three subsystems of the WWTP are selected to be $1.18x_{1s}$, $1.15x_{2s}$ and $1.29x_{3s}$, respectively.

Additive process disturbances to the WWTP and random noise contaminating the output measurements are taken into account. Specifically, the disturbance vector w_i which affects the dynamics of subsystem i is generated following normal distribution with zero mean and standard deviations $0.06x_{is}$, and is constrained by $-0.12x_{is}$ and $0.12x_{is}$. Moreover, random noise to the vector of output measurements y_i is assumed to be Gaussian white noise with zero mean and standard deviations $0.05y_{is}$ where the elements in $y_{i,s}$ take the steady-state values of y_i, $i = 1, 2, 3$. We also take into account hard constraints on the process dynamics; that is, the estimates of the system states which denote different concentrations should not be negative.

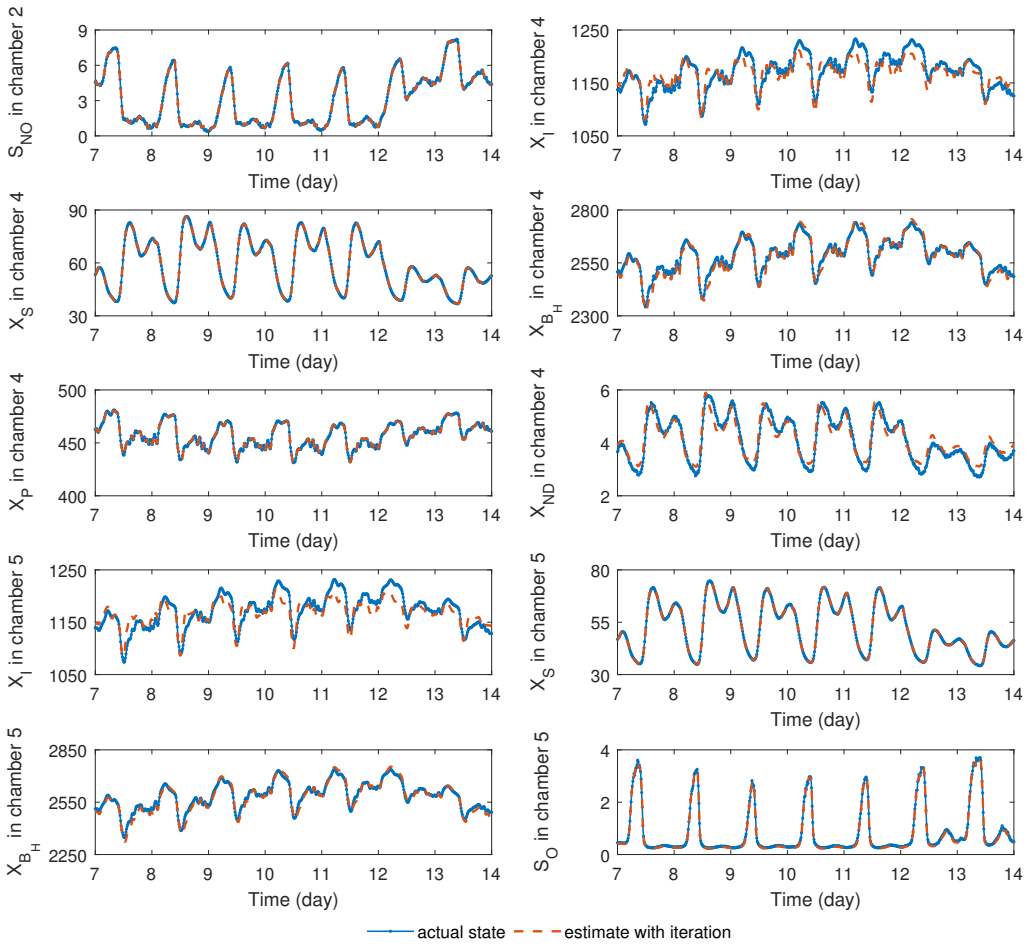

Figure 6.4: The trajectories of the actual states (blue dash dot lines) and the state estimates given by iterative distributed MHE (red dashed lines) in dry weather

The thresholds on the innovations of the three subsystems in Step 2.3 of Algorithm 3 are selected as $T_1 = 1.83$, $T_2 = 1.96$ and $T_3 = 1.70$. The values of δ_i, $i = 1, 2, 3$, are selected to be $\delta_1 = 0.35$, $\delta_2 = 0.40$, $\delta_3 = 0.65$. In terms of the MHE estimators for the three subsystems, the weighting matrices are selected to be $Q_i = \text{diag}(\sigma_{w,i}^{\circ 2})$ where $\sigma_{w,i} := 0.06 x_{is}$ and $R_i = \text{diag}(\sigma_{v,i}^{\circ 2})$ where $\sigma_{v,i} := 0.05 y_{is}$. Unless stated otherwise, the initial guess of each MHE estimator i is set to be $1.08 x_{is}$, $i \in \mathbb{I}$. The length of the estimation horizon of each MHE estimator is determined to be: $N_1 = N_2 = 10$ and $N_3 = 20$. Note that in the evaluation of the MHE estimators, each system state $x_{i,j}$ (which denotes the j-th element of state vector

Table 6.7: Actual values and the corresponding estimates of EQ and OCI in different weather conditions

	Actual value in dry weather	Estimate in dry weather	Actual value in rain	Estimate in rain	Actual value in storm	Estimate in storm
EQ	6.3062×10^3	6.3092×10^3	8.4994×10^3	8.6314×10^3	7.5390×10^3	7.6243×10^3
OCI	1.6223×10^4	1.6222×10^4	1.5936×10^4	1.5968×10^4	1.7061×10^4	1.7062×10^4

x_i) is scaled following

$$\bar{x}_{i,j}(t_k) = \frac{x_{i,j}(t_k) - x_{i,j}^{\min}}{x_{i,j}^{\max} - x_{i,j}^{\min}}$$

where $x_{i,j}^{\max} := \max\{x_{i,j}(t_k) : k \in \mathbb{K}_+\}$ and $x_{i,j}^{\min} := \min\{x_{i,j}(t_k) : k \in \mathbb{K}_+\}$.

Remark 39. *The size of the estimation horizon for the MHE estimator of subsystem 3 (i.e., N_3) is made larger than N_1 and N_2 because: (1) there are more state variables in subsystem 3 than the other two subsystems; (2) the MHE estimator for subsystem 3 is a linear one such that the computational complexity of the evaluation of this estimator will still be tractable with a larger-sized estimation horizon.*

6.4.2 Results of dry weather condition

We first apply the proposed distributed MHE approach to WWTP under the dry weather condition. The state estimate and the actual state trajectories for certain process states in the second week of the operation are presented in Figure 6.4 and Figure 6.5. The results indicate that the proposed scheme can provide satisfying state estimates in the presence of external disturbances and constraints. Then, EQ and OCI are calculated based on the actual system states and the state estimates given by the iterative distributed state estimation scheme. The values are presented in Table 6.7. The results show that the EQ and OCI can be accurately estimated based on the proposed method.

Next, comparisons are made between the proposed iterative distributed MHE and a decentralized MHE scheme in terms of estimation accuracy and computational complexity. In

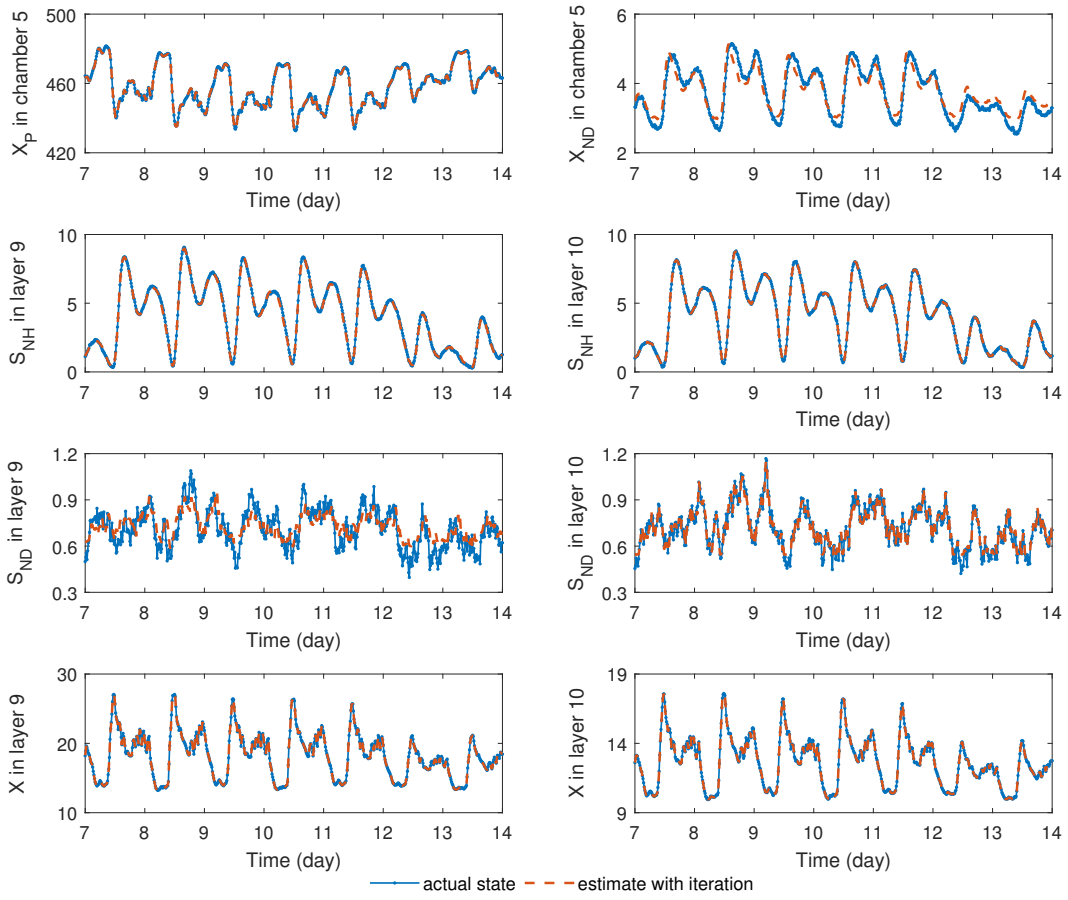

Figure 6.5: The trajectories of the actual states (blue dash dot lines) and the state estimates given by iterative distributed MHE (red dashed lines) in dry weather

the decentralized scheme, the local estimators do not communicate with each other to exchange subsystem state estimates. Steady-state values of the process states are calculated based on a 100-day open-loop operation subject to constant inputs following [149]. These steady-state values are used in the local MHEs of the decentralized scheme to conservatively compensate for the time-varying subsystem interaction [129]. The mean value of the Euclidean norm of the normalized estimation error and the average computation time required for the one-sampling-time evaluation of the two schemes are given in Table 6.8 for comparison. The average computation time required by the decentralized scheme is slightly shorter, which is mainly because each local estimator within the decentralized framework is evaluated

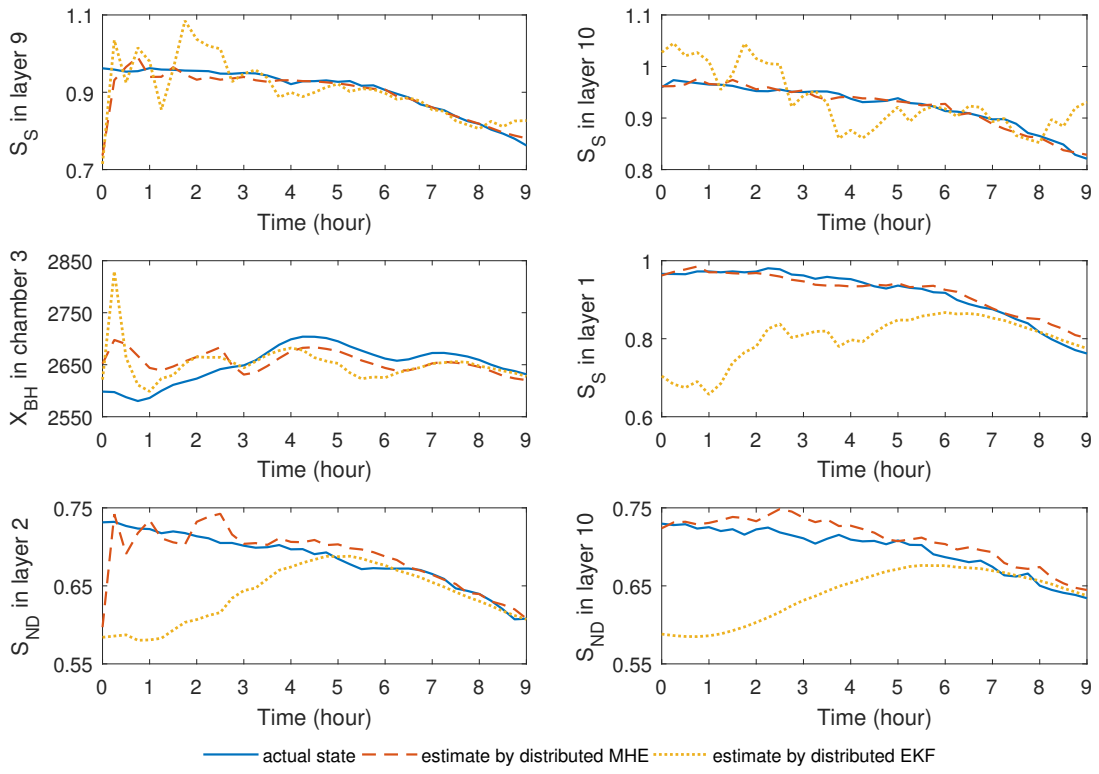

Figure 6.6: The trajectories of the actual states (blue solid lines) and the state estimates by iterative distributed MHE (red dashed lines) and the state estimates given by distributed EKF (yellow dotted lines) at the initial stage in dry weather

only once at one sampling time. However, the estimation error norm given by the decentralized scheme is significantly larger than that of the proposed method, which indicates that the proposed distributed MHE can give much improved estimation performance compared to the decentralized scheme. Note that in Table 6.8, the average computation time for the distributed MHE/decentralized MHE is the average time required by the MHE estimator for subsystem 1. This is because the evaluation of this local estimator requires the longest computation time among the three estimators.

We also make performance comparisons between the proposed iterative distributed MHE scheme and centralized MHE. The estimation horizon of the centralized MHE is picked as $N = 36$. Note that the length of the estimation horizon may be further tuned for possibly better performance, which is not within the scope of this work. The mean value of the Euclidean norm of the normalized estimation error and the average computation time for

Table 6.8: Mean values of the Euclidean norm of the normalized estimation error and the average computation time required for the one-sampling-time evaluation for iterative distributed MHE and centralized MHE

	Iterative distributed MHE	Decentralized MHE	Centralized MHE
Mean of the error norm	3.3081	61.5810	3.1196
Average computation time	3.59 sec	3.46 sec	41.75 sec

the centralized MHE scheme is also given in Table 6.8. While the centralized MHE scheme can provide slightly better estimation accuracy compared with the proposed approach, the use of the proposed iterative distributed MHE leads to significantly reduced computation time (saving 91.40% of the computation time compared with the considered centralized MHE scheme).

In addition, we show that the proposed approach can provide better estimates for the WWTP based on BSM1 compared to the distributed EKF method in [125]. In this set of simulations, the initial guess of each estimator for subsystem i (both MHE estimator i and EKF estimator i) is set to be $0.3x_{is}$, $i \in \mathbb{I}$. The state estimates for certain states given by the two approaches are given in Figure 6.6. The results show that the proposed distributed MHE scheme can provide much better estimates for certain states at the initial stage when the initial guess is not good.

Remark 40. *In the simulations, we see that the distributed MHE gives improved estimation performance compared with the distributed EKF. This may due to the inherent advantages of MHE over EKF. First, MHE can incorporate the information of physical constraints on system states in the estimation optimization problem. However, the constraints are not explicitly considered in the formulation of EKF. Also, EKF is not robust to disturbance or poor initial guess due to the way nonlinearity is treated in EKF [127]. Moreover, the convergence of EKF requires the initial estimation error to be sufficiently accurate as shown in [140, 141].*

6.4.3 Results in rainy and stormy weather conditions

In different weather conditions, the flow rate and the concentration of the wastewater entering the process (the influent) are much different. It is necessary to verify the effectiveness of the proposed method in different weather conditions. In this section, we consider two common weather conditions, i.e., rainy and stormy weather. The length of the data corresponding to each weather condition is two weeks. The data of rainy weather consist of one week of dynamic dry weather data and a long rain event during the second week. The data of stormy weather also consists of two portions: the first portion is one week of dynamic dry weather data, while the rest portion describes two storm events superimposed on the dry weather data during the second week. More detailed information about the different weather conditions is given in [130]. The data can be found in [139].

The proposed approach is also applied to the WWTP when it is operated in rainy and stormy weather conditions. The actual system states and the corresponding state estimates given by the proposed approach are obtained under the two considered weather conditions. The trajectories of the actual state and the estimate for certain process states are shown in Figure 6.7 (for rainy weather) and Figure 6.8 (for stormy weather). Further, we show the actual trajectories of EQ_f generated based on both the actual states and the state estimates in Figure 6.9. We also present the values of EQ and OCI which are calculated based on actual process states and state estimates in the two weather conditions as in Table 6.7. The results confirm that the proposed method can provide good state estimates under different weather conditions.

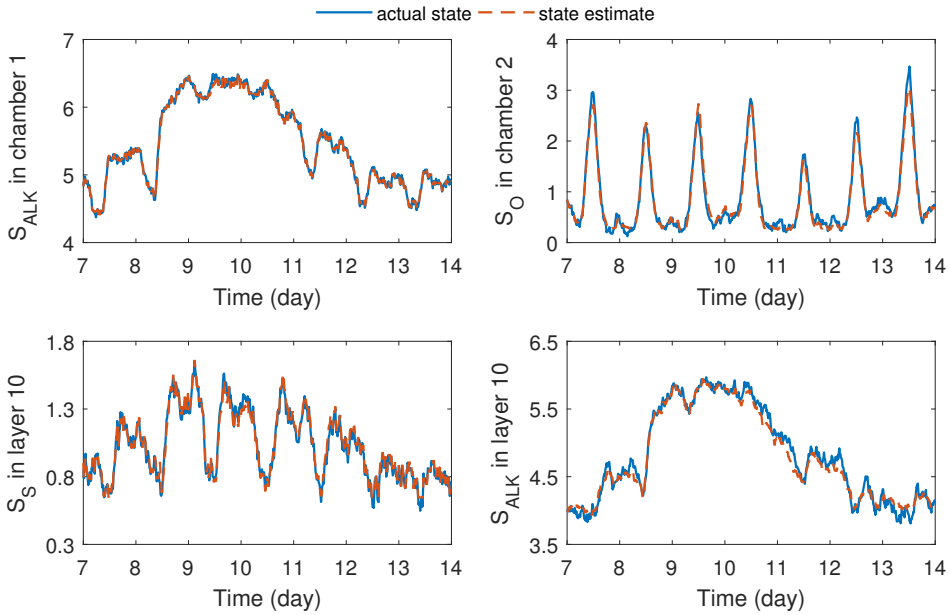

Figure 6.7: The trajectories of the actual states (blue solid lines) and the state estimates given by iterative distributed MHE (red dashed lines) in rainy weather

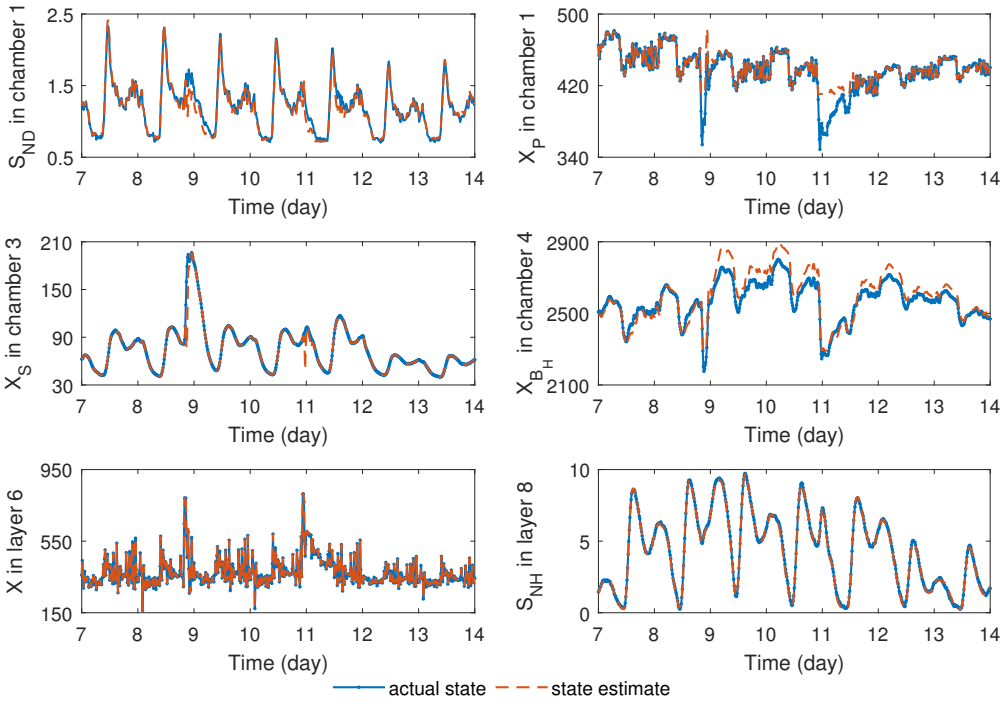

Figure 6.8: The trajectories of the actual states (blue solid lines) and the state estimates given by iterative distributed MHE (red dashed lines) in stormy weather

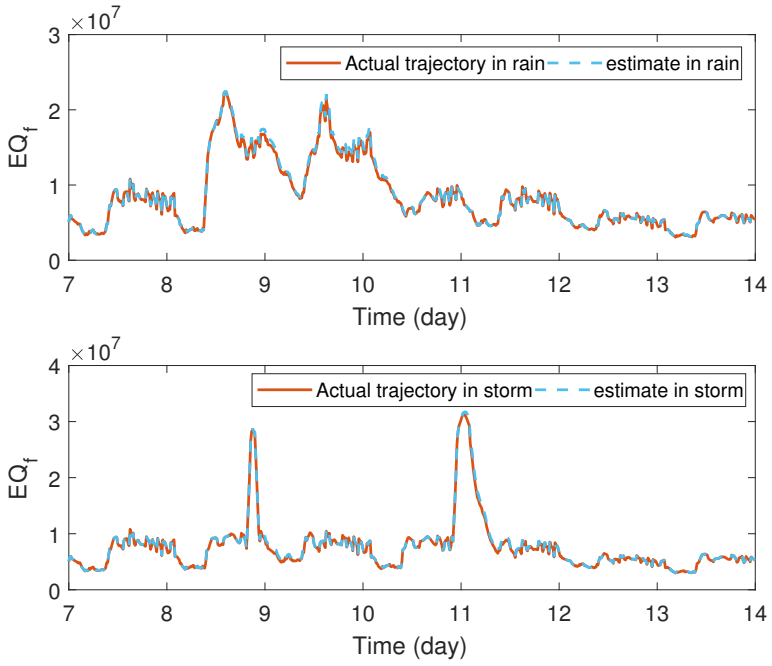

Figure 6.9: The trajectories of EQ_f calculated based on actual states (solid red lines) and state estimates (blue dashed lines) in rainy and stormy weather

Remark 41. *The distributed MHE and the distributed EKF considered in performance comparison are developed based on the same subsystem decomposition. The aim of this comparison is to show that the proposed distributed MHE can perform better than the distributed EKF. We note that the decomposed subsystem structure obtained and used in this work is not unique for distributed state estimation, and there could be other subsystem decompositions that can be used to achieve better estimation performance for the WWTP. How to optimally decompose the WWTP into subsystems for distributed state estimation is a very important problem and will be considered in our future work.*

6.5 Summary

In this chapter, the state estimation problem for the WWTP based on BSM1 was addressed. An approach on subsystem decomposition and configuration was proposed based on structural closeness. The WWTP was decomposed into 13 groups of state and output mea-

surement variables using the proposed approach. The 13 groups were grouped to form 3 subsystems by considering the interaction between different groups as well as the nonlinearity of the state dynamics of each group. A distributed MHE scheme based on iterative evaluation was proposed for the WWTP. Innovation-based triggering conditions were developed for the evaluation of the distributed scheme to reduce its computational complexity. Both nonlinear and linear MHE approaches were used to design subsystem MHE estimators. Simulations were carried out under different weather conditions. The proposed scheme can provide good state estimates in different weather. We also made performance comparisons with non-iterative distributed MHE, centralized MHE and distributed EKF. The results show that the iterative distributed MHE can provide more balanced performance compared to its counterparts. Specifically, the proposed approach provides better estimates than non-iterative distributed MHE and distributed EKF do, while it requires much lower computational cost compared with centralized MHE.

Appendix II: the model of the WWTP

In this appendix, the differential equations describing the dynamics of the WWTP based on BSM1 are presented [130]. The dynamics of the biological reactor are described as follows:

- For the chamber k ($k = 1$) of the reactor:

$$\frac{dZ_1}{dt} = \frac{1}{V_1}(Q_a Z_a + Q_r Z_r + Q_0 Z_0 + r_1 V_1 - Q_1 Z_1)$$
$$Q_1 = Q_a + Q_r + Q_0$$

- For the chamber k ($k = 2, \ldots, 5$) of the reactor:

$$\frac{dZ_k}{dt} = \frac{1}{V_k}(Q_{k-1} Z_{k-1} + r_k V_k - Q_k Z_k)$$
$$Q_k = Q_{k-1}$$

- Special case for the concentration of dissolved oxygen in chamber k (denoted by $S_{O,k}$), $k = 1, \ldots, 5$:

$$\frac{dS_{O,k}}{dt} = \frac{1}{V_k}\Big(Q_{k-1}S_{O,k-1} + r_k V_k + K_L a_k V_k \left(S_O^* - S_{O,k}\right) - Q_k S_{O,k}\Big)$$

where Z_k represents a state defined in Table 6.1 (except S_O), V_k is the volume, r_k denotes the observed reaction rate, $K_L a_k$ is the oxygen transfer coefficient, S_O^* denotes the saturation concentration of dissolved oxygen. Note that subscript k denotes a chamber in the reactor ($k = 1, \ldots, 5$), and the first two chambers are non-aerated, thus $K_L a_1 = K_L a_2 = 0$ d^{-1}.

The mass balance model of the sludge in the settler is established as follows:

$$\begin{aligned}
\frac{dX_1}{dt} &= \frac{1}{z_1}\Big(v_{up}(X_2 - X_1) + J_{clar,1}\Big) \\
\frac{dX_p}{dt} &= \frac{1}{z_p}\Big(v_{up}(X_{p+1} - X_p) + J_{clar,p-1} - J_{clar,p}\Big) \quad p = 2, \ldots, 4 \\
\frac{dX_5}{dt} &= \frac{1}{z_5}\left(\frac{Q_f X_f}{A} + J_{clar,4} - \left(v_{up} - v_{dn}\right)X_5 - J_{s,6}\right) \\
\frac{dX_6}{dt} &= \frac{1}{z_6} v_{dn}(X_5 - X_6) \\
\frac{dX_7}{dt} &= \frac{1}{z_7}\Big(v_{dn}(X_6 - X_7) + J_{s,6} - J_{s,8}\Big) \\
\frac{dX_8}{dt} &= \frac{1}{z_8} v_{dn}(X_7 - X_8) \\
\frac{dX_9}{dt} &= \frac{1}{z_9}\Big(v_{dn}(X_8 - X_9) + J_{s,8} - J_{s,9}\Big) \\
\frac{dX_{10}}{dt} &= \frac{1}{z_{10}}\Big(v_{dn}(X_9 - X_{10}) + J_{s,9}\Big)
\end{aligned}$$

where subscript p represents a layer in the settler and ranges from $p = 1, \ldots, 10$, X_p is the concentration of the sludge in the p-th layer of the settler, z_p is the height of each layer, $J_{clar,p}$ denotes the the flux in the clarification layer p ($p = 1, \ldots, 4$), $J_{s,p}$ denotes the solid flux due to gravity in layer p ($p = 5, \ldots, 10$), $v_{dn} = \frac{Q_r + Q_w}{A}$ and $v_{up} = \frac{Q_e}{A}$ with A being the cross sectional area of the settler.

The dynamics of the concentrations of the soluble materials in the settler are described as follows:

- For layers ($p = 1, \ldots, 4$):

$$\frac{dZ_{s,p}}{dt} = \frac{1}{z_p} v_{up} \Big(Z_{s,p+1} - Z_{s,p} \Big)$$

- For the feed layer ($p = 5$):

$$\frac{dZ_{s,5}}{dt} = \frac{1}{z_5} \left(\frac{Q_f Z_f}{A} - (v_{dn} + v_{up}) Z_{s,5} \right)$$

- For layers ($p = 6, \ldots, 10$):

$$\frac{dZ_{s,p}}{dt} = \frac{1}{z_p} v_{dn} \Big(Z_{s,p-1} - Z_{s,p} \Big)$$

where $Z_{s,p}$ is the concentration of each soluble material in the p-th layer of settler.

The parameters of this process model can be found in [130].

Chapter 7

Subsystem decomposition of process networks for simultaneous distributed state estimation and control

An appropriate subsystem configuration is a prerequisite for a successful distributed control/state estimation design. Existing subsystem decomposition methods are not designed to handle simultaneous distributed estimation and control.we focus on subsystem decomposition of nonlinear process networks for simultaneous distributed state estimation and distributed control. To achieve this goal, in this chapter, we propose a systematic approach based on the concept of community structure detection. We resort to the measure of modularity to quantitatively assess the quality of different community structures. Specifically, the state, manipulated input and measured output variables of a process are taken into account and are viewed as vertices in a network. The ways to construct a directed graph containing all the vertices and the corresponding adjacency matrix are presented. An implementation procedure based on approximate optimization of modularity is developed, such that subsystem models for simultaneous distributed state estimation and distributed control can be established by allocating vertices into communities based on modularity. The proposed

method is applied to two process examples with different complexities. Candidate subsystem structures that can be used for simultaneous distributed estimation and control are recommended by the proposed method for each process example. This chapter is a revised version of [73].

7.1 Preliminaries

7.1.1 Notation

$L_f h$ represents the Lie derivative of function h with respect to function f, defined as $L_f h(x) = \frac{\partial h}{\partial x} f(x)$. $L_f^r h$ represents the r-th order Lie derivative of function f, defined as $L_f^r h(x) = L_f L_f^{r-1} h(x)$. δ_{ij} is the Kronecker delta function of two variables i and j such that $\delta_{ij} = 1$ if $i = j$ or $\delta_{ij} = 0$ if $i \neq j$.

7.1.2 System model

In this work, we consider a class of process network systems that can be described by the following general form:

$$\dot{x}(t) = f\left(x(t), u(t), p(t)\right) \tag{7.1a}$$

$$y(t) = h\left(x(t)\right) \tag{7.1b}$$

where $x \in \mathbb{R}^{n_x}$ denotes the state vector of the nonlinear system, $y \in \mathbb{R}^{n_y}$ is the vector containing all the measured outputs, $u \in \mathbb{R}^{n_u}$ is the vector of the manipulated inputs, $p \in \mathbb{R}^{n_p}$ is the vector of time-varying parameters (i.e., time-varying disturbances to the system that are exactly known.).

7.1.3 Subsystem model

In this work, the objective is to decompose the entire process network system described in (7.1) into subsystems that are appropriate for simultaneous distributed estimation and control. Specifically, we aim to decompose system (7.1) into q subsystems in the following form:

$$\dot{x}^{(i)}(t) = \mathbf{f}^{(i)}\left(x^{(i)}(t), X^{(i)}(t), u^{(i)}(t), \bar{u}^{(i)}(t), p_i(t)\right) \tag{7.2a}$$

$$y^{(i)}(t) = \mathbf{h}^{(i)}(x^{(i)}(t)) \tag{7.2b}$$

where $x^{(i)} \in \mathbb{R}^{n_{x^{(i)}}}$ is the state vector of the i-th subsystem, $u^{(i)} \in \mathbb{R}^{n_{u^{(i)}}}$ and $\bar{u}^{(i)} \in \mathbb{R}^{n_{\bar{u}^{(i)}}}$ together (i.e., $\left[{u^{(i)}}^{\mathrm{T}} \ {\bar{u}^{(i)}}^{\mathrm{T}}\right]^{\mathrm{T}}$) account for manipulated inputs to the i-th subsystem, $y^{(i)} \in \mathbb{R}^{n_{y^{(i)}}}$ is the vector of measured outputs of subsystem i, $p^{(i)} \in \mathbb{R}^{n_{p^{(i)}}}$ is a vector containing time-varying parameters that explicitly affect the dynamics of $x^{(i)}$, $X^{(i)} \in \mathbb{R}^{n_{X^{(i)}}}$ represents a vector containing the states of neighboring subsystems that directly affect the dynamics of $x^{(i)}$, $i \in \{1, \ldots, q\}$. For example, if the dynamics of subsystem 4 is affected directly by the states of subsystem 1, subsystem 3 and subsystem 4, then one has $X^{(4)} = \left[{x^{(1)}}^{\mathrm{T}} \ {x^{(3)}}^{\mathrm{T}}\right]^{\mathrm{T}}$. Note that the manipulated input vector of the i-th subsystem comprises two parts: $u^{(i)}$ denotes manipulated inputs determined by the local controller of subsystem i, and $\bar{u}^{(i)}$ denotes manipulated inputs that affect subsystem i but determined by controllers of other subsystems. This means that, for the local controller of subsystem i, $i = 1, \ldots, q$, $u^{(i)}$ contains its decision variables and $\bar{u}^{(i)}$ contains known input information generated by and sent from the controller(s) of the interacting subsystem(s).

7.1.4 Directed graph

In the proposed approach, directed graph is used to determine the connectivity between state, manipulated and output variables. For system (7.1), a directed graph can be constructed by treating the state, manipulated, output variables as vertices that are connected via directed

edges. Let x_i denote the i-th element of vector x, $i = 1, \ldots, n_x$, y_j denote the j-th measured output of vector y, $j = 1, \ldots, n_y$, and u_k denote the k-th input variable of vector u, $k = 1, \ldots, n_u$. Further, let us denote by f_i the i-th element of the vector field f, $i = 1, \ldots, n_x$, and denote by h_j the j-th element of the vector field h, $j = 1, \ldots, n_y$. The edges can be placed based on the following rules [118, 132]:

- State-to-state edge: there exists a unidirectional edge from x_i to x_k, if $\partial f_k(x, u)/\partial x_i \neq 0$, $k, i = 1, \ldots, n_x$.
- Input-to-state edge: there exists a unidirectional edge from u_i to x_k, if $\partial f_k(x, u)/\partial u_i \neq 0$, $k = 1, \ldots, n_x$, $u = 1, \ldots, n_u$.
- State-to-output edge: there exists a bidirectional edge between x_i and y_j, if $\partial h_j(x)/\partial x_i \neq 0$, $j = 1, \ldots, n_y$, $i = 1, \ldots, n_x$.

Remark 42. *We note that the rules used in this work for constructing edges between variables are slightly different from those adopted in existing literature [118, 137, 61, 72, 132]. Specifically, while the state-to-state edges and input-to-state edges are still unidirectional, the edges that connect the state and measured output variables are bidirectional. This is due to the consideration that the output measurement equations (i.e., h_i, $1, \ldots, n_y$) are algebraic equations instead of differential equations.*

7.1.5 Observability of nonlinear systems

When state estimation is considered, observability of the states based on given output measurements is important. One approach to check the observability of a nonlinear system is to check whether the full-rank condition is satisfied by the corresponding nonlinear observability

matrix $Q(x)$, which is defined as follows [30]:

$$Q(x) = \frac{d}{dx} \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{bmatrix}. \tag{7.3}$$

However, when the order of a system is large, it may be challenging to obtain higher order Lie derivatives or the observability matrix may be ill-conditioned. In these cases, a practical approach is to linearize the nonlinear system at different points along typical operating trajectories and check the observability of the linearized systems. If the linearized models are observable, it can be concluded that the nonlinear system is locally observable along the considered trajectories.

The above methods will be used to test the observability of a considered process as well as the observability of each decomposed subsystem.

7.1.6 Community structure detection

In this section, we briefly review the method of community structure detection for directed networks, which will be used for subsystem decomposition. The community structure detection method has been considered as an effective tool to divide a large-scale network into communities/groups such that the connection within each group is denser while the inter-group connection is made sparser [144, 145]. In these approaches, the quality of a certain community structure is quantitatively assessed by the measure of modularity [142]. In this work, we use the approach proposed in [145] for directed networks.

Specifically, consider a network of n interconnected vertices. The modularity of a com-

munity structure can be calculated as follows:

$$Q = \frac{1}{m} \sum_{i,j=1,\ldots,n} \left(A_{ij} - \frac{k_i^{\text{in}} k_j^{\text{out}}}{m} \right) \delta_{c_i,c_j} \tag{7.4}$$

where k_i^{in} and k_i^{out} are the in-degrees and the out-degrees of vertex i (i.e., the numbers of edges entering and leaving vertex i, $i = 1, \ldots, n$), respectively, m is the total number of edges in the directed graph, c_i is an integer labelling the community to which vertex i, $i = 1, \ldots, n$, belongs, A_{ij} is the i-th element of the j-th of the adjacency matrix for the considered network. $A_{ij} = 1$ if there is a directed edge from vertex j to vertex i ($j, i \in \{1, \cdots, n\}$) and $A_{ij} = 0$ otherwise. Note that in this work, self edges (i.e., directed edges, each of which exits from/enters the same vertex) are not taken into account in the construction of an adjacency matrix; that is, the diagonal elements of an adjacency matrix are always zero regardless of the existence of self edges. In (7.4), the term $\frac{k_i^{\text{in}} k_j^{\text{out}}}{m}$ measures the probability of having an edge from vertex j to vertex i edges between the two vertices in the case of a randomized network where the edges are randomly placed. The value of modularity can be either positive or negative, with positive modularity values indicating that there are more within-community edges than expected. Since larger positive values indicate better community structures, the problem of finding the best community structure is equivalent to maximizing the modularity value over all possible community structure candidates. The optimal community structure can be found by finding the feasible label sequence $\{c_1, \ldots, c_n\}$ such that Q in Eq.(7.4) is maximized given a known adjacency matrix. In practice, a modularity value greater than 0.3 often indicates a good community structure [146].

7.2 Proposed subsystem decomposition method

In this section, we present the proposed subsystem decomposition method for process networks for simultaneous distributed estimation and control. From the perspective of community structure detection, the original nonlinear system can be viewed as a large network,

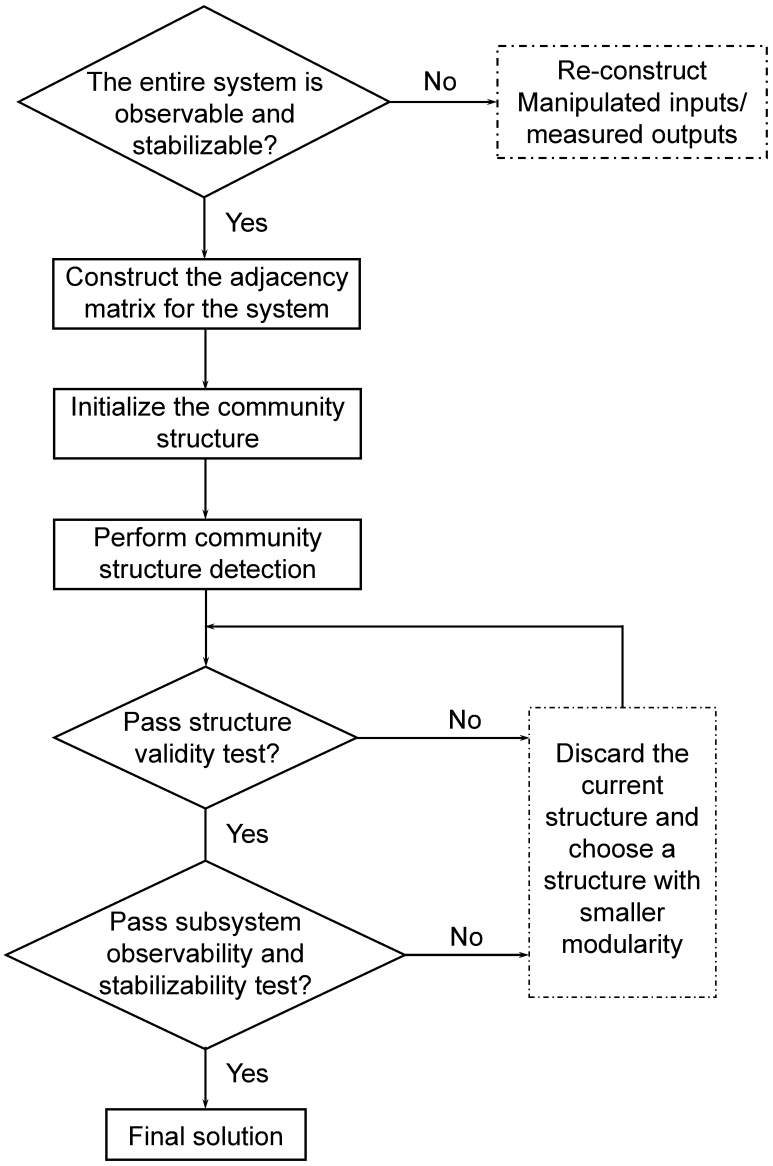

Figure 7.1: A flowchart of the proposed subsystem decomposition method

with state, manipulated input and measured output variables being vertices in the network. Accordingly, each configured subsystem consisting of certain variables with stronger connections is a community of the network. A flowchart that describes the key steps in the proposed method is shown in Figure 7.1.

7.2.1 Consideration on observability and stabilizability

To design a state estimation based control system, it is necessary to ensure that the entire system satisfies detectability (or observability) and stabilizability (or controllability) simultaneously in the operating region. In the first step of the proposed method, we check if system (7.1) is both observable and stabilizable. To verify the observability of system (7.1), the method described in Section 7.1.5 is used. Note that we can check the stabilizability of the system in a similar way at each point along the typical operating trajectories. If system (7.1) is either unobservable or unstabilizable, further steps in the proposed method cannot be carried out. In this case, the measured outputs or the manipulated inputs should be re-constructed accordingly and then the observability and stabilizability of the system should be examined again.

Remark 43. *It is more often to consider detectability and stabilizability of a nonlinear process. In this work, we require the entire system in (7.1) to be observable along typical trajectories, which is stricter than requiring the system to be detectable. This is because time-varying parameters are considered in system (7.1). When the system is not under a steady-state operation, detectability may not suffice for obtaining accurate state estimates. Therefore, we require that the system is locally observable at each point along typical operating trajectories.*

7.2.2 Adjacency matrix construction

After ensuring the entire system satisfies observability and stabilizability, an adjacency matrix should be constructed for system (7.1) for subsystem decomposition. To account for both state estimation and control, all the state, manipulated input and measured output variables are considered as vertices. In total, there are n_d (where $n_d = n_x + n_y + n_u$) vertices in the directed graph constructed based on (7.1). Let $\mathbf{c}_d$ denote an augmented vector such that $\mathbf{c}_d = [x_1, \ldots, x_{n_x}, u_1, \ldots, u_{n_u}, y_1, \ldots, y_{n_y}]$. The i-th element of $\mathbf{c}_d$ cor-

responds the i-th vertex in the directed graph of system (7.1). Accordingly, we use $\mathbf{s}_d$ ($\mathbf{s}_d = [c_{x_1}, \ldots, c_{x_{n_x}}, c_{u_1}, \ldots, c_{u_{n_u}}, c_{y_1}, \ldots, c_{y_{n_y}}]$) to denote a n_d-dimensional vector in which the i-th element is the community label of the i-th vertex in $\mathbf{c}_d$, $i = 1, \ldots, n_d$.

An $n_d \times n_d$ adjacency matrix $\mathbf{A}_d$ involving all the vertices in $\mathbf{c}_d$ can be generated based on the directed edges between the vertices in $\mathbf{c}_d$. However, the construction of $\mathbf{A}_d$ via examining the existence of directed edges between each two vertices in a directed graph can be demanding when n_d is not small. Alternatively, we can obtain $\mathbf{A}_d$ based on Jacobian matrices of the vector fields f and h in (7.1). Specifically, let us define:

$$\bar{A} = \left.\frac{\partial f(x,u)}{\partial x}\right|_{(x_s,u_s)}, \quad \bar{B} = \left.\frac{\partial f(x,u)}{\partial u}\right|_{(x_s,u_s)}, \quad \bar{C} = \left.\frac{\partial h(x)}{\partial x}\right|_{x_s} \tag{7.5}$$

where (x_s, u_s) denotes an equilibrium point of system (7.1).

Then, a matrix is constructed as follows:

$$\bar{\mathbf{A}}_d = \begin{bmatrix} \bar{A} & \bar{B} & \bar{C}^{\mathrm{T}} \\ \mathbf{0}_{n_u \times n_x} & \mathbf{0}_{n_u \times n_u} & \mathbf{0}_{n_u \times n_y} \\ \bar{C} & \mathbf{0}_{n_y \times n_u} & \mathbf{0}_{n_y \times n_y} \end{bmatrix}_{n_d \times n_d}$$

$\mathbf{A}_d$ can be obtained based on $\bar{\mathbf{A}}_d$ by: (1) substituting all non-zero elements in $\bar{\mathbf{A}}_d$ by 1; (2) setting all the diagonal elements in $\bar{\mathbf{A}}_d$ to be zero.

Remark 44. *In this work, the adjacency matrix constructed for subsystem decomposition can be considered as a non-trivial extension of the ones defined in existing literature [132, 102]. Measured output variables are also incorporated as vertices for community structure detection in order to handle distributed state estimation. Both $\bar{C}^T$ and $\bar{C}$ are used in constructing the adjacency matrix, which accounts for the bidirectional edges between state and measured output variables as defined in Section 7.1.4. Also, the way used in this work to construct the adjacency matrix is more efficient due to the use of Jacobian matrices instead of examining the existence of directed edges between each two vertices.*

Remark 45. *In applications, there can be elements with very small absolute values in the Jacobian matrices $\bar{A}$, $\bar{B}$ and $\bar{C}$ due to numerical precision. This issue may be handled by specifying a small positive threshold such that an element in A_d is not treated as zero only if the absolute value of the element in A_d is greater than the threshold.*

7.2.3 Initialization of the community structure

Based on the constructed adjacent matrix, we can move forward with community detection for subsystem decomposition. To perform community structure detection via finding a higher modularity value defined in (7.4), an initial community structure is always required. Consider an initial structure $\mathbf{s}_d(0) = [c_{x_1}(0), \ldots, c_{x_{n_x}}(0), c_{u_1}(0), \ldots, c_{u_{n_u}}(0), c_{y_1}(0), \ldots, c_{y_{n_y}}(0)]$ where each element represents the community label of the corresponding variable. The values of the elements in $s_d(0)$ are dependent on different approximation methods that can be adopted to solve the problem. In this work, the fast unfolding algorithm [143] is used for community structure detection, and the community structure is initialized by assigning the i-th vertex in $\mathbf{s}_d$ to the i-th community, $i = 1, \ldots, n_d$; that is, $\mathbf{s}_d(0) = [1, 2, \ldots, n_d]$.

While this type of initialization can help achieve a higher modularity value via implementing the fast unfolding algorithm, constraints on the subsystem structure due to the form of the measured output model in (7.2b) are not taken into account. Specifically, from the model in (7.2b), it is seen that each measured output of a subsystem is only dependent on the state(s) of the same subsystem. This implies that for each measured output y_i, $i = 1, \ldots, n_y$, all the state variables involved in the argument of scalar function $h_i(\cdot)$ (i.e., the state variables that directly affect y_i) should be finally assigned to the same community that y_i belongs to. To handle this issue, we update the initial structure $\mathbf{s}_d(0)$ by incorporating the above constraint. This is done through two steps:

- For each state variable x_i, $i = 1, \ldots, n_x$, find all the measured outputs that have state-to-output edges with respect to x_i, and assign these measured outputs to the same community.

- For each measured output y_j, $j = 1, \ldots, n_y$, find all the state variables that have state-to-output edges with respect to y_j, and assign these state variables to the same community to which y_i belongs.

The updated community structure is denoted as $\mathbf{s}'_d(0)$. $\mathbf{s}'_d(0)$ will be used as the initial community structure in subsystem decomposition instead of $\mathbf{s}_d(0)$.

7.2.4 Community detection for subsystem configuration

In the proposed procedure, the desired number of subsystems should be pre-specified. Once the number of subsystems is specified, the proposed procedure finds subsystem configurations (in which each subsystem is a community) that maximize the modularity that is defined in Section 7.1.6. To address the computational issue of modularity maximization, the fast folding algorithm proposed in [143] is adopted in our work. Based on the initialization in Section 7.2.3 and the pre-specified number of subsystems, we carry out the following steps to detect the communities for subsystem decomposition:

1. (***Initial aggregation of vertices***) For the initial community structure, calculate the modularity value following Eq.(7.4), and aggregate the vertices from each community (if there are more than one vertex) into an aggregated vertex. Set $k = 0$.

2. (***Community detection based on modularity maximization***) Compare $n_c(k)$ and q. If $n_c(k) > q$, go to Step 2.1. Else, go to Step 4.

   2.1 Repeat the following steps for each vertex (including aggregated vertex) i, $i = 1, \ldots, n_c(k)$:

      2.1.1 For each neighboring vertex (including aggregated vertex) of vertex i, i.e., vertex j, calculate the change in the modularity value (i.e., ΔQ) by moving vertex i from its current community to the community of vertex j.

2.1.2 Find the maximum $\Delta Q > 0$ that can be achieved in the previous step and place each vertex i in the corresponding community. And aggregate the vertices in that community into a new aggregated vertex.

2.2 Set $n_c(k+1)$ to be the number of communities after vertex aggregation. If $n_c(k+1) < n_c(k)$, go to Step 2.3. Else, go to Step 3.

2.3 If $n_c(k+1) \geq q$, set $k = k+1$ and go to Step 2. Else if $n_c(k+1) < p$, find the community structure of p communities (equivalently, p vertices) with the largest modularity value among all the structures of p communities. Go to Step 4.

3. (***Community detection subject to smallest decrease in modularity***) Compare $n_c(k)$ and q. If $n_c(k) > q$, go to Step 3.1. Else, go to Step 4.

3.1 Repeat the following steps for each vertex i (including aggregated vertex), $i = 1, \ldots, n_c(k)$:

3.1.1 For each neighboring vertex (including aggregated vertex) of vertex i, i.e., vertex j, calculate the change in the modularity value (i.e., ΔQ) by moving vertex i from its current community to the community of vertex j.

3.1.2 Find the minimum $|\Delta Q|$ (denoted as $|\Delta Q(i)|_{\min}$) that can be achieved among all the neighboring vertices.

3.2 Find the minimum decrease in the modularity value (denoted as $|\Delta Q|_{\min}$) with respect to all the currently existing vertices, i.e., $|\Delta Q|_{\min} := \min\{|\Delta Q(i)|_{\min} : i = 1, \ldots, n_c(k)\}$, and find the corresponding vertex placement.

3.3 Place the corresponding vertex in the appropriate community which leads to $|\Delta Q|_{\min}$. And aggregate the vertices in this community into a new vertex.

3.4 Set $n_c(k+1)$ to be the number of vertices after aggregation. Set $k = k+1$ and go to Step 3.

4. (***Forming subsystem models based on the communities***) Form each subsystem

model in the form of (7.2) by selecting elements of the vector field f that correspond to the state variables in each community.

5. For each subsystem i, form vector $u^{(i)}$ using the manipulated input variables in the corresponding community and then form vector $\bar{u}^{(i)}$.

The following numerical example is used to better explain Step 4 and Step 5 in the above procedure.

Example 4. Let us consider the following linear state-space model:

$$\dot{x}_1 = -4x_1 + 2.5x_2 + u_1 + 0.3u_2 \tag{7.6a}$$

$$\dot{x}_2 = 0.1x_1 - 5x_2 \tag{7.6b}$$

$$\dot{x}_3 = 13x_1 - 22x_3 + 2x_4 + 0.1u_1 + u_2 \tag{7.6c}$$

$$\dot{x}_4 = 13x_2 - 0.1x_3 - 23x_4 \tag{7.6d}$$

$$\dot{y}_1 = x_1 \tag{7.6e}$$

$$\dot{y}_2 = x_3 \tag{7.6f}$$

where x_1, x_2, x_3 and x_4 are the system states, u_1 and u_2 are two manipulated inputs, and y_1 and y_2 are the two measured outputs.

Let us consider decomposing the above system into two subsystems. Following Steps 1-3 of the procedure, two communities can be found as follows: the first community contains vertices x_1, x_2, u_1 and y_1 while the second community contains vertices x_3, x_4, u_2 and y_2. The state variables in the first community are x_1 and x_2, so that the first two differential equations in Eq.(7.6) (i.e., Eq.(7.6a) and Eq.(7.6a) constitute the model of subsystem 1 according to Step 4 of the above procedure. Further, since both u_1 and u_2 exist in the model of subsystem 1 and only u_1 is in the first community, we have $u^{(1)} = u_1$ and $\bar{u}^{(1)} = u_2$ according to Step 5. Similarly, Eq.(7.6c) and Eq.(7.6d) account for the model of subsystem 2, and we have $u^{(2)} = u_2$ and $\bar{u}^{(2)} = u_1$.

Remark 46. *Step 3 is carried out when $n_c(k+1) = n_c(k)$. This means that the maximum modularity value has been achieved, and further improvement in the modularity value cannot be made (no $\Delta Q > 0$ can be achieved via moving a vertex to any other community). In this case, if the number of communities is greater than p, we need to further combine communities although the modularity value will not increase any more. Step 3 aims to reduce the number of communities by one in each iteration with minimal decrease in the modularity value, until the requirement on the number of subsystems (communities) is satisfied.*

7.2.5 Structure validity test

Following the algorithm in the previous section, we can get candidates of subsystem configuration. We need to perform a structure validity test to confirm that the subsystem configuration candidates are proper. Specifically, this test is to verify that all the input variables placed in community i are included in the input vector $u^{(i)}$ of the corresponding subsystem.

For subsystem i, one implication is that vector $u^{(i)}$ needs to contain all the manipulated input variables in the corresponding community. This requires that the available input variables in the model of subsystem i formed based on the procedure proposed in Section 7.2.4 should include all the input variables assigned to the corresponding community. Therefore, we use this test to verify that the configured subsystem model does not violate the structure of the corresponding community.

7.2.6 Subsystem observability and stabilizability test

The obtained subsystem configurations with maximal modularity values are also required to satisfy observability (or detectability) and controllability (or stabilizability) criteria. After the obtained subsystem configurations pass the "structure validity test", subsystem observability and stabilizability are examined. For nonlinear systems, it can be difficult to check observability and stabilizability directly when the scale of the nonlinear system is not small.

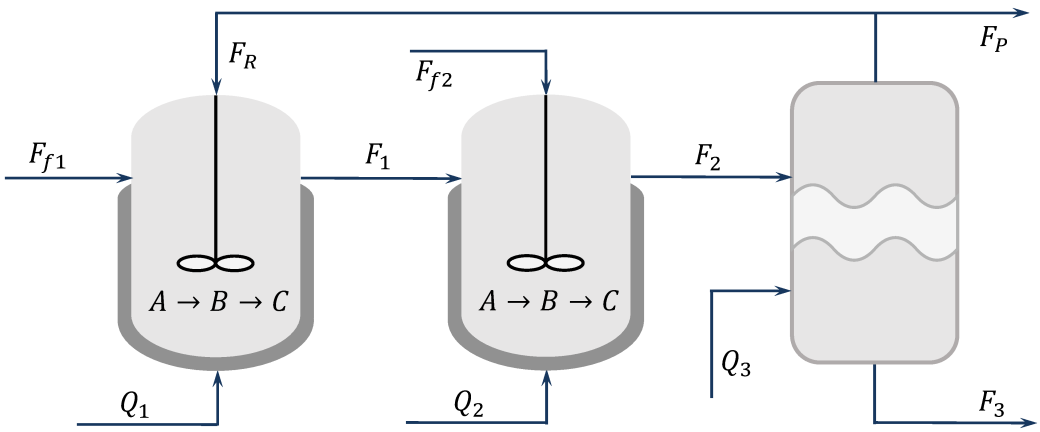

Figure 7.2: A schematic of the reactor-separator process

An alternative is to perform successive linearization of the nonlinear system at different points along typical operating trajectories, and then check the observability and stabilizability of the linearized systems at different points instead.

Remark 47. *By executing the procedure in Section 7.2.3, several candidate subsystem configurations can be obtained. If any of the two tests introduced in Section 7.2.5 and Section 7.2.6 is not passed by the configured subsystems, the subsystem configuration is discarded. In this case, we consider other possible subsystem configurations with the next largest modularity value and perform the above two tests again until we find an appropriate configuration.*

7.3 Application to a reactor-separator example

In this section, a benchmark chemical process example is used to illustrate the proposed method. The process consists of three vessels: two well-mixed non-isothermal continuous-stirred tank reactors (CSTR) and one separator that are connected in series (as shown in Figure 7.2). This process has been widely used as an illustrative example in distributed state estimation [15], distributed control [147, 5, 148], and subsystem decomposition problems [103, 102].

In this process, two first-order reactions take place in the two reactors; that is $A \rightarrow B$ and $B \rightarrow C$, where A is the reactant, B is the desired product while C is the side product.

A dynamic model of 12 differential equations is established to describe the dynamics of the process. The model, the definitions of the variables involved in the model and a more detailed description of the process can be found in [5, 103]. The values of the model parameters used in the current work are the same as those in [102, 103]. This process has 12 state variables, including x_{Ai} (the mass fraction of material A in the i-th vessel), x_{Bi} (the molar fraction of material A in the i-th vessel), T_i (the temperature in the i-th vessel), V_i (the liquid holdup volume in the i-th vessel), $i = 1, 2, 3$. The manipulated inputs include the flow rates of the feed streams to the first and the second reactor (F_{f1} and F_{f2}, respectively), the flow rates of the effluent streams of the three vessels (F_i, $i = 1, 2, 3$), the recycle flow rate (i.e., F_R) and the heating inputs to the three vessels (Q_i, $i = 1, 2, 3$). It is assumed that the mass fractions are not measured and only temperatures and holdup volumes (T_i, V_i, $i = 1, 2, 3$) are measured online. Let us consider that we want to decompose the entire system to 3 subsystems (i.e., $q = 3$).

Subsystem decomposition of this process was also studied in [102, 132, 103]. However, in [102, 132, 103], subsystem decomposition of this process was investigated without explicitly considering distributed state estimation.

Table 7.1: Subsystem decompositions for the reactor-separator process using the method in [132, 103]

Decomposition A.1 (*Decomposition 1 in [103]*)			
Subsystem	States	Manipulated inputs	Available output measurements
1	V_1, T_1, x_{A1}, x_{B1}	F_{f1}, F_R, Q_1	T_1, V_1
2	V_2, T_2, x_{A2}, x_{B2}	F_{f2}, F_1, Q_2	T_2, V_2
3	V_3, T_3, x_{A3}, x_{B3}	F_2, F_3, Q_3	T_3, V_3
Decomposition A.2 (*Decomposition 2 in [103]*)			
Subsystem	States	Manipulated inputs	Available output measurements
1	V_1, T_1, x_{A1}, x_{B1}	F_{f1}, Q_1	T_1, V_1
2	V_2, T_2, x_{A2}, x_{B2}	F_{f2}, F_1, Q_2	T_2, V_2
3	V_3, $\boldsymbol{x_{A3}}$, $\boldsymbol{x_{B3}}$	F_2, F_3, F_R	V_3
4	T_3	Q_3	T_3
Decomposition A.3 (*Decomposition 4 in [103]*)			
Subsystem	States	Manipulated inputs	Available output measurements
1	$\boldsymbol{x_{A1}}$, $\boldsymbol{x_{A2}}$, $\boldsymbol{x_{A3}}$, $\boldsymbol{x_{B1}}$, $\boldsymbol{x_{B2}}$, $\boldsymbol{x_{B3}}$	F_{f1}, F_{f2}, F_R	n/a
2	V_1, V_2, V_3	F_1, F_2, F_3	V_1, V_2, V_3
3	T_1, T_2, T_3	Q_1, Q_2, Q_3	T_1, T_2, T_3

* State variables in bold cannot be estimated using the available measurements of the corresponding subsystem

** "n/a" means that no variable is assigned to the corresponding subsystem

Table 7.1 and Table 7.2 show the subsystem decomposition results reported in [103, 132] and [102], respectively. Among all the decompositions, Decomposition A.1 which was reported in [103, 132] is the only one that can be used for simultaneous distributed state estimation and control given the available measurements. In Decomposition A.2, the states x_{A3} and x_{B3} assigned to subsystem 3 cannot be estimated based on the single measurement of V_3. In Decomposition A.3, no subsystem measurement is available, such that the states

Table 7.2: Subsystem decompositions for the reactor-separator process based on [102]

Decomposition A.4 (*First structure recommended by [102]*)			
Subsystem	States	Manipulated inputs	Available output measurements
1	x_{B1}, T_1	F_{f1}, Q_1	T_1
2	x_{B2}, T_2, T_3	F_{f2}, Q_2, Q_3	T_2, T_3
3	V_1, V_2	F_1	V_1, V_2
4	$\boldsymbol{x_{B3}}$, V_3	F_2, F_R, F_3	V_3
Decomposition A.5 (*Second structure recommended by [102]*)			
Subsystem	States	Manipulated inputs	Available output measurements
1	$\boldsymbol{x_{B1}}$, $\boldsymbol{x_{B2}}$	F_{f1}, F_{f2}	n/a
2	T_1, T_2, T_3	Q_1, Q_2, Q_3	T_1, T_2, T_3
3	V_1, V_2	F_1	V_1, V_2
4	x_{B3}, V_3	F_2, F_3, F_R	V_3

* State variables in bold cannot be estimated using the available measurements of the corresponding subsystem

** "n/a" means that no variable is assigned to the corresponding subsystem

assigned to subsystem 1 cannot be estimated. Similarly, certain state(s) in Subsystem 4 of Decomposition A.4 and Subsystem 1 of Decomposition A.5 (highlighted in bold) cannot be estimated due to insufficient subsystem measurement(s). Neither decomposition obtained using [102] can be used for simultaneous distributed state estimation and control given the available measurements. While the method in [132, 103] indeed gives one decomposition (Decomposition A.1 in Table) that can be used for simultaneous distributed state estimation and control, it may fail to give any feasible decompositions for larger-scale processes when distributed state estimation is also considered as demonstrated in the example discussed in the next section.

Next, the proposed method is applied to the process. Two subsystem configurations are

Table 7.3: Subsystem decomposition for the reactor-separator process based on the proposed method

Decomposition A.6 ($Q = 0.4458$)			
Subsystem	States	Manipulated inputs	Available output measurements
1	V_1, T_1, x_{A1}, x_{B1}	F_{f1}, Q_1	T_1, V_1
2	V_2, T_2, x_{A2}, x_{B2}	F_{f2}, F_1, Q_2	T_2, V_2
3	V_3, T_3, x_{A3}, x_{B3}	F_2, F_3, $\boldsymbol{F_R}$, Q_3	T_3, V_3
Decomposition A.7 ($Q = 0.4417$)			
Subsystem	States	Manipulated inputs	Available output measurements
1	V_1, T_1, x_{A1}, x_{B1}	F_{f1}, $\boldsymbol{F_R}$, Q_1	T_1, V_1
2	V_2, T_2, x_{A2}, x_{B2}	F_{f2}, F_1, Q_2	T_2, V_2
3	V_3, T_3, x_{A3}, x_{B3}	F_2, F_3, Q_3	T_3, V_3

found and are presented in Table 7.3. The only difference between the two decompositions is that the recycle flow F_R is assigned to different subsystems for control. Both decompositions capture the physical topology of the process well. This may be because the interaction of the state dynamics within each vessel is relatively strong, while the interaction between the vessels (via effluent and recycle streams) is comparatively weaker. Note that the second decomposition (Decomposition A.7) obtained based on the proposed method is consistent with Decomposition A.1 reported in [103, 132] in terms of state and manipulated input variables.

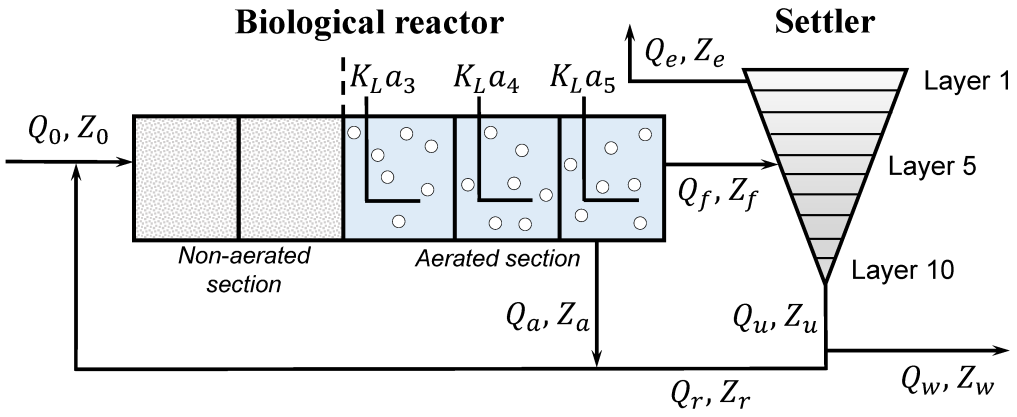

Figure 7.3: A schematic of the wastewater treatment plant

7.4 Application to a wastewater treatment plant

7.4.1 Model description

A wastewater treatment plant (WWTP) described by the Benchmark Simulation Model No.1 (BSM1) is considered. This plant consists of a multi-chamber biological activated sludge reactor and a secondary settler. A schematic of the plant is given in Figure 7.3 [130]. The biological reactor comprises five interconnected chambers: the first two anoxic chambers are the non-aerated section and the remaining three aerobic chambers are the aerated section. In the non-aerated section, pre-denitrification reactions where nitrate is converted into nitrogen take place, while in the aerated section, nitrification reactions take place such that ammonium is oxidized into nitrate [130].

Wastewater enters the plant through the first chamber of the biological reactor at concentration Z_0 and flow rate Q_0, which are time-varying parameters dependent on the current weather condition. A portion of the effluent discharged from the fifth chamber of the reactor enters the settler at flow rate Q_f and concentration Z_f, while the remaining portion is recycled to the first chamber at flow rate Q_a. The secondary settler is constituted by 10 nonreactive layers, the 5-th layer of which is the feed layer. The outlets of the settler are discharged in three ways: (a) the overflow of the settler containing purified water is removed continuously through the top layer of the settler with concentration Z_e and flow rate Q_e; (b)

a portion of the underflow of the settler is sent back to the first chamber at flow rate Q_r and concentration Z_r; (c) the rest of the underflow is discharged through the bottom layer at flow rate Q_w. In the process model, we take into account eight major biological reactions, and 13 major compounds are considered in these reactions. The concentrations of the 13 compounds in the five chambers are the state variables of the model of the biological reactor. The 13 state variables for each chamber are listed in Table 7.4. The model of the settler is established based on mass balances of the sludge considering solid flux due to gravity [133]. Specifically, 8 states are taken into account in the model of each layer of the settler, including S_O, S_{ALK}, S_{NH}, S_{NO}, S_S, S_I, S_{ND} and X. X represents the concentration of suspended solids of a layer, which is the summation of X_S, X_I, X_{B_A}, X_{B_H}, X_P and X_{ND} in the corresponding layer. In the remainder, we use $x_{s_{i,l}}$, $i = 1, \ldots, 8$, to denote the concentrations S_O, S_{ALK}, S_{NH}, S_{NO}, S_S, S_I, S_{ND} and X in l-th layer ($l = 1, \ldots, 10$) of the settler, respectively. In total, 145 states are used to capture the dynamics of the WWTP based on BSM1. A more detailed description, the process model and the model parameters of the WWTP are given in [130].

7.4.2 Manipulated inputs and measured outputs

In this work, four manipulated inputs are taken into account: the flow rate of the recirculation stream (denoted by Q_a) and the oxygen transfer rate in each of the aerobic chambers of the biological reactor (denoted by K_La_3, K_La_4 and K_La_5, respectively).

For WWTP described by the BSM1, we consider that there are 56 measured output variables available for state estimation. In each chamber of the biological reactor, the concentrations of dissolved oxygen, free and saline ammonia (i.e., NH3 and $NH4^+$), nitrate and nitrate nitrogen, alkalinity, chemical oxygen demand (COD), filtered chemical oxygen demand (COD_f), biological oxygen demand (BOD) and the concentration of suspended solids can be measured online [123, 130]. In the settler, the states of the top layer and the bottom layer are measured online. The output measurements in each chamber of the biological

Table 7.4: State variables of the i-th chamber ($i = 1, \ldots, 5$) of the biological reactor

State	Definition	Unit	Notation
X_I	inert particulate organic matter	g COD $\cdot$ m^{-3}	$x_{c_{i,1}}$
S_I	inert soluble organic matter	g COD $\cdot$ m^{-3}	$x_{c_{i,2}}$
X_s	slowly biodegradable and soluble substrate	g COD $\cdot$ m^{-3}	$x_{c_{i,3}}$
S_s	readily biodegradable and soluble substrate	g COD $\cdot$ m^{-3}	$x_{c_{i,4}}$
X_P	particulate generated from decay of organisms	g COD $\cdot$ m^{-3}	$x_{c_{i,5}}$
X_{ND}	particulate biodegradable organic nitrogen	g N $\cdot$ m^{-3}	$x_{c_{i,6}}$
X_{B_A}	biomass of active autotrophs	g COD $\cdot$ m^{-3}	$x_{c_{i,7}}$
X_{B_H}	biomass of active heterotrophs	g COD $\cdot$ m^{-3}	$x_{c_{i,8}}$
S_{NO}	nitrite nitrogen and nitrate	g N $\cdot$ m^{-3}	$x_{c_{i,9}}$
S_{NH}	free and saline ammonia	g N $\cdot$ m^{-3}	$x_{c_{i,10}}$
S_O	dissolved oxygen	g (-COD) $\cdot$ m^{-3}	$x_{c_{i,11}}$
S_{ALK}	alkalinity	mol $\cdot$ m^{-3}	$x_{c_{i,12}}$
S_{ND}	biodegradable and soluble organic nitrogen	g N $\cdot$ m^{-3}	$x_{c_{i,13}}$

reactor are described in Table 7.5. In this table, $y_{c_{i,l}}$ represents the l-th measurement of chamber i, $i = 1, \ldots, 5$, $l = 1, \ldots, 8$. The output measurements in the top/bottom layer of the reactor are shown in Table 7.6. In this table, $y_{s_{i,l}}$ denotes the measurement of state $x_{s_{i,l}}$ ($i = 1, \ldots, 8$) in the top layer (when $l = 1$) or the bottom layer (when $l = 10$).

7.4.3 Subsystem decomposition

Now, we apply the proposed method to decompose the WWTP into subsystems for simultaneous distributed state estimation and distributed control. There are 205 vertices (145 states, 4 manipulated inputs and 56 measured outputs). Suppose that we would like to decompose the plant into six subsystems (i.e., $q = 6$). The top three decomposition candidates are given in Table 7.7.

In Decomposition B.1, the modularity value is large ($Q = 0.6068$), which indicates a very good subsystem structure. Subsystem 1 contains all the states and measurements of the

Table 7.5: The output measurements in the i-th chamber ($i = 1, \ldots, 5$) of the biological reactor

Output measurement variable	Expression in the form of system state(s)	Symbol
concentration of dissolved oxygen	S_O	$y_{c_{i,1}}$
concentration of nitrate and nitrite nitrogen	S_{NO}	$y_{c_{i,2}}$
concentration of NH3 and $NH4^+$	S_{NH}	$y_{c_{i,3}}$
concentration of alkalinity	S_{ALK}	$y_{c_{i,4}}$
COD	$S_S + S_I + X_S + X_I + X_{B_A} + X_{B_H}$	$y_{c_{i,5}}$
concentration of suspended solids	$X_S + X_I + X_{B_A} + X_{B_H} + X_P + X_{ND}$	$y_{c_{i,6}}$
COD_f	$S_S + S_I$	$y_{c_{i,7}}$
BOD	$S_S + X_S$	$y_{c_{i,8}}$

Table 7.6: Output measurements in the top layer ($l = 1$) and bottom layer ($l = 10$) of the settler

Output measurement variable	Expression in the form of system state(s)	Symbol
concentration of dissolved oxygen	S_O	$y_{s_{1,l}}$
concentration of alkalinity	S_{ALK}	$y_{s_{2,l}}$
concentration of NH3 and $NH4^+$	S_{NH}	$y_{s_{3,l}}$
concentration of nitrate and nitrite nitrogen	S_{NO}	$y_{s_{4,l}}$
concentration of readily biodegradable and soluble substrate	S_S	$y_{s_{5,l}}$
concentration of inert soluble organic matter	S_I	$y_{s_{6,l}}$
concentration of biodegradable and soluble organic nitrogen	S_{ND}	$y_{s_{7,l}}$
concentration of suspended solids	$X_S + X_I + X_{B_A} + X_{B_H} + X_P + X_{ND}$	$y_{s_{8,l}}$

first and the fifth chambers of the reactor, the states associated with the concentration of X in the ten layers of the settler, the measured concentrations of X in the top and bottom

layers, , and the manipulated input $K_L a_5$. Subsystem 2 contains the state and measurement variables of chamber 2 and the manipulated input variable Q_a. Subsystem 3 to subsystem 4 each contains the state, measurement and manipulated variables of chamber 3 to chamber 4, respectively. As stated in [130, 149, 150], Q_a is often used to control the nitrate level in the second chamber and $K_L a$ is to control the dissolved oxygen in the corresponding chamber. Therefore, the allocation of the manipulated inputs aligns with common control strategies for the WWTP. Subsystem 5 contains the states associated with S_I, S_S, S_{NH} and S_{ND} in the ten layers of the settler and the associated measurements, while Subsystem 6 contains the states and the measurements associated with S_O, S_{NO} and S_{ALK} in the settler. No manipulated inputs are included in either Subsystem 5 or Subsystem 6. Note that this does not deteriorate local stabilizability of the two subsystems. This is because Subsystem 5 (Subsystem 6) has linear local dynamics when treating its interacting dynamics as known inputs, and the local dynamics of Subsystem 5 (Subsystem 6) is open-loop asymptotically stable.

The modularity values for the other two decompositions are slightly smaller. As shown in Table 7.7, these two decompositions (Decomposition B.2 and B.3) are the same as Decomposition B.1 except in the last two subsystems, and serve as backup options.

Next, the method in [132, 103] is also applied to the WWTP. In the construction of the adjacency matrix, the states and the manipulated inputs are considered as vertices while the measurement outputs are excluded. A structure of 14 communities leads to the maximal achievable modularity value (0.5784) based on this method. However, there are several subsystems that are not assigned with any available output measurements, which leads to unobservability of the subsystems. Totally, nine subsystem structures with different number of subsystems are obtained at different levels of sub-divisions after fine-tuning. Given the available measured outputs described in Section 7.4.2, none of these candidate decompositions can pass the subsystem observability test. The above results show the necessity of using the method proposed in this work for subsystem decomposition for simultaneous distributed

state estimation and control.

Table 7.7: Six-subsystem structures for the WWTP

Decomposition B.1 ($Q = 0.6068$)			
Subsystem	States	Manipulated inputs	Output measurements
1	$x_{c_{1,i}}$, $x_{c_{5,i}}$ $(i = 1, \ldots, 13)$, $x_{s_{8,j}}$ $(j = 1, \ldots, 10)$	$K_L a_5$	$y_{c_{1,i}}$, $y_{c_{5,i}}$ $(i = 1, \ldots, 8)$, $y_{s_{8,1}}$, $y_{s_{8,10}}$
2	$x_{c_{2,i}}$ $(i = 1, \ldots, 13)$	Q_a	$y_{c_{2,i}}$ $(i = 1, \ldots, 8)$
3	$x_{c_{3,i}}$ $(i = 1, \ldots, 13)$	$K_L a_3$	$y_{c_{3,i}}$ $(i = 1, \ldots, 8)$
4	$x_{c_{4,i}}$ $(i = 1, \ldots, 13)$	$K_L a_4$	$y_{c_{4,i}}$ $(i = 1, \ldots, 8)$
5	$x_{s_{3,i}}$, $x_{s_{5,i}}$, $x_{s_{6,i}}$, $x_{s_{7,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{3,1}}$, $y_{s_{3,10}}$, $y_{s_{5,1}}$, $y_{s_{5,10}}$ $y_{s_{6,1}}$, $y_{s_{6,10}}$, $y_{s_{7,1}}$, $y_{s_{7,10}}$
6	$x_{s_{1,i}}$, $x_{s_{2,i}}$, $x_{s_{4,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{1,1}}$, $y_{s_{1,10}}$, $y_{s_{2,1}}$, $y_{s_{2,10}}$ $y_{s_{4,1}}$, $y_{s_{4,10}}$
Decomposition B.2 ($Q = 0.60544$)			
Subsystem	States	Manipulated inputs	Output measurements
1	$x_{c_{1,i}}$, $x_{c_{5,i}}$ $(i = 1, \ldots, 13)$, $x_{s_{8,j}}$ $(j = 1, \ldots, 10)$	$K_L a_5$	$y_{c_{1,i}}$, $y_{c_{5,i}}$ $(i = 1, \ldots, 8)$, $y_{s_{8,1}}$, $y_{s_{8,10}}$
2	$x_{c_{2,i}}$ $(i = 1, \ldots, 13)$	Q_a	$y_{c_{2,i}}$ $(i = 1, \ldots, 8)$
3	$x_{c_{3,i}}$ $(i = 1, \ldots, 13)$	$K_L a_3$	$y_{c_{3,i}}$ $(i = 1, \ldots, 8)$
4	$x_{c_{4,i}}$ $(i = 1, \ldots, 13)$	$K_L a_4$	$y_{c_{4,i}}$ $(i = 1, \ldots, 8)$
5	$x_{s_{3,i}}$, $x_{s_{6,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{3,1}}$, $y_{s_{3,10}}$, $y_{s_{6,1}}$, $y_{s_{6,10}}$
6	$x_{s_{1,i}}$, $x_{s_{2,i}}$, $x_{s_{4,i}}$, $x_{s_{5,i}}$, $x_{s_{7,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{1,1}}$, $y_{s_{1,10}}$, $y_{s_{2,1}}$, $y_{s_{2,10}}$, $y_{s_{4,1}}$, $y_{s_{4,10}}$ $y_{s_{5,1}}$, $y_{s_{5,10}}$, $y_{s_{7,1}}$, $y_{s_{7,10}}$
Decomposition B.3 ($Q = 0.60544$)			
Subsystem	States	Manipulated inputs	Output measurements
1	$x_{c_{1,i}}$, $x_{c_{5,i}}$ $(i = 1, \ldots, 13)$, $x_{s_{8,j}}$ $(j = 1, \ldots, 10)$	$K_L a_5$	$y_{c_{1,i}}$, $y_{c_{5,i}}$ $(i = 1, \ldots, 8)$, $y_{s_{8,1}}$, $y_{s_{8,10}}$
2	$x_{c_{2,i}}$ $(i = 1, \ldots, 13)$	Q_a	$y_{c_{2,i}}$ $(i = 1, \ldots, 8)$
3	$x_{c_{3,i}}$ $(i = 1, \ldots, 13)$	$K_L a_3$	$y_{c_{3,i}}$ $(i = 1, \ldots, 8)$
4	$x_{c_{4,i}}$ $(i = 1, \ldots, 13)$	$K_L a_4$	$y_{c_{4,i}}$ $(i = 1, \ldots, 8)$
5	$x_{s_{5,i}}$, $x_{s_{7,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{3,1}}$, $y_{s_{3,10}}$, $y_{s_{6,1}}$, $y_{s_{6,10}}$
6	$x_{s_{1,i}}$, $x_{s_{2,i}}$, $x_{s_{3,i}}$, $x_{s_{4,i}}$, $x_{s_{6,i}}$ $(i = 1, \ldots, 10)$	n/a	$y_{s_{1,1}}$, $y_{s_{1,10}}$, $y_{s_{2,1}}$, $y_{s_{2,10}}$, $y_{s_{3,1}}$, $y_{s_{3,10}}$ $y_{s_{4,1}}$, $y_{s_{4,10}}$, $y_{s_{6,1}}$, $y_{s_{6,10}}$

Chapter 8

Conclusions and Future work

8.1 Conclusions

In this thesis, we have proposed systematic approaches on subsystem decomposition and distributed state estimation for nonlinear process networks.

In Chapter 2, two-time-scale nonlinear systems was considered and a distributed state estimation method was proposed. Specifically, the nonlinear system was decomposed into a fast system and several slow subsystems. Local MHE estimators were designed for the slow subsystems and the fast system. The estimator for the fast system is not required to send out any information. The convergence and ultimate boundedness of the estimation error norm has been proved. The effectiveness of the proposed approach was illustrated using a separator-reactor process example.

In Chapter 3, general nonlinear systems that can be decomposed into smaller subsystems were considered. A systematic method to design distributed state estimation networks was proposed. Specifically, decentralized local estimators were assumed to exist for the subsystems. A compensator that compensates for the dynamics of subsystem interaction was designed and an augmented estimator for the corresponding subsystem was formed based on the compensator. The augmented estimators were connected together via discrete-time in-

formation exchange to form a distributed estimation scheme. Sufficient conditions were given for the convergence and boundedness of the estimation error. Several application examples were used to demonstrate the proposed approach.

In Chapter 4, distributed output-feedback FDI for nonlinear cascade processes was addressed. First, a distributed state estimation was designed based on the distributed state estimation method proposed in Chapter 3. The convergence to zero of the estimation error of the distributed estimation system in the fault-free context was proved. Then, a residual generator was developed for each subsystem to generate residual signals for FDI. A distributed FDI mechanism applicable to both actuator faults and sensor faults was developed by evaluating the subsystem residual signals. Potential faults that can be detected and isolated by the developed mechanism were also characterized. A froth flotation process was used to verify the effectiveness of the developed FDI mechanism.

In Chapter 5, a subsystem decomposition procedure based on structural closeness for distributed state estimation of nonlinear processes was proposed. Observability test of the entire system, identification of observable states given each measured output, relative degree analysis and sensitivity analysis between output measurements and system states were included in the proposed procedure. Process examples were used to demonstrate the applicability of the proposed techniques/procedure. The decomposition results for the examples are reasonable and consistent with process topology.

In Chapter 6, the distributed state estimation problem for the wastewater treatment plant was considered. Subsystem decomposition and distributed estimation design were systematically addressed in this chapter. Several key steps in the decomposition approach proposed in Chapter 5 were taken advantage of. The plant was decomposed into smaller subsystems. Then, a distributed MHE scheme was proposed for the WWTP. Innovation-based iterative evaluation was also considered to reduce its computational complexity. Simulations were carried out under different weather conditions. The proposed scheme can provide good state estimates in different weather. Extensive comparisons with other state estimation methods

were made to illustrate the advantages of the proposed method.

In Chapter 7, a systematic procedure was proposed for subsystem decomposition of general nonlinear process networks for simultaneous distributed state estimation and distributed control. The proposed procedure resorts to community structure detection based on modularity maximization. A fast folding algorithm that approximately maximizes the modularity was used in the proposed procedure to find candidate subsystem configurations. Two criteria were proposed to find valid solutions from the candidate subsystem configurations for simultaneous distributed estimation and control. We introduced two chemical process examples of different complexities to illustrate the effectiveness and applicability of the proposed approach. The proposed method was also compared with the existing methods that focus primarily on distributed control through the application to the two examples. The results showed the advantage of the proposed procedure in terms of addressing simultaneous distributed state estimation and control.

8.2 Future work

- *Event-triggered distributed moving horizon estimation.* As discussed in this thesis, distributed moving horizon estimation (DMHE) has attracted great research attention and has been considered to be a very effective solution to the state estimation problem of large-scale processes. However, computation and communication resource required for the implementation of a DMHE scheme is non-negligible and could be unaffordable. Based on this consideration, it would be favorable to incorporate event-triggered conditions in DMHE designs to reduce the evaluation and the communication frequency. Event-triggered conditions should be properly determined. In addition, rigorous theoretical analysis of the stability of the estimation error dynamics needs to be carried out for event-trigged DMHE methods, which have not been available in existing literature.

- *Distributed output-feedback fault detection and isolation for general nonlinear systems.*

This thesis has investigated the problem of distributed fault detection and isolation using output-feedback. However, this work focused on cascade systems but not general nonlinear systems. Sometimes one has to use a general nonlinear model to describe the dynamics of a chemical process (e.g., when recycle streams exist). However, in a distributed framework, the isolation of a possible fault is more challenging when only output-feedback information is available. The design of an effective distributed FDI mechanism for general nonlinear systems and the selection of appropriate detection and isolation criteria are part of our future work.

- *Distributed moving horizon estimation subject to unreliable communication/asynchronous measurements.* The implementation of a common DMHE design relies on a communication network that coordinates information exchange among local estimators. From a practical point of view, there are some practice issues that may be encountered in a communication network, including time-varying delays, data loss and asynchronous measurements. In future work, it is worth investigating how to propose an effective DMHE scheme to handle the above imperfections with respect to communication.

Bibliography

[1] P. D. Christofides, R. Scattolini, D. Muñoz de la Peña, and J. Liu. Distributed model predictive control: A tutorial review and future research directions. *Computers & Chemical Engineering*, 51:21–41, 2013.

[2] P. Daoutidis, M. Zachar, and S. S. Jogwar. Sustainability and process control: A survey and perspective. *Journal of Process Control*, 44:184–206, 2016.

[3] J. B. Rawlings and B. T. Stewart. Coordinating multiple optimization-based controllers: New opportunities and challenges. *Journal of Process Control*, 18:839–845, 2008.

[4] R. Scattolini. Architectures for distributed and hierarchical model predictive control - A review. *Journal of Process Control*, 19:723–731, 2009.

[5] B. T. Stewart, A. N. Venkat, J. B. Rawlings, S. J. Wright, and G. Pannocchia. Coorperative distributed model predictive control. *Systems and Control Letters*, 59(8):460–469, 2010.

[6] M. Farina, G. Betti, L. Giulioni, and R. Scattolini. An approach to distributed predictive control for tracking-theory and applications. *IEEE Transactions on Control Systems Technology*, 22(4):1558–1566, 2014.

[7] W. Qi, J. Liu, and P. D. Christofides. Distributed supervisory predictive control of distributed wind and solar energy systems. *IEEE Transactions on Control Systems Technology*, 21(2):504–512, 2013.

[8] E. Camponogara, D. Jia, B. H. Krogh, and S. Talukdar. Distributed model predictive control. *IEEE Control Systems Magazine*, 22(1):44–52, 2002.

[9] A. N. Venkat, I. A. Hiskens, J. B. Rawlings, and S. J. Wright. Distributed mpc strategies with application to power system automatic generation control. *IEEE Transactions on Control Systems Technology*, 16(6):1192–1206, 2008.

[10] J. Liu, D. Muñoz de la Peña, and P. D. Christofides. Distributed model predictive control of nonlinear systems subject to asynchronous and delayed measurements. *Automatica*, 46(1):52–61, 2010.

[11] S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Consensus based overlapping decentralized estimation with missing observations and communication faults. *Automatica*, 45(6):1397–1406, 2009.

[12] A. Haber and M. Verhaegen. Moving horizon estimation for large-scale interconnected systems. *IEEE Transactions on Automatic Control*, 58(11):2834-2847, 2013.

[13] A. Alessandri, M. Baglietto, G. Battistelli and V. Zavala. Advances in moving horizon estimation for nonlinear systems. In *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, GA, 5681 - 5688, 2010.

[14] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Moving horizon estimation for distributed nonlinear systems with application to cascade river reaches. *Journal of Process Control*, 21(5):767–774, 2011.

[15] J. Zhang and J. Liu. Distributed moving horizon estimation for nonlinear systems with bounded uncertainties. *Journal of Process Control*, 23(9):1281–1295, 2013.

[16] J. Zhang and J. Liu. Lyapunov-based MPC with robust moving horizon estimation and its triggered implementation. *AIChE Journal*, 59(11): 4273-4286, 2013.

[17] V. W. Weekman and D. M. Nace. Kinetics of catalytic cracking selectivity in fixed, moving, and fluid bed reactors. *AIChE Journal*, 16:397–404, 1970.

[18] I. S. Han and C. B. Chung. Dynamic modeling and simulation of a fluidized catalytic cracking process. Part II: Property estimation and simulation. *Chemical Engineering Science*, 56(5):1973–1990, 2001.

[19] P. Kokotovic, H. K. Khalil, and J. O'Reilly. *Singular Perturbation Methods in Control: Analysis and Design*. Academic Press, London, 1986.

[20] P. D. Christofides. Robust output feedback control of nonlinear singularly perturbed systems. *Automatica*, 36(1):45–52, 2000.

[21] X. Chen, M. Heidarinejad, J. Liu, D. Muñoz de la Peña, and P. D. Christofides. Model predictive control of nonlinear singularly perturbed systems: Application to a large-scale process network. *Journal of Process Control*, 21(9):1296–1305, 2011.

[22] A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, and J. Ragot. State estimation of two-time scale multiple models. application to wastwater treatment plant. *Control Engineering Practice*, 19(11):1354–1362, 2011.

[23] H. K. Khalil. *Nonlinear systems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[24] J. Liu. Moving horizon state estimation for nonlinear systems with bounded uncertainties. *Chemical Engineering Science*, 93:376–386, 2013.

[25] C. Kravaris, J. Hahn and Y. Chu. Advances and selected recent developments in state and parameter estimation. *Computers and Chemical Engineering*, 51(5):111-123, 2013.

[26] J. P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems: applications to bioreactors. *IEEE Transactions on Automatic Control*, 37(6):875–880, 1992.

[27] N. Boizot, E. Busvelle and J.-P. Gauthier. An adaptive high-gain observer for nonlinear systems. *Automatica*, 46(9):1483-1488, 2010.

[28] J. H. Ahrens and H. K. Khalil. High-gain observers in the presence of measurement noise: A switched-gain approach. *Automatica*, 45(4):936-943, 2009.

[29] D. Dochain. State and parameter estimation in chemical and biochemical processes: A tutorial. *Journal of Process Control*, 13(8):801-818, 2003.

[30] G. Ciccarella, M. Dalla Mora, and A. Germani. A Luenberger-like observer for nonlilear systems. *International Jounal of Control*, 57(3):537–556, 1993.

[31] B. Açıkmeşe, M. Mandić, and J. L. Speyer. Decentralized observers with consensus filters for distributed discrete-time linear systems. *Automatica*, 50(4):1037–1052, 2014.

[32] Y. Hong, G. Chen, and L. Bushnell. Distributed observers design for leader-following control of multi-agent networks. *Automatica*, 44(3):846–850, 2008.

[33] U. A. Khan and J. M. F. Moura. Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, 56(10):4919–4935, 2008.

[34] R. Vadigepalli and F. J. Doyle III. A distributed state estimation and control algorithm for plantwide processes. *IEEE Transactions on Control Systems Technology*, 11(1):119–127, 2003.

[35] R. Olfati-Saber and P. Jalalkamali. Coupled distributed estimation and control for mobile sensor networks. *IEEE Transactions on Automatic Control*, 57(10):2609–2614, 2012.

[36] S. Roshany-Yamchi, M. Cychowski, R. R. Negenborn, B. De Schutter, K. Delaney, and J. Connell. Kalman filter-based distributed predictive control of large-scale multi-rate systems: Application to power networks. *IEEE Transactions on Control Systems Technology*, 21(1):27–39, 2013.

[37] G. Battistelli and L. Chisci. Stability of consensus extended Kalman filter for distributed state estimation. *Automatica*, 68:169–178, 2016.

[38] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Distributed moving horizon estimation for linear constrained systems. *IEEE Transactions on Automatic Control*, 55(11):2462–2475, 2010.

[39] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Moving-horizon partition-based state estimation of large-scale systems. *Automatica*, 46(5):910–918, 2010.

[40] R. Schneider, H. Scheu, and W. Marquardt. An iterative partition-based moving horizon estimator for large-scale linear systems. In *proceedings of the European Control Conference*, pages 2621–2626, Zürich, Switzerland, July 2013.

[41] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Distributed moving horizon estimation for nonlinear constrained systems. *International Journal of Robust and Nonlinear Control*, 22:123–143, 2012.

[42] X. Yin and J. Liu. Distributed moving horizon state estimation of two-time-scale nonlinear systems. *Automatica*, 79:152–161, 2017.

[43] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou. Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, 57(2):275–290, 2012.

[44] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou. Distributed fault diagnosis with overlapping decompositions: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, 54(4):794–799, 2009.

[45] X. Zhang and Q. Zhang. Distributed fault diagnosis in a class of interconnected nonlinear uncertain systems. *International Journal of Control*, 85(11):1644–1662, 2012.

[46] Q. Zhang and X. Zhang. Distributed sensor fault diagnosis in a class of interconnected nonlinear uncertain systems. *Annual Reviews in Control*, 37(1):170–179, 2013.

[47] X. Yin, J. Zeng, and J. Liu. From decentralized to distributed state estimation. In *Proceedings of American Control Conference*, pages 1904–1909, Seattle, WA, 2017.

[48] L. Ji, J. B. Rawlings, W. Hu, A. Wynn, and M. Diehl. Robust stability of moving horizon estimation under bounded disturbances. *IEEE Transactions on Automatic Control*, 61(11):3509–3514, 2016.

[49] M. A. Müller. Nonlinear moving horizon estimation in the presence of bounded disturbances. *Automatica*, 79:306–314, 2017.

[50] E. D. Sontag and Y. Wang. Output-to-state stability and detectability of nonlinear systems. *Systems & Control Letters*, 29(5):279–290, 1997.

[51] Y. Sun and N. H. El-Farra. Quasi-decentralized model-based networked control of process systems. *Computers and Chemical Engineering*, 32:2016–2029, 2008.

[52] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Dover Publications, New York, 1972.

[53] D. Nešić, A Teel, and P. Kokotovic. Sufficient conditions for stabilization of sampled-data nonlinear systems via discrete time approximations. *Systems and Control Letters*, 38:259–270, 1999.

[54] P. Tabuada and X. Wang. Preliminary results on state-trigered scheduling of stabilizing control tasks. In *Proceedings of IEEE Conference on Decision and Control*, pages 282–287, 2006.

[55] J. P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems: applications to bioreactors. *IEEE Transactions on Automatic Control*, 37(6):875–880, 1992.

[56] N. Boizot, E. Busvelle, and J.-P. Gauthier. An adaptive high-gain observer for nonlinear systems. *Automatica*, 46(9):1483–1488, 2010.

[57] N. Kazantzis and C. Kravaris. Nonlinear observer design using Lyapunov's auxiliary theorem. *Systems & Control Letters*, 34:241–247, 1998.

[58] M. Fang, H. Kodamana, B. Huang, and N. Sammaknejad. A novel approach to process operating mode diagnosis using conditional random fields in the presence of missing data. *Computers & Chemical Engineering*, 111:149–163, 2018.

[59] M. Rashedi. *Distributed adaptive high-gain extended Kalman filtering for nonlinear systems*. PhD thesis, Department of Chemical and Materials Engineering, University of Alberta, Alberta, AB, 2016.

[60] G. S Canright, C. H. Brown, G. O. Allgood Jr, and W. R. Hamel. *Dynamic modeling and control analysis of froth floatation and clean coal filtration as applied to coal beneficiation*. Technical report, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1981.

[61] X. Yin, K. Arulmaran, J. Liu, and J. Zeng. Subsystem decomposition and configuration for distributed state estimation. *AIChE Journal*, 62(6):1995–2003, 2016.

[62] M. Du and P. Mhaskar. Isolation and handling of sensor faults in nonlinear systems. *Automatica*, 50:1066–1074, 2014.

[63] B. Huang. Detection of abrupt changes of total least squares models and application in fault detection. *IEEE Transactions on Control Systems Technology*, 9(2):357–367, 2001.

[64] F. Boem, R. M. G. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou. A distributed networked approach for fault detection of large-scale systems. *IEEE Transactions on Automatic Control*, 62:18–33, 2017.

[65] J. C. Hoskins, K. M. Kaliyur, and D. M. Himmelblau. Fault diagnosis in complex chemical plants using artificial neural networks. *AIChE Journal*, 37:137–141, 1991.

[66] X. Xu, B. Huang, and S. Dubljevic. Optimal continuous-time state estimation for linear finite and infinite-dimensional chemical process systems with state constraints. *Journal of Process Control*, 35:127–142, 2015.

[67] P. M. Frank. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy. *Automatica*, 26:459–474, 1990.

[68] C. P. Tan and C. Edwards. Sliding mode observers for detection and reconstruction of sensor faults. *Automatica*, 38:1815–1821, 2002.

[69] H. Alwi, C. Edwards, and C. P. Tan. Sliding mode estimation schemes for incipient sensor faults. *Automatica*, 45:1679–1685, 2009.

[70] C. D. Persis and A. Isidori. A geometric approach to nonlinear fault detection and isolation. *IEEE Transactions on Automatic Control*, 46:853–865, 2001.

[71] P. Mhaskar, C. McFall, A. Gani, P. D. Christofides, J. F. Davis. Isolation and handling of actuator faults in nonlinear systems. *Automatica*, 44:53–62, 2008.

[72] X. Yin, K. Arulmaran, and J. Liu. Subsystem decomposition for distributed state estimation of nonlinear systems. In *Proceedings of American Control Conference*, pages 5569–5574, Boston, MA, 2016.

[73] X. Yin and J. Liu. Subsystem decomposition of process networks for simultaneous distributed state estimation and control. *AIChE Journal*, submitted.

[74] Y. Zhang and S. J. Qin. Improved nonlinear fault detection technique and statistical analysis. *AIChE Journal*, 54:3207–3220, 2008.

[75] X. Zhang, M. M. Polycarpou, and T. Parisini. Fault diagnosis of a class of nonlinear uncertain systems with lipschitz nonlinearities using adaptive estimation. *Automatica*, 46:290–299, 2010.

[76] X. Zhang. Sensor bias fault detection and isolation in a class of nonlinear uncertain systems using adaptive estimation. *IEEE Transactions on Automatic Control*, 56:1220–1226, 2011.

[77] A. M. Pertew, H. J. Marquez, and Q. Zhao. LMI-based sensor fault diagnosis for nonlinear Lipschitz systems. *Automatica*, 43:1464–1469, 2007.

[78] C. Keliris, M. M. Polycarpou, and T. Parisini. A robust nonlinear observer-based approach for distributed fault detection of input-output interconnected systems. *Automatica*, 53:408–415, 2015.

[79] S. X. Ding, P. Zhang, C. Chihaia, W. Li, Y. Wang, and E. Ding. Advanced design scheme for fault tolerant distributed networked control systems. In *Proceedings of the 17th World Congress of the International Federation of Automatic Control*, Seoul, Korea, 13569–13574, 2008.

[80] I. Shames, A. M. H. Teixeira, H. Sandberg, K. H. Johansson. Distributed fault detection for interconnected second-order systems. *Automatica*, 47:2757–2764, 2011.

[81] R. M. G. Ferrari, T. Parisini, and M. M. Polycarpou. Distributed fault detection and isolation of large-scale discrete-time nonlinear systems: An adaptive approximation approach. *IEEE Transactions on Automatic Control*, 57:275–290, 2012.

[82] C. Keliris, M. M. Polycarpou, and T. Parisini. A distributed fault detection filtering approach for a class of interconnected continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 58:2032–2047, 2013.

[83] R. J. Patton, C. Kambhampati, A. Casavola, P. Zhang, S. X. Ding , and D. Sauter. A generic strategy for fault-tolerance in control systems distributed over a network. *European Journal of Control*, 13:280–296, 2007.

[84] G. S Canright, C. H. Brown, G. O. Allgood Jr, and W. R. Hamel. *Dynamic modeling and control analysis of froth floatation and clean coal filtration as applied to coal benefictiation*. Technical report, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 1981.

[85] X. Yin and J. Liu. Distributed fault detection and isolation of nonlinear systems using output feedback. In *International Symposium on Advanced Control of Industrial processes*, Taipei, 547–552, May 2017.

[86] T. Ahmed-Ali, E. Cherrier, and F. Lamnabhi-Lagarrigue. Cascade high gain predictors for a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 57:221–226, 2012.

[87] A. N. Atassi and H. K. Khalil. A separation principle for the control of a class of nonlinear systems. *IEEE Transactions on Automatic Control*, 46:742–746, 2001.

[88] E. D. Sontag. Comments on integral variants of ISS. *Systems & Control Letters*, 34:93–100, 1998.

[89] H. R. Shaker and J. Stoustrup. Control configuration selection for multivariable descriptor systems. In *Proceedings of American Control Conference*, pages 6294–9299, Montreal, QC, 2012.

[90] D. R. Lewin and A. Parag. A constrained genetic algorithm for decentralized control system structure selection and optimization. *Automatica*, 39:1801–1807, 2003.

[91] J. B. Jorgensen and S. B. Jorgensen. Towards automatic decentralized control structure selection. *Computers & Chemical Engineering*, 24:841–846, 2000.

[92] W. Al-Gherwi, H. Budman, and A. Elkamel. Selection of control structure for distributed model predictive control in the presence of model errors. *Journal of Process Control*, 20:270–284, 2010.

[93] S. Heo, W. A. Marvin, and P. Daoutidis. Automated synthesis of control configurations for process networks based on structural coupling. *Chemical Engineering Science*, 136:76–87, 2015.

[94] X. Yin and J. Liu. Input–output pairing accounting for both structure and strength in coupling. *AIChE Journal*, 63(4):1226–1235, 2017.

[95] S. Heo and P. Daoutidis. Control-relevant decomposition of process networks via optimization-based hierarchical clustering. *AIChE Journal*, 62(9):3177–3188, 2016.

[96] L. Kang, W. Tang, Y. Liu, and P. Daoutidis. Control configuration synthesis using agglomerative hierarchical clustering: A graph-theoretic approach. *Journal of Process Control*, 46:43–54, 2016.

[97] X. Yin and J. Liu. Distributed output-feedback fault detection and isolation of cascade process networks. *AIChE Journal*, 63(10):4329–4342, 2017.

[98] M. S. Chiu and Y. Arkun. Decentralized control structure selection based on integrity considerations. *Industrial & Engineering Chemistry Research*, 29:369–373, 1990.

[99] L. Narraway and J. Perkins. Selection of process control structure based on economics. *Computers & Chemical Engineering*, 18, Supplement 1:S511–S515, 1994.

[100] J. H. Lee, R. D. Braatz, M. Morari, and A. Packard. Screening tools for robust control structure selection. *Automatica*, 31:229–235, 1995.

[101] M. Hovd and S. Skogestad. Simple frequency-dependent tools for control system analysis, structure selection and design. *Automatica*, 28:989–996, 1992.

[102] W. Tang and P. Daoutidis. Network decomposition for distributed control through community detection in input–output bipartite graphs. *Journal of Process Control*, 64:7–14, 2018.

[103] D. B. Pourkargar, A. Almansoori, and P. Daoutidis. Impact of decomposition on distributed model predictive control: A process network case study. *Industrial & Engineering Chemistry Research*, 56(34):9606–9616, 2017.

[104] W. Tang, D. B. Pourkargar, and P. Daoutidis. Relative time-averaged gain array (RTAGA) for distributed control-oriented network decomposition. *AIChE Journal*, 2018. doi:10.1002/aic.16130.

[105] W. Tang, A. Allman, D. B. Pourkargar, and P. Daoutidis. Optimal decomposition for distributed optimization in nonlinear model predictive control through community detection. *Computers & Chemical Engineering*, 111:43–54, 2018.

[106] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *Proceedings of 46th IEEE Conference on Decision and Control*, pages 5492–5498, New Orleans, LA, 2007.

[107] R. Vadigepalli and III Doyle, F.J. A distributed state estimation and control algorithm for plantwide processes. *IEEE Transactions on Control Systems Technology*, 11:119–127, 2003.

[108] M. V. Subbotin and R. S. Smith. Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies. *Automatica*, 45:2491–2501, 2009.

[109] B. Chen and W. Wang. Robust stabilization of nonlinearly perturbed large-scale systems by decentralized observer-controller compensators. *Automatica*, 26:1035–1041, 1990.

[110] G. Antonelli, F. Arrichiello, F. Caccavale, and A. Marino. A decentralized controller-observer scheme for multi-agent weighted centroid tracking. *IEEE Transactions on Automatic Control*, 58:1310–1316, 2013.

[111] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas in Communications*, 26:622–633, 2008.

[112] R. Schneider, H. Scheu, and W. Marquardt. An iterative partition-based moving horizon estimator for large-scale linear systems. In *Proceedings of the European Control Conference*, pages 2621–2626, Zurich, Switzerland, 2013.

[113] J. Zhang and J. Liu. Observer-enhanced distributed moving horizon state estimation subject to communication delays. *Journal of Process Control*, 24:672–686, 2014.

[114] Y. Shardt, R. Gonzalez, and A. Tulsyan. *Process Identification Experiment Lab Manual*. University of Alberta, Edmonton, AB, 2009.

[115] H. K. Khalil. *Nonlinear systems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[116] M. Ellis and P. D. Christofides. Selection of control configurations for economic model predictive control systems. *AIChE Journal*, 60:3230–3242, 2014.

[117] X. Chen, M. Heidarinejad, J. Liu, D. Muñoz de la Peña, and P. D. Christofides. Model predictive control of nonlinear singularly perturbed systems: Application to a large-scale process network. *Journal of Process Control*, 21:1296–1305, 2011.

[118] P. Daoutidis and C. Kravaris. Structural evaluation of control configurations for multivariable nonlinear processes. *Chemical Engineering Science*, 47:1091–1107, 1992.

[119] O. Ore. *Theory of graphs*, volume 38. American Mathematical Soc., 1962.

[120] X. Qu, P. J. J. Alvarez, and Q. Li. Applications of nanotechnology in water and wastewater treatment. *Water Research*, 47(12):3931–3946, 2013.

[121] H. Han and J. Qiao. Hierarchical neural network modeling approach to predict sludge volume index of wastewater treatment process. *IEEE Transactions on Control Systems Technology*, 21(6):2423–2431, 2013.

[122] A. M. Nagy Kiss, B. Marx, G. Mourot, G. Schutz, and J. Ragot. State estimation of two-time scale multiple models. application to wastewater treatment plant. *Control Engineering Practice*, 19(11):1354–1362, 2011.

[123] J. Busch, D. Elixmann, P. Kühl, C. Gerkens, J. P. Schlöder, H. G. Bock, and W. Marquardt. State estimation for large-scale wastewater treatment plants. *Water Research*, 47(13):4774–4787, 2013.

[124] X. Yin and J. Liu. State estimation of wastewater treatment plants based on model approximation. *Computers & Chemical Engineering*, 111:79–91, 2018.

[125] J. Zeng, J. Liu, T. Zou, and D. Yuan. Distributed extended Kalman filtering for wastewater treatment processes. *Industrial & Engineering Chemistry Research*, 55(28):7720–7729, 2016.

[126] X. Yin and J. Liu. State estimation of wastewater treatment plants based on reduced-order model. In *10th IFAC International Symposium on Advanced Control of Chemical Processes*, Shenyang, China, 566–571, 2018.

[127] E. L. Haseltine and J. B. Rawlings. Critical evaluation of extended Kalman filtering and moving-horizon estimation. *Industrial & engineering chemistry research*, 44(8):2451–2460, 2005.

[128] E. Arnold and S. Dietze. Nonlinear moving horizon state estimation of an activated sludge model. *IFAC Proceedings Volumes*, 34(8):545–550, 2001.

[129] X. Yin, J. Zeng, and J. Liu. Forming distributed state estimation network from decentralized estimators. *IEEE Transactions on Control Systems Technology*, accepted.

[130] J. Alex, L. Benedetti, J. Copp, K. V. Gernaey, U. Jeppsson, I. Nopens, M. N. Pons, L. Rieger, C. Rosen, J. P. Steyer, P. Vanrolleghem, and S. Winkler. Benchmark simulation model no. 1 (BSM1). *Technical Report, Department of Industrial Electrical Engineering and Automation, Lund University*, 2008.

[131] X. Yin, B. Decardi-Nelson, and J. Liu. Subsystem decomposition and distributed moving horizon estimation of wastewater treatment plants. *Chemical Engineering Research and Design*, 134:405–419, 2018.

[132] S. S. Jogwar and P. Daoutidis. Community-based synthesis of distributed control architectures for integrated process networks. *Chemical Engineering Science*, 172:434–443, 2017.

[133] I. Takács, G. G. Patry, and D. Nolasco. A dynamic model of the clarification-thickening process. *Water Research*, 25(10):1263–1271, 1991.

[134] P. Daoutidis and C. Kravaris. Structural evaluation of control configurations for multivariable nonlinear processes. *Chemical engineering science*, 47(5):1091–1107, 1992.

[135] H. K. Khalil. *Nonlinear systems*. Prentice Hall, Upper Saddle River, NJ, 3rd edition, 2002.

[136] P. Daoutidis, W. Tang, and S. S. Jogwar. Decomposing complex plants for distributed control: Perspectives from network theory. *Computers and Chemical Engineering*, in press, doi:10.1016/j.compchemeng.2017.10.015.

[137] M. Ellis and P. D. Christofides. Selection of control configurations for economic model predictive control systems. *AIChE Journal*, 60:3230–3242, 2014.

[138] C. V. Rao, J. B. Rawlings, and J. H. Lee. Constrained linear state estimation - A moving horizon approach. *Automatica*, 37:1619–1628, 2001.

[139] International Water Association. http://www.benchmarkwwtp.org.

[140] K. Reif, S. Günther, E.Yaz and R. Unbehauen. Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic Control*, 44:714-728, 1999.

[141] R. Huang, S. C. Patwardhan and L. T. Biegler. Robust stability of nonlinear model predictive control with extended Kalman filter and target setting. *International Journal of Robust and Nonlinear Control*, 23:1240-1264, 2013.

[142] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.

[143] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

[144] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[145] E. A. Leicht and M. E. J. Newman. Community structure in directed networks. *Physical Review Letters*, 100(11):118703, 2008.

[146] A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Physical Review E*, 70(6):066111, 2004.

[147] J. Liu, D. Muñoz de la Peña, and P. D. Christofides. Distributed model predictive control of nonlinear process systems. *AIChE Journal*, 55:1171–1184, 2009.

[148] M. J. Tippett and J. Bao. Distributed model predictive control based on dissipativity. *AIChE Journal*, 59:787–804, 2013.

[149] J. Zeng and J. Liu. Economic model predictive control of wastewater treatment processes. *Industrial & Engineering Chemistry Research*, 54(21):5710–5721, 2015.

[150] B. Holenda, E. Domokos, A. Rédey, and J. Fazakas. Dissolved oxygen control of the activated sludge wastewater treatment process using model predictive control. *Computers & Chemical Engineering*, 32(6):1270–1278, 2008.