Distributed Adaptive High-Gain Extended Kalman Filtering For Nonlinear Systems

by

Mohammad Rashedi

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

Recently, increasing attention has been given to the theoretical and practical analysis of large-scale networked systems. Large-scale systems are usually composed of several interconnected subsystems connected through material and energy flows. Due to the scale of these systems and the interactions among subsystems, the design of appropriate process monitoring and control systems is challenging. To handle the scale and interactions of large-scale networked systems in process monitoring and control, distributed predictive control and distributed moving horizon estimation approaches have been developed. The distributed framework can improve the performance of the decentralized network and outperform the centralized framework in terms of fault tolerance. Most of the existing distributed control and process monitoring strategies require the availability of the state measurements of all subsystems; however this requirement may not be satisfied in many applications.

In this thesis, we propose a distributed adaptive high-gain extended Kalman filtering approach for nonlinear systems. Specifically, we consider a class of nonlinear systems that are composed of several subsystems interacting with each other via their states. In the proposed approach, an adaptive high-gain extended Kalman filter is designed for each subsystem. The distributed filters communicate with each other to exchange subsystems' estimates. First, assuming continuous communication among the distributed filters, an implementation strategy which specifies how the distributed filters should communicate is designed and the detailed design of the subsystem filter is described. Second, we consider the case where the subsystem filters communicate to exchange information at discrete-time instants. Following this, the

problem of time-varying delays and data losses in communications between subsystems' estimators is considered. For these two latter cases, a state predictor is used in each subsystem filter to provide predictions of the states of other subsystems. Also, to reduce the number of information transmission among the filters and prevent data trafficking, a triggered communication strategy is developed. The stability properties of the proposed distributed estimation schemes with the described communication types are analyzed. Finally, the effectiveness and applicability of the proposed schemes are illustrated via the applications to simulated chemical processes and a Three-Tank experimental system.

Preface

This thesis is an original work by Mohammad Rashedi. The materials presented in this thesis are part of the research project under the supervision of Dr. Biao Huang and Dr. Jinfeng Liu, and are funded by Alberta Innovates Technology Futures (AITF) and Natural Sciences and Engineering Research Council (NSERC) of Canada.

Chapter 2 of this thesis is briefly presented and published in the 9th International Symposium on Advanced Control of Chemical Processes, *ADCHEM 2015*.

Chapter 3 of this thesis is under revision to be published in *International Journal of Robust and Nonlinear Control*.

Chapter 4 of this thesis is published in *AICHE Journal*.

Chapter 5 of this thesis is under revision to be published in *IEEE Transactions on Industrial Informatics*

To my lovely parents and family, for their support and encouragement.

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. Biao Huang and Dr. Jinfeng Liu for their appreciable inspiration, support and patience throughout my research and acting as a mentor to my overall professional development. I would like to thank Dr. Liu for the constant motivation and imparting me the big picture view of my research during the meetings we had on campus. I would like to thank Dr. Huang for providing me excellent opportunities for attending international technical conference and interacting with external research group members. The critical reviews and recommendations that they provided for the improvement of my papers and presentations were very valuable. They were great supervisors and friends and it was a great pleasure to have the opportunity to work under their supervision. Also, I acknowledge for the financial support that Dr. Huang provided me throughout my graduate studies.

During my research studies, I had close interactions with several external professors and researchers. I sincerely thank Dr. Alireza Fatehi, who spent significant amount of time for my research during his visit to our research group. I thank Dr. Amiya Jana, a visiting professor who helped me in modeling of SAGD process. I express my sincere thanks to Dr. Omid Namaki Shoushtari, Dr. Shima Khatibisepehr, Dr. Fadi Ibrahim and Dr. Nima Sammaknejad who helped me whenever I needed.

I would like to thank my group of friends on Campus for making my graduate student life a wonderful experience, both personally and professionally. Also I should thank all CPC group members who knowingly supported me during my studies.

Last but not the least, I would like to gratefully thank my mother who always gave me courage to pursue my education and handle difficulties. In addition to my parents, I owe a lot of thanks to my grandmothers, brother and my wife who gave me their encouragement and moral support.

Contents

List of Tables

List of Figures

List of Symbols

x	Concatenated vector of states for the entire system in the normal form
$\hat{x}$	Concatenated vector of states for the entire original system
z	Concatenated vector of state estimates for the entire system
e	Prediction error
ϵ	Estimation error
f	Nonlinear system function
w	Concatenated vector of state noise of the entire system
u	Concatenated vector of manipulated input variables of the entire system
y	Concatenated vector of measured output variables of the entire system
v	Concatenated vector of measurement noise of the entire system
ϕ_w	Upper bound of the bounded process noise
ϕ_v	Upper bound of the bounded measurement noise
L	Constant for the associated Lipschitz property
t	Continuous time
t_k	Discrete time instant
$L_f h$	Lie derivative of function h in the direction of function f
Υ	Observability set for the entire system
A	Linear system matrix in normal form
b	Nonlinear function in normal form
C	Output matrix in normal form
n	Number of states in centralized system

n_{x_i}	Number of states in subsystem i
n^*	Maximum of the number of states in subsystems
n_y	Number of outputs and subsystems
Φ	Observability vector
J	Innovation term
d	Forgetting horizon used in the calculation of innovation
$\mathbb{X}$	Convex constraint set for the operation of stable states
θ	High-gain parameter
S	Riccati matrix
R	Measurement noise covariance matrix
Q	Process noise covariance matrix
q_m	Diagonal elements of matrix Q
μ	Sigmoid function variable
β	Constant coefficient in Sigmoid function exponential
m	Constant bias in Sigmoid function exponential
λ	Positive constant in gain adaptation function
ΔT	Rising time of the high-gain parameter
θ_m	Half of maximum possible high-gain parameter
Θ_i	Coefficient matrix for change of coordinates of subsystem i, Chapter 2
Θ_c	Coefficient matrix for change of coordinates for entire system, Chapter 2
sup	Supremum value of variable
$\mathbb{I}$	Set of subsystem indices for the entire system
$\mathbb{I}_i$	Set of subsystem indices which have interaction with subsystem i
α_{min}	Lower bound of Riccati matrix, Chapters 2 and 5
α_{max}	Upper bound of Riccati matrix, Chapters 2 and 5
δ_{min}	Lower bound of Riccati matrix, Chapters 3 and 4
δ_{max}	Upper bound of Riccati matrix, Chapters 3 and 4

T	The time when the high-gain parameter rises to maximum value, Chapter 2
T^*	the time when the error becomes small enough to converge exponentially to zero, Chapter 2
T_i	Temperature of CSTR i in the 2-CSTR and 4-CSTR systems
C_{Ai}	Concentration of CSTR i in 2-CSTR and 4-CSTR systems
ρ	Density
c_p	Heat capacity
ΔH	Enthalpy of reaction
F	Flow-rate
Q_{h_i}	Heat input to CSTR i in 2-CSTR and 4-CSTR systems
V	Volume of CSTR in 2-CSTR and 4-CSTR systems
k_{i0}	Pre-exponential constant of reaction i in 2-CSTR and 4-CSTR systems
E	Activation energy of reaction in 2-CSTR and 4-CSTR systems
p	Superscript indicating the predicted variable
Δ	Constant discrete communication interval
V	Lyapunov function in the stability analyses
η	Time instant when the maximum estimation error occurs within communication interval, Chapter 3
ζ	Time instant when the maximum estimation error occurs within communication interval, Chapter 4
Λ	Maximum of Lyapunov function within communication interval, Chapter 3
Ω	Maximum of Lyapunov function within communication interval, Chapter 4
f_e	Upper bound of prediction error
D	Maximum allowable sample time delay
T_m	Maximum allowable consecutive sample time data losses
$d_{i,j}$	Associated time delay for information of filter j to be available for filter i
L_z	Triggering condition threshold

Chapter 1

Introduction

1.1 Motivation

Improvement of operating performance in today's industry is highly demanded, and huge amount of capital is invested to enhance the productivity and profitability of industrial processes. Model predictive control is an advanced strategy in process control which can optimize process operation in real time with the ability to handle constraints and nonlinearities. However, advanced control systems typically require measuring the entire process states which in general is difficult. State estimation is a technique that reconstructs entire process state estimates based on a process model and output measurements to improve the control performance and assist in process monitoring.

In the literature, there are many results on the evaluation of centralized and decentralized schemes in state estimator designs for different classes of systems [1, 2, 3, 4, 5]. However, the performance of these structures may suffer from poor fault tolerance, high complexity and the inability to compensate for interactions in systems composed of interconnected subsystems. In order to resolve the mentioned issues, distributed estimation is proposed which not only maintains the flexibility of decentralized scheme, but also handles the complexity and poor fault tolerance associated with the centralized framework.

Moving horizon estimation and Kalman filter are two main categories of state estimation which have been designed in distributed framework for linear and nonlinear large-scale networked systems. The existing results on distributed Kalman filtering are mostly developed based on consensus algorithms with the application to sensor networks [6, 7, 8, 9], however none of them take the advantage of high-gain extended Kalman fitler to provide the global convergence for nonlinear systems. In order to hold both the global convergence of high-gain EKF and the noise smoothing property associated with regular EKF, an adaptive-gain EKF is proposed in [10]. On the other hand, to handle the nonlinearities in large-scale systems explicitly, a number of researches have been conducted through the application of distributed moving horizon estimation subject to communication issues [11, 12, 13, 14]. Despite the mentioned capability, distributed moving horizon estimation is much more computationally expensive than Kalman filters. Motivated by the above considerations, this thesis proposes a distributed adaptive high-gain EKF for a class of nonlinear systems that are composed of several interacting subsystems.

1.2 Background

Environmental responsibilities, process safety and profitability are highly associated with the improvement of modern process control systems in today's industry. The operation of complex large-scale chemical processes has attracted significant attention in industry to increase the operating efficiency and profits. Large-scale chemical processes include several unit operations (subsystems) which interact through material and energy flows. In the design of automatic control systems for such complex large-scale processes, the increased scale of the process and interaction among subsystems cause many challenges in fulfillment of the fundamental safety, environmental sustainability and profitability requirements [15].

One approach to improve the performance of large scale processes is to deploy optimal plant-wide strategies to ensure safe and efficient operations. Traditionally, the control and

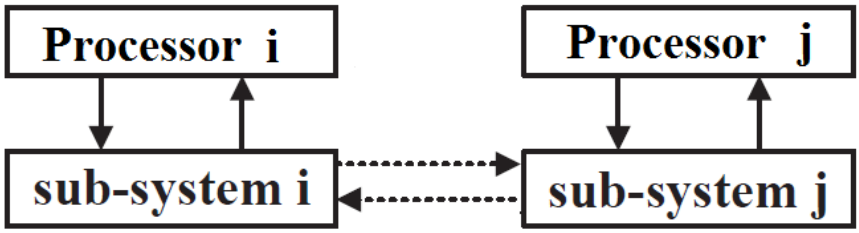

Figure 1.1: Decentralized framework

monitoring of large-scale networked systems were established in either centralized or decentralized framework. For applications involving large-scale systems, centralized algorithms are in general not favorable due to organizational difficulties, high computational complexity and poor fault tolerance. To handle the issues associated with the centralized scheme, the burden of computation can be distributed to several units in the decentralized framework. As shown in Figure 1.1, a separate computational unit is assigned to each subsystem, and the units make the decisions in parallel based on the information received from local subsystems. In addition, this architecture takes the advantage of simpler design for individual units and increased robustness to component failures. Despite the potential advantage of the decentralized framework in improving the fault tolerance, it fails to account for the interactions among the subsystems in large-scale systems. This leads to the plantwide problem towards a suboptimal solution which may lead to lost closed loop stability in predictive control [16, 17]. The distributed framework is a middle ground between the centralized framework and the decentralized framework. While preserving the flexibility of the decentralized framework, the distributed structure achieves improved performance due to the ability of computational units in coordinating their actions via information communication [18, 19, 13]. This framework is achieved by taking a small modification step in the decentralized network which provides information transmission among distributed computational units as shown in Figure 1.2.

In the context of model predictive control, distributed scheme has provided attractive alternative to attain the maximum plant-wide performance by exploiting the interactions of

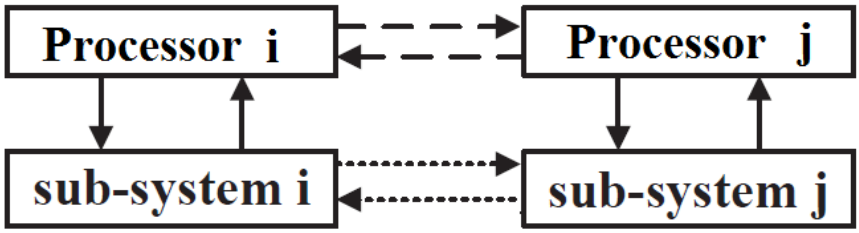

Figure 1.2: Distributed framework

large-scale complex chemical processes [16, 20, 21]. In [22] a distributed MPC is designed for a system composed of decoupled subsystems with the local states and control variables of subsystems being coupled in the objective function. Although the DMPC problem was initially designed for linear systems [23], the extension to the nonlinear constrained systems seems more practical [24, 25]. However most advanced process control strategies require that all the states of the systems are available, which is not always possible. A state estimator may be used to obtain reliable state estimates based on output measurements.

In particular, distributed state estimation of large-scale systems composed of coupled subsystems has attracted significant attention in process control applications [26, 27, 13] as well as in many other engineering control applications [15, 28]. In this approach, a team of collaborated estimators are used to estimate the states of a large-scale system via local knowledge of the system structure and parameters, local measurements and distributed computation. In the literature, distributed state estimation has been studied primarily under two frameworks: distributed Kalman filtering [29] and distributed moving horizon state estimation [27, 13].

Kalman filters are widely used for state estimation in many applications ranging from industrial processes [30, 31] to aerospace navigation systems [32]. Distributed Kalman filtering has been investigated extensively in the past decade [29]. A large portion of the existing results on distributed Kalman filtering (DKF) has been developed for sensor networks [33, 15, 34, 28]. Within the application to sensor networks, the optimality of the distributed filters may be determined based on the type of information transmission to the

fusion centers [35, 36]. Within process control, a distributed Kalman filtering algorithm was formulated for multirate sampled-data systems for plantwide control of processes in [26]. A method for decomposing large-scale processes for distributed Kalman filtering and distributed control was presented in [37]. Although distributed Kalman filtering has been formulated for plant-wide control [26] and decomposition of large-scale processes [37], its application is confined to linear systems. When nonlinear systems are present, extended Kalman filters (EKF) are typically used in the design of distributed state estimation algorithms [38]. As in the centralized EKF, the global stability of the error dynamics of the distributed EKF is difficult to establish. However, taking the advantage of high-gain EKF in the design of each subsystem's filter, the global stability of distributed framework can be guaranteed [39].

In recent years, moving horizon state estimation (MHE) has been adopted in distributed state estimation. One advantage of MHE is its ability to account for state constraints leading to improved estimates [40]. In [11], a distributed MHE algorithm was developed for linear systems which was extended to nonlinear systems in [12]. In [41, 27], distributed MHE schemes based on subsystem models were developed for both linear and nonlinear systems. In [42], an iterative sensitivity-driven partition-based distributed MHE was developed. In [13], an observer-enhanced distributed MHE design with potentially tunable convergence rate was introduced. A method for handling communication delays in distributed MHE was also developed in [43]. While distributed MHE algorithms are able to handle system nonlinearities explicitly, they are typically much more computational demanding than Kalman filters.

In order to handle the above issues, adaptive high-gain extended Kalman filtering (AHG-EKF) method can be adopted in the distributed framework. In the centralized framework, this approach not only removes the computational complexities associated with MHE, but also guarantees the global convergence in state estimation for nonlinear systems by adaptively tuning the Kalman gain [44, 45]. In reference [39], a distributed AHG-EKF design was proposed based on the assumption of continuous communication among state estimators.

Due to the importance of communication between the distributed estimators, the consideration of possible implementation issues like communication data losses and delays is critical. Recently, some approaches are developed to accommodate time-varying delays [43] together with data losses [46] in distributed MHE. However, these results are not directly applicable to distributed AHG-EKF. The handling of communication delays and losses needs to be carefully addressed in the distributed AHG-EKF framework.

All the above scenarios are developed with the distributed filters communicating through the network periodically, however the limited capacity of the network may impede the applicability of the above DAHGEKF. Furthermore, the robustness of the distributed estimation may be reduced by extensive information transmission due to data dropouts in the communication network. In order to address the drawbacks of the periodic communication paradigm, an algorithm is proposed in Chapter 5 to reduce the number of information exchange between local filters based on DAHGEKF designed in [39] via triggered communication. In the literature, this strategy is widely used in control system design with shared communication resources [47, 48, 14]. In the context of state estimation, event-triggered approaches have also been used in wireless sensor networks to attenuate the frequency of information exchange in the network without loss of stability and performance [49, 50]. Moreover, the quasi-decentralized framework is utilized in some other designs to minimize the information exchange for networked control systems by adopting an adaptive forecast-triggered communication algorithm [51].

Motivated by the above considerations, in this work, we propose a distributed adaptive high-gain extended Kalman filtering approach for nonlinear systems. The proposed approach is able to achieve ensured ultimate boundedness and convergence of the estimate and is computationally efficient. Specifically, we consider a type of continuous-time nonlinear systems that are composed of several subsystems interacting with each other via their states. In the proposed approach, an adaptive high-gain extended Kalman filter (AHG-EKF) is designed for each subsystem. The distributed Kalman filters communicate with each other to

exchange estimated subsystem states. In each chapter, the implementation strategy of the proposed distributed state estimation is discussed. It specifies how the distributed filters transmit the information regarding the mentioned scenarios and what information should be exchanged while the detailed design of the subsystem filter is described. Moreover, considering discrete or delayed communication among filters, state predictors are introduced to predict the states of other subsystems in each communication interval. The stability properties of the distributed state estimation are analyzed for all continuous, discrete, delayed and triggered communications. Finally, the effectiveness and applicability of the proposed designs are illustrated via simulations of chemical process examples.

1.3 Terms and definitions

In this section, the definitions of some common terms throughout the thesis are provided for proper clarifications.

Lipschitz property: The Lipschitz property is defined for a continuous function $f : \mathbb{R}^q \to \mathbb{R}^r$ on a set $\Xi \subset \mathbb{R}^q$, if a strictly positive constant exists such that:

$$\|f(x_1) - f(x_2)\| \leq L\|x_1 - x_2\| \tag{1.1}$$

Lie derivative: The Lie derivative of a function $h(x)$ with respect to the vector field $f(x)$ is denoted by the symbol $L_f h(x)$ and is defined as:

$$L_f h(x) = \frac{\partial h(x)}{\partial x} f(x) \tag{1.2}$$

Accordingly, the r-th order Lie derivative is defined as $L_f^r h(x) = L_f L_f^{r-1} h(x)$.

Two-norm: The two-norm of a vector $x \in \mathbb{R}^n$ and a diagonal matrix $S \in \mathbb{R}^{n \times n}$ are

defined as follows:

$$\begin{aligned} \|x\| &= \sqrt{(x_1^2 + x_2^2 + \ldots + x_n^2)} \\ \|S\| &= \sqrt{\lambda_{max}} \end{aligned} \tag{1.3}$$

where λ_{max} is the maximum eigenvalue of $S^H S$, with S^H being the conjugate transpose of S.

1.3.1 Transformation of systems to normal form

We deal with input-affine systems of the form

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(x(t))u(t) \\ y(t) &= h(x(t)) \end{aligned} \tag{1.4}$$

where $x \in \mathbb{R}^n$, and also we consider a subset $\Upsilon \subset \mathbb{R}^n$ under which the system (1.4) is observable. In order to transform the observable system (1.4) into the normal form, the new coordinates $z \in \mathbb{R}^n$ can be defined as:

$$z = \begin{bmatrix} h(x) \\ L_f(h(x)) \\ \vdots \\ L_f^{n-1}(h(x)) \end{bmatrix} = \phi(x) \tag{1.5}$$

Based on (1.5), the system dynamics in the new coordinates will be transformed into

$$\begin{aligned} \dot{z} &= Az + b(z, u) \\ y &= Cz \end{aligned} \tag{1.6}$$

where $C = [1, 0, \ldots, 0]_{1\times n}$ and

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & & \ldots & & 0 \end{bmatrix}_{n\times n}, \; b(z,u) = \begin{bmatrix} L_g h(\phi^{-1}(z))u \\ L_g L_f h(\phi^{-1}(z))u \\ \vdots \\ L_g L_f^{n-2} h(\phi^{-1}(z))u \\ L_f^n h(\phi^{-1}(z)) + L_g L_f^{n-1} h(\phi^{-1}(z))u \end{bmatrix} \tag{1.7}$$

1.3.2 Transformation of subsystems to normal form

Let us consider the system (1.4) is composed of p interacting subsystems with the following state space equations:

$$\begin{cases} \dot{x}_1 = f_1(x) + g_1(x)u \\ y_1 = h_1(x) \end{cases}$$
$$\vdots \tag{1.8}$$
$$\begin{cases} \dot{x}_p = f_p(x) + g_p(x)u \\ y_p = h_p(x) \end{cases}$$

where $x = [x_1^T, \ldots, x_p^T]^T$ and $x_i = [x_{i,1}, \ldots, x_{i,n_{x_i}}]^T \in \mathbb{R}^{n_{x_i}}$ with $i \in \{1, \ldots, p\}$. In order to transform the dynamics of subsystems into the normal form, a similar strategy to subsection 1.3.1 can be taken to find new coordinates and change the dynamics of each subsystem to normal form. Within the network of subsystems, the new coordinates are defined as

$$\Phi(x) = \begin{bmatrix} \Phi_1(x) \\ \Phi_2(x) \\ \vdots \\ \Phi_p(x) \end{bmatrix} \tag{1.9}$$

where $\Phi_i(x) = [h_i(x), L_f h_i(x), \ldots, L_f^{n_{x_i}-1} h_i(x)]^T$. Using the transformation in (1.9), the state space equations of subsystem i will be transformed to

$$\begin{aligned} \dot{z}_i &= A_i z_i + b_i(z,u) \\ y_i &= C_i z_i \end{aligned} \tag{1.10}$$

where $C_i = [1, 0, \ldots, 0]_{1\times n_{x_i}}$ and

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & & \ldots & & 0 \end{bmatrix}_{n_{x_i}\times n_{x_i}}, \; b_i(z,u) = \begin{bmatrix} L_g h_i(\Phi^{-1}(z))u \\ L_g L_f h_i(\Phi^{-1}(z))u \\ \vdots \\ L_g L_f^{n_{x_i}-2} h_i(\Phi^{-1}(z))u \\ L_f^{n_{x_i}} h_i(\Phi^{-1}(z)) + L_g L_f^{n_{x_i}-1} h_i(\Phi^{-1}(z))u \end{bmatrix} \tag{1.11}$$

Notations. The operator $\|\cdot\|$ represents the two-norm of matrix or vector and $|\cdot|_G$ denotes the weighted two-norm which is defined as $|x|_G = \sqrt{x^T G x}$ with G being a positive definite square matrix. The notation $M = \text{diag}\{[m_1, \ldots, m_p]\}$ represents a diagonal matrix M whose diagonal elements are m_i with $i \in \mathbb{I}$ and $\mathbb{I} = \{1, \ldots, p\}$. For a system that is composed of p subsystems, the set $\mathbb{I}_i \subset \mathbb{I}$ denotes the set of corresponding indices of subsystems which have interaction with subsystem i.

1.4 Thesis outline and contributions

The chapters of this thesis are organized as follows.

In Chapter 2, a distributed adaptive high-gain extended Kalman filtering (DAHGKEF) approach is presented for a class of deterministic nonlinear systems that are composed of several interconnected subsystems. For each subsystem a local adaptive high-gain EKF (AHGEKF) is designed and filters communicate to compensate for subsystems' interactions.

In this design, the distributed filters are assumed to transmit information continuously. An implementation strategy describes how the distributed filters should communicate, and the design of local filters is an extension of centralized AHGEKF in [10] for distributed framework. A rigorous stability analysis is carried out to provide sufficient conditions under which the exponential convergence of state estimates to actual states is guaranteed. The efficiency of the proposed method is demonstrated via the application to a 2-CSTR process.

In chapter 3, to take more realistic consideration, the proposed DAHGEKF is designed for a network of coupled subsystems where the distributed filters communicate at discrete time instants. Within the communication interval, each filter incorporates the latest information transmitted from other filters. To improve the performance, a state predictor in each subsystem filter is used to provide predictions of states of other subsystems between two consecutive communication instants. The filters are designed based on both deterministic and stochastic forms of subsystems. In the deterministic form, sufficient conditions are provided to guarantee the convergence of estimation error to zero. In addition, the convergence properties of the proposed distributed estimation schemes under the stochastic form of subsystems is analyzed. The proposed approach is applied to a 4-CSTR process example to illustrate its applicability and effectiveness.

Chapter 4 deals with the DAHGEKF schemes developed in Chapter 3 to handle time-varying communication delays and data losses. An open-loop state predictor is designed for each subsystem which operates based on a two-step predict-update strategy. In other words, each predictor uses the centralized model to predict the overall states and then updates the predictions whenever a new information is received from other filters. The stability of the distributed filtering framework is analyzed under worst case scenario within both deterministic and stochastic forms of subsystems. Sufficient conditions are derived under which the convergence within deterministic structure of subsystems and the boundedness of the estimation error within the stochastic form of subsystems are established. In order to evaluate the applicability and effectiveness of the proposed approach under random delays

and data losses in communications, the 2-CSTR process example is used.

In Chapter 5, a DAHGEKF scheme is designed for nonlinear interconnected subsystems with triggered communication. In this design, although the distributed filters transmit information at discrete time instants, a trigger is designed for each subsystem to reduce the information transmission frequency. The triggering condition is determined based on the difference between the current state estimate and the latest transmitted one. This condition designates if a trigger lets the corresponding filter send information to other filters. The conditions are derived based on which the convergence and ultimate boundedness of the estimation error is ensured. To demonstrate the performance and applicability of the design based on triggering condition, the simulated 4-CSTR process is utilized.

In Chapter 6, the applicability of the proposed distributed estimation framework is verified through the application to a Three-Tank experimental system. In order to design the distributed filters, first a dynamical model is developed and unknown model parameters are identified. Then the identified model is decomposed into three interacting subsystems and an adaptive high-gain EKF is designed for each subsystem. Finally, the performance of distributed adaptive high-gain EKF is compared with that of distributed regular EKF in estimating the states of the experimental system. Moreover, the effects of communication among the local filters in the designed distributed framework are evaluated.

Chapter 7 briefly explains the conclusions of the main results of the thesis and describes future research directions.

Chapter 2

Distributed adaptive high-gain extended Kalman filters with continuous communication

2.1 Introduction

In this chapter, we introduce the proposed distributed state estimation design with continuous communication. The system under consideration is a deterministic system composed of several interacting subsystems. A schematic of the proposed design is shown in Figure 2.1. In the proposed approach, an AHG-EKF is designed for each subsystem. In order to design each AHGEKF, a change of coordinates is required to transform the subsystem into the normal form which is described in Chapter 1. The distributed AHG-EKFs communicate with each other to exchange information. The filter of a subsystem estimates the subsystem state based on the subsystem output measurements and information received from other filters. In this section, we assume that the distributed filters can communicate and exchange information continuously. In Chapter 3, we will extend the results to the case that distributed filters can only communicate at discrete-time instants. With continuous communication, the

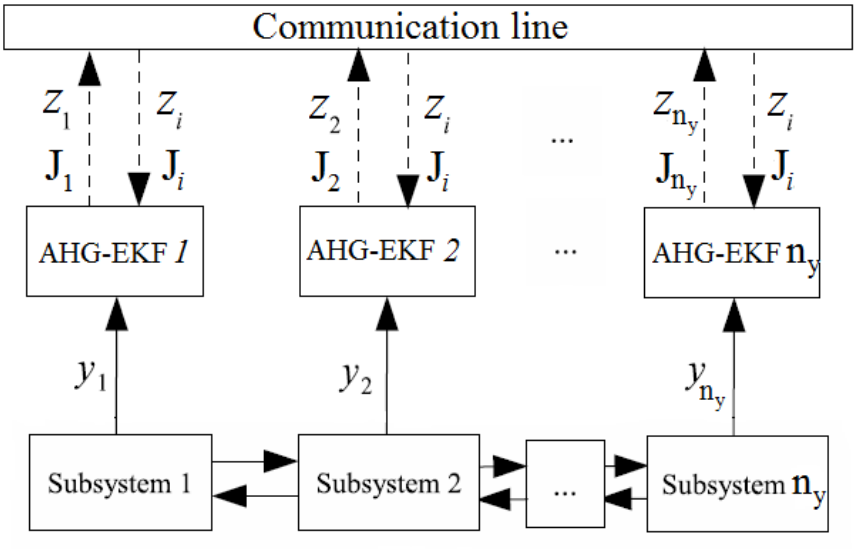

Figure 2.1: Proposed distributed state estimation design with continuous communiation. In this design, an AHG-EKF is designed for each subsystem and the filters communicate with each other.

asymptotic and exponential stability of the distributed estimation approach can be guaranteed under certain conditions. The effectiveness of the proposed approach with continuous communication is illustrated via the application to a chemical process example.

2.2 System description

In this section, we consider a class of deterministic nonlinear systems composed of n_y interacting multi-input single-output subsystems. The dynamics of each subsystem is defined as follows:

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i(t) + b_i(x(t), u(t)) \\ y_i(t) &= C_i x_i(t) \end{aligned} \tag{2.1}$$

where $i \in \mathbb{I}$ with $\mathbb{I} = \{1, \ldots, n_y\}$, $x_i(t) \in \mathbb{R}^{n_{x_i}}$ is the state vector of subsystem i and $y_i(t) \in \mathbb{R}$ is the output of subsystem i. The input $u(t) \in \mathbb{R}^{n_u}$ is assumed to be bounded for all times.

The matrices A_i, C_i and $b_i(x(t), u(t))$ are defined as follows:

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & & \dots & & 0 \end{bmatrix}_{n_{x_i} \times n_{x_i}}, \; b_i(x,u) = \begin{bmatrix} b_{i,1}(x_{i,1}, u) \\ b_{i,2}(x_{i,1}, x_{i,2}, u) \\ \vdots \\ b_{i,n_{x_i}-1}(x_{i,1}, x_{i,2}, \dots, x_{i,n_{x,i}-1}, u) \\ b_{i,n_{x_i}}(x, u) \end{bmatrix} \tag{2.2}$$

$$C_i = [1, 0, \dots, 0]_{1 \times n_{x_i}}$$

and $b_i(x, u)$ is Lipschitz with respect to x and uniform with respect to u. Note that to simplify the analysis without loss of generality, we assume that each subsystem has only one measured output.

The entire system state vector x and measured output vector y are defined as follows: $x = [x_1^T \cdots x_i^T \cdots x_{n_y}^T]^T \in \mathbb{R}^n$, $y = [y_1 \cdots y_i \cdots y_{n_y}]^T \in \mathbb{R}^{n_y}$. The dynamics of the entire system can be described by the following state-space model:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x, u) \\ y(t) &= Cx(t) \end{aligned} \tag{2.3}$$

where $A = \text{diag}\{[A_1, \dots, A_{n_y}]\}$, $b(x, u) = [b_1^T(x, u), \dots, b_{n_y}^T(x, u)]^T$ and

$$C = \begin{bmatrix} C_1 & 0 & \dots & 0 \\ 0 & C_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & C_{n_y} \end{bmatrix}_{n_y \times n} \tag{2.4}$$

It is assumed that the subsystem state x_i satisfies the following constraint:

$$x_i \in \mathbb{X}_i \tag{2.5}$$

for all $i \in \mathbb{I}$. It is also assumed that measurements of the outputs of the subsystems are available continuously.

2.3 Implementation strategy

At the initial time instant (i.e., $t = 0$), each filter needs to be initialized. Specifically, in the initialization, filter i ($i \in \mathbb{I}$) is initialized with initial subsystem state guesses of all the subsystems (i.e., $z_j(0)$ with $j \in \mathbb{I}$), the actual subsystem output measurement (i.e., $y_i(0)$), and the initial value of the adaptive gain (i.e., $\theta(0) = 1$).

After the initialization, at a time instant, each filter needs to carry out the following steps continuously:

1. Filter i ($i \in \mathbb{I}$) receives its local output measurement $y_i(t)$, and the latest state estimates from all the other filters; that is, z_j with $j \in \mathbb{I} \setminus \{i\}$.

2. Filter i ($i \in \mathbb{I}$) calculates its subsystem state estimate $z_i(t)$ and updates its innovation term J_i.

3. Filter i ($i \in \mathbb{I}$) sends z_i and J_i to other subsystems.

4. Filter i ($i \in \mathbb{I}$) updates the adaptive gain θ based on its J_i and J_j ($j \in \mathbb{I}\setminus\{i\}$) from other subsystems.

These steps will be further clarified in the following subsystem filter design section.

2.4 Design of subsystem filters

For each subsystem, an AHG-EKF is designed. The subsystem AHG-EKF design is based on the centralized AHG-EKF presented in [52] and [44] with appropriate modifications to account for interactions between subsystems. In the following, we exclude the notation of time dependency from the state, state estimate and input and denote them as x, z and u.

The proposed design of filter i ($i \in \mathbb{I}$) is formulated as follows:

$$\dot{z}_i = A_i z_i + b_i(z,u) - S_i^{-1} C_i^T R_{\theta_i}^{-1} (C_i z_i - y_i) \tag{2.6}$$

where $R_{\theta_i} = \frac{1}{\theta}\theta^{2(n^*-n_{x_i})} R_i$ with θ being the mutual adaptive gain of filters, $n^* = \max_i\{n_{x_i}\}$ and R_i being a positive scalar. In (2.6), S_i is the solution to the following matrix Riccati equation:

$$\dot{S}_i \;\; = \;\; -\left(A_i + b_i^*(z,u)\right)^T S_i - S_i \left(A_i + b_i^*(z,u)\right) + C_i^T R_{\theta_i}^{-1} C_i - S_i Q_{\theta_i} S_i \tag{2.7}$$

where $b_i^*(z,u)$ denotes the Jacobian of $b_i(z,u)$ with respect to z_i (i.e. $b_i^*(z,u) = \frac{\partial b_i(z,u)}{\partial z_i}$) and $Q_{\theta_i} = \theta \Theta_i^{-1} Q_i \Theta_i^{-1}$ with $\Theta_i = diag\left([\theta^{-(n^*-n_{x_i})}, \theta^{-(n^*-n_{x_i}+1)}, \ldots, \theta^{-(n^*-1)}]\right)$ and Q_i being a n_{x_i} by n_{x_i} symmetric positive definite matrix. It should be noted that in (2.7), the Jacobian of $b_i(z,u)$ with respect to z_j, $j \neq i$, is ignored for simplicity. Indeed, the filters are used in the distributed framework with interacted subsystems but each subsystem's filter is designed based on the Riccati equation as mentioned in (2.7).

In (2.6), the first two terms on the right-hand-side are from subsystem model (2.1) and the last term is a correction term based on the difference between the subsystem measurement y_i and its estimate. The correction term has a time-varying gain which is determined following (2.8) and (2.9).

In the design of (2.6)-(2.7), the parameter θ involved in R_{θ_i} and Q_{θ_i} is an adaptive parameter whose adaptation depends on the innovation information which will be defined. When the innovation indicates that the state estimate is close to the actual system state, θ will be small (i.e., close to 1). When $\theta = 1$, the design of (2.6)-(2.7) essentially reduces to the standard extended Kalman filter (with interactions taken into account). When the innovation indicates that the state estimate is far away from the actual system state, θ will evolve to large values to ensure convergence of the estimate to the actual system state.

Specifically, the equation governing the adaptation of θ is as follows:

$$\begin{aligned} \dot{\theta} &= \mu(J)D(\theta) + (1-\mu(J))\lambda(1-\theta) \\ &\triangleq F(\theta, J) \end{aligned} \tag{2.8}$$

where $\mu(s) = \left[1 + e^{-\beta(s-m)}\right]^{-1}$ is a β and m parameterized sigmoid function, λ is a positive constant, and

$$D(\theta) = \begin{cases} \dfrac{1}{\Delta T}\theta^2, & \text{if} \quad \theta \leq \theta_m \\ \dfrac{1}{\Delta T}(\theta - 2\theta_m)^2, & \text{if} \quad \theta > \theta_m \end{cases} \tag{2.9}$$

with $\Delta T > 0$, $\theta_m > 1$. In (2.8), the variable J is the innovation for the overall system. The innovation is used to drive the adaptation mechanism since it upper bounds the estimation error as proven in Lemma 50 of [10]. Considering a unique forgetting horizon, d, the innovation for the whole system can be defined as

$$J(t) = \int_{t-d}^{t} \|y(s) - y(t-d, z(t-d), s)\|^2 ds = \sum_{i=1}^{p} J_i(t) \tag{2.10}$$

where $J_i(t)$ denotes the innovation for subsystem i, defined as follows:

$$J_i(t) = \int_{t-d}^{t} \|y_i(t-d, x(t-d), \tau) - y_i(t-d, z(t-d), \tau)\|^2 d\tau \tag{2.11}$$

where $y_i(t_0, x_0, \tau)$ is the output of subsystem (2.1) at time τ with $x(t_0) = x_0$. Note that $y_i(t-d, z(t-d), \tau)$ is an open-loop predicted filter output and is not the actual filter output. In the calculation of J_i, $y_i(t-d, x(t-d), \tau)$ for $\tau \in [t-d, t]$ is measured and $y_i(t-d, z(t-d), \tau)$ for $\tau \in [t-d, t]$ needs to be evaluated. In the evaluation of $y_i(t-d, z(t-d), \tau)$, the entire system state estimate $z(t-d)$ is needed and the entire system model should be used due to the interaction between subsystems.

Remark 1. *In order to obtain a proper adaptation function in the local filters, the effective parameters can be tuned systematically. First, we set the gain $\theta = 1$ and tune the Q_i and R_i*

matrices based on standard EKF tuning methods [53, 54, 55]. Next, we tune the maximum value of the high-gain parameter $2\theta_m$ in the pure high-gain EKF mode (i.e., $\theta = 2\theta_m$). In pure high-gain EKF model, the maximum value $2\theta_m$ is tuned in a way such that the estimation error converges to a small neighborhood of the actual system state fast. However, the gain should not be too large such that the noise dominates. Subsequently, we tune the parameters in the adaptation function. In this step, the forgetting horizon d in the innovation should be tuned. With small d, fewer previous measurements are taken into account and the innovation is more sensitive to measurement and process noise, while larger d provides larger delay in the adaptation of the gain. In the tuning of d a balance between noise filtering and adaptation delay must be achieved. The parameters m_i and β_i determine the mean and the slope of the sigmoid function, respectively. In other words, when the innovation is higher than parameter m_i, the parameter θ starts increasing; and the parameter β_i controls the duration of the transition part of the sigmoid. ΔT is a quantity which determines the rising time of θ (the smaller ΔT, the shorter rising time of θ). The value of parameter λ_i determines how fast the gain must decrease when the innovation is close to zero. In the tunning, the parameters should be tuned such that excessive oscillation of θ should be avoided in the presence of noise/disturbances. Also, the adaptation of θ is faster than the dynamics of the system. The tuning process is an iterative process and should be based on offline simulations. The whole parameter tuning procedure is summarized in a flowchart as shown in Figure 2.2. It should be noted that when there are sufficient large online operational changes such that the original model (or the set of model parameters) is not suitable for describing the new operation, it is recommended to re-tune the parameters (together with parameter re-identification).

Remark 2. *It should be noted that the chemical processes may be subject to online operational changes. Within small enough changes, the proposed distributed estimation approach has inherent robustness to handle the model uncertainties. On the other hand, in the presence of sufficient large online operational changes, the parameters are required to be retuned.*

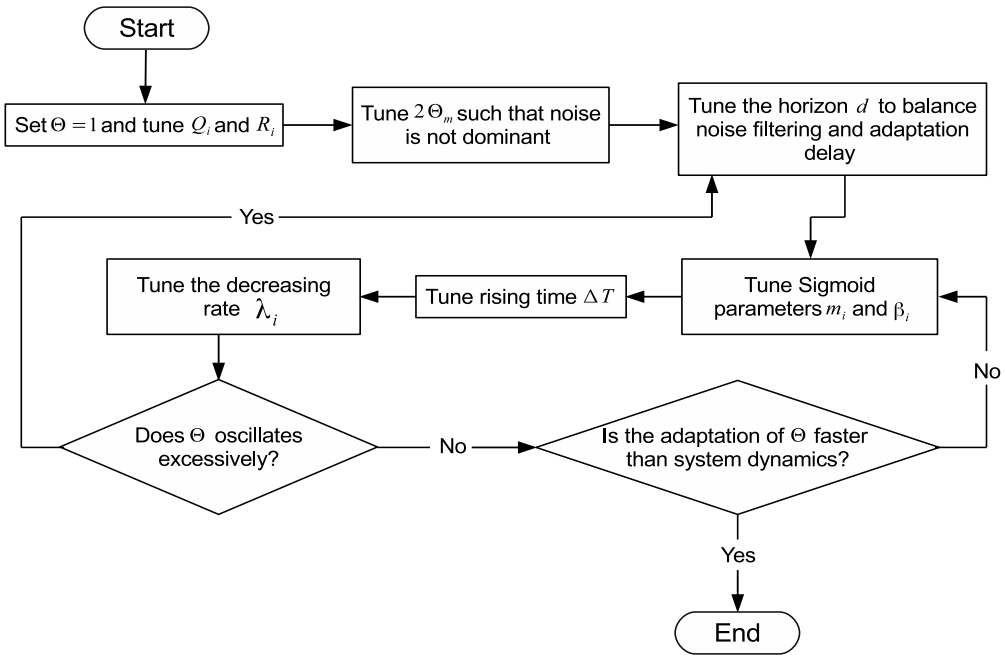

Figure 2.2: Parameter tuning procedure for distributed adaptive-gain filters

2.5 Stability analysis

In this section, we analyze the stability of the proposed distributed AHG-EKF design. Let us define the estimation error for each subsystem as $\epsilon_i = z_i - x_i$, $i \in \mathbb{I}$. The error dynamics is:

$$\dot{\epsilon}_i = \dot{z}_i - \dot{x}_i = \left(A_i - S_i^{-1} C_i^T R_{\theta_i}^{-1} C_i\right) \epsilon_i + b_i(z,u) - b_i(x,u) \tag{2.12}$$

For better explanation of the high-gain effects on the reduction of estimation error, we consider the following change of variables for $i \in \mathbb{I}$:

$$\begin{aligned} &\tilde{x}_i = \Theta_i x_i,\ \tilde{z}_i = \Theta_i z_i,\ \tilde{\epsilon}_i = \Theta_i \epsilon_i,\ \tilde{S}_i = \Theta_i^{-1} S_i \Theta_i^{-1}, \\ &\tilde{b}_i(\cdot,u) = \Theta_i b_i(\Theta_c^{-1}\cdot, u),\ \tilde{b}_i^*(\cdot,u) = \Theta_i b_i^*(\Theta_c^{-1}\cdot, u)\Theta_i^{-1} \end{aligned} \tag{2.13}$$

where $\Theta_c = \mathrm{diag}([\Theta_1, \ldots, \Theta_{n_y}])$. From the definition of Θ_i, it can be verified that:

$$\begin{aligned} &\Theta_i A_i = \theta A_i \Theta_i,\ A_i^T \Theta_i = \theta \Theta_i A_i^T,\ A_i \Theta_i^{-1} = \theta \Theta_i^{-1} A_i,\ \Theta_i^{-1} A_i^T = \theta A_i^T \Theta_i^{-1}, \\ &\dot{\Theta}_i = -\frac{\dot{\theta}}{\theta} N_i \Theta_i,\ \frac{d}{dt}\left(\Theta_i^{-1}\right) = \frac{\dot{\theta}}{\theta} N_i \Theta_i^{-1},\ \Theta_i^{-1} C_i^T R_{\theta_i}^{-1} C_i \Theta_i^{-1} = \theta C_i^T R_i C_i \end{aligned} \tag{2.14}$$

where $N_i = \text{diag}\left([n^* - n_{x_i}, n^* - n_{x_i} + 1, \ldots, n^* - 1]\right)$. Based on (2.13) and (2.14), the error dynamics after the change of variables are:

$$\dot{\tilde{\epsilon}}_i = \dot{\Theta}_i \epsilon_i + \Theta_i \dot{\epsilon}_i = \theta \left[-\tfrac{\dot{\theta}}{\theta^2} N_i \tilde{\epsilon}_i + A_i \tilde{\epsilon}_i - \tilde{S}_i^{-1} C_i^T R_i^{-1} C_i \tilde{\epsilon}_i + \tfrac{1}{\theta} \Big(\tilde{b}_i(\tilde{z}, u) - \tilde{b}_i(\tilde{x}, u) \Big) \right] \tag{2.15}$$

The Riccati equation becomes:

$$\begin{aligned} \dot{\tilde{S}}_i &= \tfrac{d}{dt}\Big(\Theta_i^{-1}\Big) S_i \Theta_i^{-1} + \Theta_i^{-1} S_i \tfrac{d}{dt}\Big(\Theta_i^{-1}\Big) + \Theta_i^{-1} \dot{S}_i \Theta_i^{-1} \\ &= \theta \Bigg[\tfrac{\dot{\theta}}{\theta^2} \Big(N_i \tilde{S}_i + \tilde{S}_i N_i \Big) - \Big(A_i^T \tilde{S}_i + \tilde{S}_i A_i \Big) + C_i^T R_i^{-1} C_i - \tilde{S}_i Q_i \tilde{S}_i - \tfrac{1}{\theta} \tilde{S}_i \tilde{b}_i^*(\tilde{z}, u) - \tfrac{1}{\theta} \tilde{b}_i^{*T}(\tilde{z}, u) \tilde{S}_i \Bigg] \end{aligned} \tag{2.16}$$

Let us pick the Lyapunov function $\tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i$ for subsystem i, $i \in \mathbb{I}$. It can be obtained that:

$$\frac{d\left(\tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i\right)}{dt} = \theta \left[-\tilde{\epsilon}_i^T \tilde{S}_i Q_i \tilde{S}_i \tilde{\epsilon}_i - \tilde{\epsilon}_i^T C_i^T R_i^{-1} C_i \tilde{\epsilon}_i + \frac{2}{\theta} \Big(\tilde{b}_i^T(\tilde{z}, u) \tilde{S}_i \tilde{\epsilon}_i - \tilde{b}_i^T(\tilde{x}, u) \tilde{S}_i \tilde{\epsilon}_i - \tilde{\epsilon}_i^T \tilde{S}_i \tilde{b}_i^*(\tilde{z}, u) \tilde{\epsilon}_i \Big) \right] \tag{2.17}$$

The convergence of the proposed distributed AHG-EKF design will be established based on (2.17). To state the convergence properties, the following lemmas are first presented.

Lemma 1. *Consider a three dimensional matrix $F \in \mathbb{R}^{n_x \times n_y \times n_z}$ and two vectors $y \in \mathbb{R}^{n_y \times 1}$ and $z \in \mathbb{R}^{n_z \times 1}$. If $U_i = [Fz]y = \sum_{j=1}^{n_y} [\sum_{k=1}^{n_z} F_{ijk} z_k] y_j$ and $U = [U_1, \ldots, U_{n_x}]^T$ then*

$$\|U\|_1 \leq n_x K_s \|z\|_1 \|y\|_1 \tag{2.18}$$

where $K_s = sup \|F_{ijk}\|_1$ with $\|\cdot\|_1$ denoting the one-norm.

Proof: Multiplying matrix F by vector z results in the following two-dimensional matrix F^z:

$$F^z = [Fz]_{n_x \times n_y} = \begin{bmatrix} \sum_{k=1}^{n_z} F_{11k} z_k & \sum_{k=1}^{n_z} F_{12k} z_k & \dots & \sum_{k=1}^{n_z} F_{1n_yk} z_k \\ \sum_{k=1}^{n_z} F_{21k} z_k & \sum_{k=1}^{n_z} F_{22k} z_k & \dots & \sum_{k=1}^{n_z} F_{2n_yk} z_k \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^{n_z} F_{n_x1k} z_k & \dots & \dots & \sum_{k=1}^{n_z} F_{n_xn_yk} z_k \end{bmatrix}_{n_x \times n_y} \tag{2.19}$$

The product of matrix F^z and vector y can lead to a vector U

$$U = F^z y = \begin{bmatrix} \sum_{j=1}^{n_y} \left(\sum_{k=1}^{n_z} F_{1jk} z_k \right) y_j \\ \sum_{j=1}^{n_y} \left(\sum_{k=1}^{n_z} F_{2jk} z_k \right) y_j \\ \vdots \\ \sum_{j=1}^{n_y} \left(\sum_{k=1}^{n_z} F_{n_xjk} z_k \right) y_j \end{bmatrix}_{n_x \times 1} \tag{2.20}$$

Based on (2.20), we would have,

$$\|U\|_1 \leq n_x sup \|F_{ijk}\|_1 \sum_{j=1}^{n_y} \sum_{k=1}^{n_z} \|z_k\|_1 \|y_j\|_1 \tag{2.21}$$

and this proves Lemma 1. □

Lemma 2. *Consider* $B = [B_1^T, \dots, B_p^T]^T$, *with* $B_i = \tilde{b}_i(\tilde{z}, u) - \tilde{b}_i(\tilde{x}, u) - \tilde{b}_i^*(\tilde{z}, u)\epsilon_i$, *where* $\tilde{b}_i(\tilde{z}, u) = \Theta_i b(\Theta_c^{-1}\tilde{z}, u)$, $\tilde{b}_i^*(\tilde{z}, u) = \Theta_i b_i^*(\Theta_c^{-1}\tilde{z}, u)\Theta_i^{-1}$, $\Theta_i = diag\left(\left[\theta^{-(n^*-n_{x_i})}, \theta^{-(n^*-n_{x_i}+1)}, \dots, \theta^{-(n^*-1)}\right]\right)$, $\Theta_c = diag([\Theta_1, \dots, \Theta_{n_y}])$ *and* $\epsilon_i = z_i - x_i$. *If* b_i *is smooth and compactly supported, and* $n^* = \max_i\{n_{x_i}\}$ *then,*

$$\|B\| \leq K_1 \|\tilde{\epsilon}\| + K_2 \theta^{n^*-1} \|\tilde{\epsilon}\|^2 \tag{2.22}$$

for some positive constants K_1 *and* K_2.

Proof: If $\tilde{b}_i$, $i \in \mathbb{I}$ are compactly supported and smooth, they can be expanded as follows:

$$\tilde{b}_i(\tilde{z}-t\tilde{\epsilon},u)=\tilde{b}_i(\tilde{z},u)-\sum_{j=1}^{n_y}\left(\int_0^t\frac{\partial\tilde{b}_i}{\partial\tilde{z}_j}(\tilde{z}-\tau\tilde{\epsilon},u)d\tau\right)\tilde{\epsilon}_j \tag{2.23}$$

where $\tilde{\epsilon}_i=\tilde{z}_i-\tilde{x}_i$ and $\tilde{\epsilon}=\tilde{z}-\tilde{x}$, $i\in\mathbb{I}$. Moreover, we can write,

$$\frac{\partial\tilde{b}_i}{\partial\tilde{z}_j}(\tilde{z}-\tau\tilde{\epsilon},u)=\frac{\partial\tilde{b}_i}{\partial\tilde{z}_j}(\tilde{z},u)-\sum_{k=1}^{n_y}\left(\int_0^\tau\frac{\partial^2\tilde{b}_i}{\partial\tilde{z}_j\partial\tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)d\eta\right)\tilde{\epsilon}_k \tag{2.24}$$

Considering $t=1$ and inserting (2.24) into (2.23), it can be obtained that

$$\tilde{b}_i(\tilde{z}-\tilde{\epsilon},u)=\tilde{b}_i(\tilde{z},u)-\sum_{j=1}^{n_y}\frac{\partial\tilde{b}_i}{\partial\tilde{z}_j}(\tilde{z},u)\tilde{\epsilon}_j+\sum_{j=1}^{n_y}\left[\sum_{k=1}^{n_y}\left(\int_0^1\int_0^\tau\frac{\partial^2\tilde{b}_i}{\partial\tilde{z}_j\partial\tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)d\eta d\tau\right)\tilde{\epsilon}_k\right]\tilde{\epsilon}_j \tag{2.25}$$

Since $\tilde{x}=\tilde{z}-\tilde{\epsilon}$, from (2.25) we obtain

$$\tilde{b}_i(\tilde{z},u)-\tilde{b}_i(\tilde{x},u)-\frac{\partial\tilde{b}_i(\tilde{z},u)}{\partial\tilde{z}_i}\tilde{\epsilon}_i=\sum_{j=1,j\neq i}^{n_y}\frac{\partial\tilde{b}_i}{\partial\tilde{z}_j}(\tilde{z},u)\tilde{\epsilon}_j-\sum_{j=1}^{n_y}\left[\sum_{k=1}^{n_y}\left(\int_0^1\int_0^\tau\frac{\partial^2\tilde{b}_i}{\partial\tilde{z}_j\partial\tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)d\eta d\tau\right)\tilde{\epsilon}_k\right]\tilde{\epsilon}_j \tag{2.26}$$

Taking one-norm of both sides of (2.26), we would have

$$\begin{aligned}\left\|\tilde{b}_i(\tilde{z},u)-\tilde{b}_i(\tilde{x},u)-\frac{\partial\tilde{b}_i(\tilde{z},u)}{\partial\tilde{z}_i}\tilde{\epsilon}_i\right\|_1 \leq{}& \left\|\sum_{j=1,j\neq i}^{n_y}\frac{\partial\tilde{b}_i(\tilde{z},u)}{\partial\tilde{z}_j}\tilde{\epsilon}_j\right\|_1\\ &+\left\|\sum_{j=1}^{n_y}\left[\sum_{k=1}^{n_y}\left(\int_0^1\int_0^\tau\frac{\partial^2\tilde{b}_i}{\partial\tilde{z}_j\partial\tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)d\eta d\tau\right)\tilde{\epsilon}_k\right]\tilde{\epsilon}_j\right\|_1\end{aligned} \tag{2.27}$$

For the first term on the right hand side of (2.27) we have

$$\left\|\sum_{j=1,j\neq i}^{n_y}\frac{\partial\tilde{b}_i(\tilde{z},u)}{\partial\tilde{z}_j}\tilde{\epsilon}_j\right\|_1\leq n_{x_i}\tilde{N}_i\sum_{j=1}^{n_y}\|\tilde{\epsilon}_j\|_1=n_{x_i}\tilde{N}_i\|\tilde{\epsilon}\|_1 \tag{2.28}$$

where $\tilde{N}_i = \sup_{\tilde{z}} \|\frac{\partial \tilde{b}_i}{\partial \tilde{z}_j}(\tilde{z},u)\|_1$. Note that $\tilde{\epsilon}_j \in \mathbb{R}^{n_{x_j}\times 1}$, $\tilde{\epsilon}_k \in \mathbb{R}^{n_{x_k}\times 1}$ and $H = \frac{\partial^2 \tilde{b}_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)$ is a three dimensional matrix ($H \in \mathbb{R}^{n_{x_i}\times n_{x_j}\times n_{x_k}}$). Using Lemma 1, it can be obtained that

$$\|[H\tilde{\epsilon}_j]\tilde{\epsilon}_k\|_1 \leq n_{x_i} K_{ijk} \sum_{m=1}^{n_{x_j}} \sum_{q=1}^{n_{x_k}} \|\tilde{\epsilon}_{j_m}\|_1 \|\tilde{\epsilon}_{k_q}\|_1 = n_{x_i} \|\tilde{\epsilon}_j\|_1 \|\tilde{\epsilon}_k\|_1 K_{ijk} \tag{2.29}$$

where $K_{ijk} = sup_{\tilde{z}_j,\tilde{z}_k} \left\| \frac{\partial^2 \tilde{b}_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u) \right\|_1$. Using (2.29), it can be obtained that

$$\left\| \sum_{j=1}^{n_y} \left[\sum_{k=1}^{n_y} \left(\int_0^1 \int_0^\tau \frac{\partial^2 \tilde{b}_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u) d\eta d\tau \right) \tilde{\epsilon}_k \right] \tilde{\epsilon}_j \right\|_1 \leq \frac{n_{x_i}}{2} \tilde{M}_i \sum_{j=1}^{n_y} \sum_{k=1}^{n_y} \|\tilde{\epsilon}_j\|_1 \|\tilde{\epsilon}_k\|_1 \tag{2.30}$$

where $\tilde{M}_i = \max_{j,k}\{K_{ijk}\} = sup_z \|\frac{\partial^2 \tilde{b}_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\tilde{z}-\eta\tilde{\epsilon},u)\|_1$. Noting that $\theta \geq 1$, it can be said that:

$$\left\| \frac{\partial^2 \tilde{b}_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\tilde{z},u) \right\|_1 \leq \theta^{n^*-1} \left\| \frac{\partial^2 b_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\Theta_c^{-1}\tilde{z},u) \right\|_1, \quad \left\| \frac{\partial \tilde{b}_i}{\partial \tilde{z}_j}(\tilde{z},u) \right\|_1 \leq \left\| \frac{\partial b_i}{\partial \tilde{z}_j}(\Theta_c^{-1}\tilde{z},u) \right\|_1 \tag{2.31}$$

From (2.27) and (2.31) we can conclude that

$$\left\| \tilde{b}_i(\tilde{z},u) - \tilde{b}_i(\tilde{x},u) - \frac{\partial \tilde{b}_i(\tilde{z},u)}{\partial \tilde{z}_i}\tilde{\epsilon}_i \right\|_1 \leq n_{x_i} N_i \|\tilde{\epsilon}\|_1 + n_{x_i} \frac{M_i}{2} \|\tilde{\epsilon}\|_1^2 \theta^{n^*-1} \tag{2.32}$$

where $M_i = \sup_z \|\frac{\partial^2 b_i}{\partial \tilde{z}_j \partial \tilde{z}_k}(\Theta_c^{-1}\tilde{z},u)\|_1$ and $N_i = \sup_z \|\frac{\partial b_i}{\partial \tilde{z}_j}(\Theta_c^{-1}\tilde{z},u)\|_1$. In the worst case, (2.32) may be changed to

$$\left\| \tilde{b}_i(\tilde{z},u) - \tilde{b}_i(\tilde{x},u) - \frac{\partial \tilde{b}_i(\tilde{z},u)}{\partial \tilde{z}_i}\tilde{\epsilon}_i \right\|_1 \leq n_{x_i} N_i \|\tilde{\epsilon}\|_1 + n_{x_i} \frac{M_i}{2} \|\tilde{\epsilon}\|_1^2 \theta^{n^*-1} \tag{2.33}$$

Considering (2.33) for the overall system,

$$\left\| \begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) - \frac{\partial \tilde{b}_1(\tilde{z},u)}{\partial \tilde{z}_1}\tilde{\epsilon}_1 \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) - \frac{\partial \tilde{b}_{n_y}(\tilde{z},u)}{\partial \tilde{z}_{n_y}}\tilde{\epsilon}_{n_y} \end{bmatrix} \right\|_1 \leq n_y (n^* N_{max} \|\tilde{\epsilon}\|_1 + n^* \frac{M_{max}}{2} \theta^{n^*-1} \|\tilde{\epsilon}\|_1^2) \tag{2.34}$$

where $N_{max} = \max_i\{N_i\}$ and $M_{max} = \max_i\{M_i\}$. On the other hand, from the norm properties we know that for a vector v of length n, its one-norm and two-norm (denoted as $\|v\|$) satisfy

$$\|v\| \leq \|v\|_1 \leq \sqrt{n}\|v\| \tag{2.35}$$

Consequently, considering $K_1 = n_y\sqrt{n}N_{max}$ and $K_2 = n_y\sqrt{n}n^*\frac{M_{max}}{2}$, from (2.34) it can be obtained that

$$\left\| \begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) - \frac{\partial \tilde{b}_1(\tilde{z},u)}{\partial \tilde{z}_1}\tilde{\epsilon}_1 \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) - \frac{\partial \tilde{b}_{n_y}(\tilde{z},u)}{\partial \tilde{z}_{n_y}}\tilde{\epsilon}_{n_y} \end{bmatrix} \right\| \leq K_1\|\tilde{\epsilon}\| + K_2\theta^{n^*-1}\|\tilde{\epsilon}\|^2 \tag{2.36}$$

and this proves Lemma 2.□

Theorem 1. *Considering Lipschitz property for the nonlinear function, $b_i(x,u)$, and its Jacobian ($\|b_i^*(x,u)\| \leq L_{b_i^*}$) with respect to x and providing the conditions in Lemma 6 of [52], for any time $T^* > 0$ and any $\epsilon^* > 0$, there exist $0 < d < T^*$, $\Delta T > 0$, $\theta_m > 1$, β_i and m_i for $i \in \mathbb{I}$ such that for all times $t \geq T^*$ and any initial condition of subsystems and observers in $\mathbb{X}_i$ (i.e. $x_i(0) \in \mathbb{X}_i$ and $z_i(0) \in \mathbb{X}_i$) for all $i \in \mathbb{I}$, the estimation error of the entire system satisfies:*

$$\|\epsilon(t)\|^2 \leq \frac{\epsilon^*}{(2\theta_m)^{2n^*-2}}e^{-a(t-T^*)} \tag{2.37}$$

where $\epsilon = [\epsilon_1^T, \ldots, \epsilon_{n_y}^T]^T$ and $a > 0$ is a constant.

Proof: It can be verified that $\tilde{\epsilon}_i^T C_i^T R_i^{-1} C_i \tilde{\epsilon}_i \geq 0$. For a positive definite matrix Q_i, there exists a $q_{m_i} > 0$ such that $Q_i \geq q_{m_i} I$. From (2.17), the following inequality can be written:

$$\frac{d\left(\tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i\right)}{dt} \leq -\theta q_{m_i}\tilde{\epsilon}_i^T\tilde{S}_i^2\tilde{\epsilon}_i + 2\tilde{\epsilon}_i^T\tilde{S}_i\left(\tilde{b}_i(\tilde{z},u) - \tilde{b}_i(\tilde{x},u) - \tilde{b}_i^*(\tilde{z},u)\tilde{\epsilon}_i\right) \tag{2.38}$$

From Lemma 6 in [52], it is known that there exist scalars $\alpha_{\min_i} > 0$, $\alpha_{\max_i} > 0$ such that $\alpha_{\min_i} I \leq \tilde{S}_i(t) \leq \alpha_{\max_i} I$. This implies that the inequality (2.38) can be further written as:

$$\frac{d\left(\tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i\right)}{dt} \leq -\theta q_{m_i} \alpha_{\min_i} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i + 2\tilde{\epsilon}_i^T \tilde{S}_i \left(\tilde{b}_i(\tilde{z},u) - \tilde{b}_i(\tilde{x},u) - \tilde{b}_i^*(\tilde{z},u)\tilde{\epsilon}_i\right) \tag{2.39}$$

for all $i \in \mathbb{I}$. Adding (2.39) for all $i \in \mathbb{I}$ together, the following inequality can be obtained:

$$\begin{aligned}
&\frac{d}{dt}\sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i \leq -\theta \min_i\{q_{m_i}\alpha_{\min_i}\} \\
&\times \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix}^T \begin{bmatrix} \tilde{S}_1 & 0 & \dots & 0 \\ 0 & \tilde{S}_2 & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{S}_{n_y} \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix} + 2 \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix}^T \begin{bmatrix} \tilde{S}_1 & 0 & \dots & 0 \\ 0 & \tilde{S}_2 & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{S}_{n_y} \end{bmatrix} \\
&\times \left(\begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) \\ \tilde{b}_2(\tilde{z},u) - \tilde{b}_2(\tilde{x},u) \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) \end{bmatrix} - \begin{bmatrix} \tilde{b}_1^*(\tilde{z},u) & 0 & \dots & 0 \\ 0 & \tilde{b}_2^*(\tilde{z},u) & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{b}_{n_y}^*(\tilde{z},u) \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix} \right)
\end{aligned} \tag{2.40}$$

Since $\tilde{b}_i(\tilde{x},u) = \Theta_i b_i(\Theta_c^{-1}\tilde{x},u)$, $\tilde{b}_i^*(\tilde{x},u) = \Theta_i b_i^*(\Theta_c^{-1}\tilde{x},u)\Theta_i^{-1}$, according to Appendix B (as well as Lemma 51 in [10]), the overall vector $\tilde{b}(\tilde{x},u)$ and the subsystem matrix $\tilde{b}_i^*(\tilde{x},u)$ have the same bound of $b(x,u)$ and $b_i^*(x,u)$, respectively, where $\tilde{b}(\tilde{x},u) = [b_1^T(\tilde{x},u),\dots,b_{n_y}^T(\tilde{x},u)]^T$. So, the Jacobian matrix $\tilde{b}_i^*(\cdot,u)$ is bounded such that $\|\tilde{b}_i^*(\cdot,u)\| \leq L_{b_i^*}$, and $\tilde{b}(\cdot,u)$ is Lipschitz such that $\|\tilde{b}(x^{\{1\}},u) - \tilde{b}(x^{\{2\}},u)\| \leq L_b\|x^{\{1\}} - x^{\{2\}}\|$, and it can be obtained that:

$$
\begin{aligned}
&\left\| \begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) \\ \tilde{b}_2(\tilde{z},u) - \tilde{b}_2(\tilde{x},u) \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) \end{bmatrix} - \begin{bmatrix} \tilde{b}_1^*(\tilde{z},u) & 0 & \dots & 0 \\ 0 & \tilde{b}_2^*(\tilde{z},u) & 0 & \dots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{b}_{n_y}^*(\tilde{z},u) \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix} \right\| \\
&\leq \left\| \begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) \\ \tilde{b}_2(\tilde{z},u) - \tilde{b}_2(\tilde{x},u) \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) \end{bmatrix} \right\| + \left\| \begin{bmatrix} \tilde{b}_1^*(\tilde{z},u) & 0 & \dots & 0 \\ 0 & \tilde{b}_2^*(\tilde{z},u) & 0 & \dots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{b}_{n_y}^*(\tilde{z},u) \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix} \right\| \\
&\leq L_b \|\tilde{\epsilon}\| + n_y L_{b^*_{max}} \|\tilde{\epsilon}\|
\end{aligned} \tag{2.41}
$$

where $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$. From Lemma 6 in [52] we have $\alpha_{min_i} I \leq \tilde{S}_i \leq \alpha_{max_i} I$ and by considering $\tilde{S} = \text{diag}([\tilde{S}_1, \dots, \tilde{S}_{n_y}])$, it can be said that $\alpha_{min} I \leq \tilde{S} \leq \alpha_{max} I$, where $\alpha_{max} = \max_i\{\alpha_{max_i}\}$ and $\alpha_{min} = \min_i\{\alpha_{min_i}\}$. Multiplying both sides of (2.41) by $2\|\tilde{\epsilon}^T\|\|\tilde{S}\|$, it can be obtained that

$$
\begin{aligned}
&2\|\tilde{\epsilon}^T\|\|\tilde{S}\| \left\| \begin{bmatrix} \tilde{b}_1(\tilde{z},u) - \tilde{b}_1(\tilde{x},u) \\ \tilde{b}_2(\tilde{z},u) - \tilde{b}_2(\tilde{x},u) \\ \vdots \\ \tilde{b}_{n_y}(\tilde{z},u) - \tilde{b}_{n_y}(\tilde{x},u) \end{bmatrix} - \begin{bmatrix} \tilde{b}_1^*(\tilde{z},u) & 0 & \dots & 0 \\ 0 & \tilde{b}_2^*(\tilde{z},u) & 0 & \dots \\ \vdots & & \ddots & \\ 0 & \dots & & \tilde{b}_{n_y}^*(\tilde{z},u) \end{bmatrix} \begin{bmatrix} \tilde{\epsilon}_1 \\ \tilde{\epsilon}_2 \\ \vdots \\ \tilde{\epsilon}_{n_y} \end{bmatrix} \right\| \leq \\
&2\left(L_b + n_y L_{b^*_{max}}\right) \|\tilde{\epsilon}^T\|\|\tilde{S}\|\|\tilde{\epsilon}\| \leq 2\left(L_b + n_y L_{b^*_{max}}\right) \alpha_{max} \|\tilde{\epsilon}\|^2 \\
&\leq 2\left(L_b + n_y L_{b^*_{max}}\right) \frac{\alpha_{max}}{\alpha_{min}} \tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}
\end{aligned} \tag{2.42}
$$

Using (2.42) in (2.40) we conclude that:

$$
\frac{d}{dt} \sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i(t) \leq -\theta \min_i\{q_{m_i} \alpha_{\min_i}\} \sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i(t) + \left(2L_b \frac{\alpha_{\max}}{\alpha_{\min}} + 2n_y L_{b^*_{max}} \frac{\alpha_{\max}}{\alpha_{\min}} \right) \sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i(t) \tag{2.43}
$$

The solution to the differential inequality in (2.43) can be obtained as

$$\sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i(t) \leq \sum_{i=1}^{n_y} \tilde{\epsilon}_i^T \tilde{S}_i \tilde{\epsilon}_i(t)\Big|_{t=0} \exp\left(\left[2L_b \frac{\alpha_{\max}}{\alpha_{\min}} + 2pL_{b^*_{max}} \frac{\alpha_{\max}}{\alpha_{\min}} - \theta \min_i\{q_{m_i}\alpha_{\min_i}\}\right]t\right) \tag{2.44}$$

and this inequality can be used globally.

Noting that $1 \leq \theta \leq 2\theta_m$, by applying Lemma 2 to (2.40) we obtain,

$$\frac{d}{dt}\left(\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}\right) \leq -\min_i\{q_{m_i}\alpha_{\min_i}\}\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon} + 2K_1\|\tilde{S}\|\|\tilde{\epsilon}\|^2 + 2K_2(2\theta_m)^{n^*-1}\|\tilde{S}\| \ \|\tilde{\epsilon}\|^3 \tag{2.45}$$

Since $\|\tilde{\epsilon}\|^3 = \left(\|\tilde{\epsilon}\|^2\right)^{3/2} \leq \left(\frac{1}{\alpha_{min}}\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}\right)^{3/2}$, the inequality (2.45) would change to

$$\frac{d}{dt}\left(\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}\right) \leq \left(-q_{m_{min}}\alpha_{min} + \frac{2K_1\alpha_{max}}{\alpha_{min}}\right)\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon} + \frac{2K_2(2\theta_m)^{n^*-1}\alpha_{max}}{(\alpha_{min})^{3/2}}\left(\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}\right)^{3/2} \tag{2.46}$$

By applying Lemma 40 in [10], we obtain that if

$$\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) \leq \frac{\left(q_{m_{min}}\alpha_{min} - \frac{2K_1\alpha_{max}}{\alpha_{min}}\right)^2 (\alpha_{min})^3}{4\left(2K_2(2\theta_m)^{n^*-1}\alpha_{max}\right)^2} \tag{2.47}$$

and

$$q_{m_{min}} > \frac{2K_1\alpha_{max}}{\alpha_{min}^2} \tag{2.48}$$

then for any $t \geq \tau$ we have

$$\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(t) \leq 4\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) e^{-\psi(t-\tau)} \tag{2.49}$$

where

$$\psi = q_{m_{min}}\alpha_{min} - \frac{2K_1\alpha_{max}}{\alpha_{min}} \tag{2.50}$$

Based on (2.47), by assuming a real γ such that

$$\gamma \leq \frac{1}{(2\theta_m)^{2n^*-2}} min\left(\frac{\alpha_{min}\epsilon^*}{4}, \frac{\psi^2\alpha_{min}^3}{4\left(2K_2\alpha_{max}\right)^2}\right) \tag{2.51}$$

then the inequality $\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) \leq \gamma$ helps to have,

$$\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(t) \leq \frac{\alpha_{min} \epsilon^*}{(2\theta_m)^{2n^*-2}} e^{-\psi(t-\tau)} \tag{2.52}$$

for any $t \geq \tau$. Based on the condition (2.47), the inequality (2.52) is satisfied when the estimation error is small enough. So (2.52) is locally exponentially convergent and in the following we will verify it in a global sense. When the estimation error is large, the gain increases and we have $\theta \geq \theta_m$. In this case, for $t \in [T, T^*]$, $T^* > T$, from the global inequality (2.44) we can obtain:

$$\begin{aligned} &\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(T^*) \leq \tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(0) \exp\Big(\Big[B_L - \min_i \{\theta_i q_{m_i} \alpha_{\min_i}\} \Big] T^* \Big) \\ &\leq M_0 \exp\Big[B_L T - T \alpha_{min} q_{m_{min}} + (B_L - \alpha_{min} q_{m_{min}} \theta_m)(T^* - T) \Big] = \\ &M_0 \exp\Big[B_L T^* - T \alpha_{min} q_{m_{min}} - \alpha_{min} q_{m_{min}} \theta_m (T^* - T) \Big] \end{aligned} \tag{2.53}$$

where $B_L = 2(L_b + n_y L_{b^*_{max}}) \frac{\alpha_{\max}}{\alpha_{\min}}$ and $M_0 = \sup_{x,z} \tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(0)$. Now we can tune two parameters θ_m and γ such that

$$M_0 \exp\Big[B_L T^* - T \alpha_{min} q_{m_{min}} - \alpha_{min} q_{m_{min}} \theta_m (T^* - T) \Big] \leq \gamma \tag{2.54}$$

and since $e^{-cte \times \theta_m} < \frac{cte}{\theta_m^{2n^*-2}}$, for θ_m being large enough, (2.51) and (2.54) would be satisfied simultaneously. In other words, even if the estimation error is large, the high-gain parameter can become large enough so that the necessary condition for exponential convergence is satisfied.

Considering a forgetting horizon d, the overall estimation error at time $t - d$ can be obtained as

$$\sum_{i=1}^{n_y} \|z_i(t-d) - x_i(t-d)\|^2 = \|z(t-d) - x(t-d)\|^2 = \|\epsilon(t-d)\|^2 \tag{2.55}$$

According to Lemma 50 of [10], for an MIMO system there exists a constant λ_t such that

$$J(t) \geq \lambda_t \|\epsilon(t-d)\|^2 \tag{2.56}$$

Now, it can be claimed that there exists $\tau \leq T^*$ such that $\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) \leq \gamma$. Indeed, if $\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) > \gamma$ for all $\tau \leq T^*$ then,

$$\gamma < \tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) \leq \alpha_{max} \|\tilde{\epsilon}(\tau)\|^2 \leq \alpha_{max} \|\epsilon(\tau)\|^2 \leq \frac{\alpha_{max}}{\lambda_t} J(\tau + d) \tag{2.57}$$

So, by defining $\gamma_1 = \frac{\lambda_t \gamma}{\alpha_{max}}$, from (2.57) it can be obtained that $J(\tau + d) \geq \gamma_1$ for $\tau \in [0, T^*]$. Consequently, $J(\tau) \geq \gamma_1$ for $\tau \in [d, T^*]$. Based on Lemma 42 of [10], when the innovation of centralized filter is greater than a bound γ_1, the centralized filter would be high gain. The equivalence of this situation in distributed framework is $\theta \geq \theta_m$ for all $i \in \mathbb{I}$ and $t \in [T, T^*]$ which makes a contradiction (i.e. $\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(T^*) \leq \gamma$) based on (2.53). In other words, even if the estimation errors of some of filters are far from the origin, when the sum of overall errors exceeds the bound γ, all the high-gain parameters increase to a value higher than θ_m and $\tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(\tau) > \gamma$.

Finally, for $t \geq \tau$ and using (2.52) we obtain

$$\begin{aligned} \|\epsilon(t)\|^2 &\leq (2\theta_m)^{2n^*-2} \|\tilde{\epsilon}(t)\|^2 \leq \frac{(2\theta_m)^{2n^*-2}}{\alpha_{min}} \tilde{\epsilon}^T \tilde{S} \tilde{\epsilon}(t) \\ &\leq \epsilon^* e^{-\psi(t-\tau)} \end{aligned} \tag{2.58}$$

and this proves Theorem 1.

2.6 Simulations on a chemical process

In this section, the proposed distributed AHG-EKF is applied to a chemical process composed of two connected continuous-stirred tank reactors (CSTRs) as shown in Figure 2.3 [17].

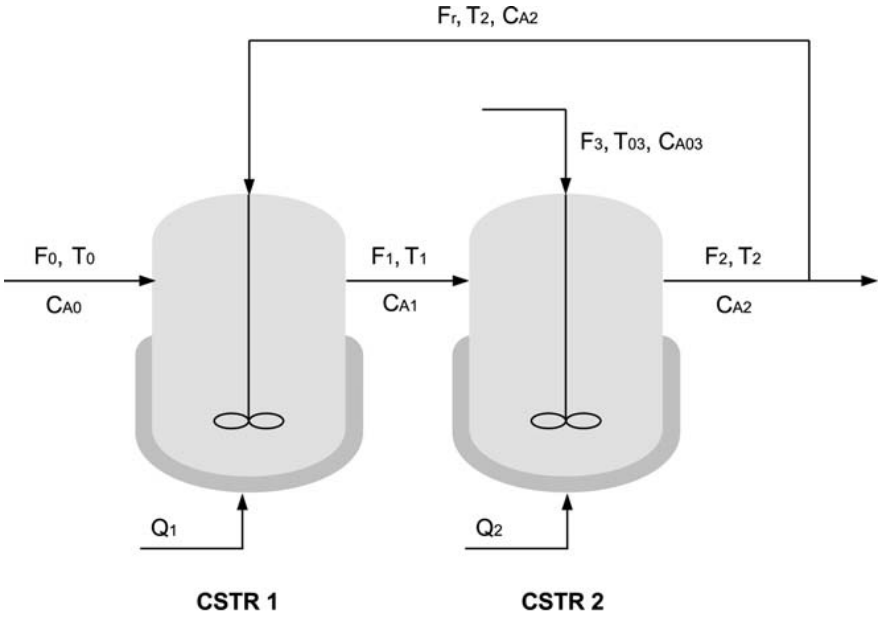

Figure 2.3: Two connected CSTRs with recycle stream.

Pure reactant A at flow rate F_0, molar concentration C_{A0} and temperature T_0 is fed into the first reactor. The effluent of the first reactor enters the second reactor at flow rate F_1, molar concentration C_{A1}, and temperature T_1. Additional pure A at flow rate F_3, molar concentration C_{A03}, and temperature T_{03} is also fed into CSTR 2. Three parallel irreversible exothermal reactions take place in the reactors: $A \to B$, $A \to C$, and $A \to D$, where B is the desired product and C, D are byproducts. A portion of the effluent of the second reactor is passed through a separator and recycled back to the first reactor at flow rate F_r, molar concentration C_{A2} and temperature T_2. Each reactor is equipped with a jacket to provide heat to the reactor. The dynamic equations, obtained from material and energy balance

Table 2.1: Process parameters for the reactors.

$F_0 = 4.998\ m^3/h$	$\Delta H_1 = -5.0 \times 10^4\ KJ/kmol$
$F_1 = 39.996\ m^3/h$	$\Delta H_2 = -5.2 \times 10^4\ KJ/kmol$
$F_3 = 30.0\ m^3/h$	$\Delta H_3 = -5.04 \times 10^4\ KJ/kmol$
$V_1 = 1.0\ m^3/h$	$k_{10} = 3.0 \times 10^6\ h^{-1}$
$V_2 = 3.0\ m^3/h$	$k_{20} = 3.0 \times 10^5\ h^{-1}$
$R = 8.314\ KJ/kmol \cdot K$	$k_{30} = 3.0 \times 10^5\ h^{-1}$
$T_0 = 300\ K$	$E_1 = 5.0 \times 10^4\ KJ/kmol$
$T_{03} = 300\ K$	$E_2 = 7.53 \times 10^4\ KJ/kmol$
$C_{A0} = 4.0\ kmol/m^3$	$E_3 = 7.53 \times 10^4\ KJ/kmol$
$C_{A03} = 2.0\ kmol/m^3$	$\rho = 1000.0\ kg/m^3$
$c_p = 0.231\ KJ/kg$	$F_r = 34.998\ m^3/h$

under standard modeling assumptions are as follows:

$$\begin{aligned}
\frac{dT_1}{dt} &= \frac{F_0}{V_1}(T_0 - T_1) + \frac{F_r}{V_1}(T_2 - T_1) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A1} + \frac{Q_{h_1}}{\rho c_p V_1} \\
\frac{dC_{A1}}{dt} &= \frac{F_0}{V_1}(C_{A0} - C_{A1}) + \frac{F_r}{V_1}(C_{A2} - C_{A1}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A1} \\
\frac{dT_2}{dt} &= \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_3}{V_2}(T_{03} - T_2) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A2} + \frac{Q_{h_2}}{\rho c_p V_2} \\
\frac{dC_{A2}}{dt} &= \frac{F_1}{V_2}(C_{A1} - C_{A2}) + \frac{F_3}{V_2}(C_{A03} - C_{A2}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A2}
\end{aligned} \tag{2.59}$$

where $T_j,\ C_{Aj},\ Q_{h_j},\ V_j,\ j = 1, 2$ denote the temperature in the reactors, the concentration of A, the rate of the heat input/removal to/from the reactors and the reactor volumes respectively, c_p and ρ denote the heat capacity and density of the mixture in the reactors, $\Delta H_i,\ k_i,\ E_i, i = 1, 2, 3$ denote the enthalpies, pre-exponential constants and activation energies of the reactions respectively. The values of these parameters are given in Table 2.1. Constant heat inputs to the two reactors are used: $Q_{h_1} = 1.8 \times 10^4\ kJ/h$ and $Q_{h_2} = 1.2 \times 10^4\ kJ/h$. These inputs ensure the stability of the process.

It is assumed that the two temperatures T_1 and T_2 are the continuously measured outputs. The objective is to estimate the entire system state based on these measurements in a

distributed manner.

The entire process is decomposed into two subsystems with respect to the two reactors. For each reactor, an AHG-EKF is designed. In the design of the AHG-EKFs, $Q_1 = Q_2 = \text{diag}([150, 5])$ and $R_1 = R_2 = 1$, respectively. The parameter θ_m is uniquely defined for all the subsystem estimators and it is selected to be $\theta_m = 20$. In the calculation of the innovation terms, the two filters use $d = 0.02h$. The parameters in the adaptation functions are selected as follows: $\Delta T = 0.001h,\ \beta_1 = 5,\ m_1 = 4,\ \lambda_1 = 100,\ \beta_2 = 5,\ m_2 = 4,\ \lambda_2 = 100$. In the simulations, the initial state of the process is $x(0) = [360, 3, 320, 2.5]^T$ and the initial guesses in the two distributed filters are $z_1(0) = [396,\ 2.6]^T$ and $z_2(0) = [352,\ 10.5]^T$.

First, we consider the proposed distributed AHG-EKF with measurement and process noise and with continuous communication between the two filters and study the effect of adaptive-gain. Consequently we compare the results of distributed adaptive-gain EKF with distributed regular EKF (DEKF). In the regular distributed EKF design, the subsystem EKF has the same design parameters (i.e., Q_1, Q_2, R_1, R_2) as in the proposed AHG-EKF, but $\theta_i = 1$ for $i = 1, 2$. Figure 2.4 shows the trajectories of the actual process states and the estimated values as well as the trajectories of the corresponding adaptive gain and the innovation. From Figure 2.4, it can be seen that the trajectories of both the proposed DAHGEKF and DEKF are able to track the actual process states, however DAHGEKF converges much faster than DEKF. It should be noted that the DAHGEKF tracks the two temperatures very fast. Also, it can be seen that when the estimation errors are large (at the initial period), the innovation values increase quickly which renders the two adaptive gains θ_1 and θ_2 to increase quickly. When the estimation errors become small, the innovation values decrease and the gains θ_1 and θ_2 decrease to one.

Next, we study the effects of communication frequency in the proposed distributed state estimation design. In this set of simulations, the two distributed filters do not exchange information continuously. They send out their state estimates to each other every $36s$. Between two communications, a filter assumes that the state of the other filter remains

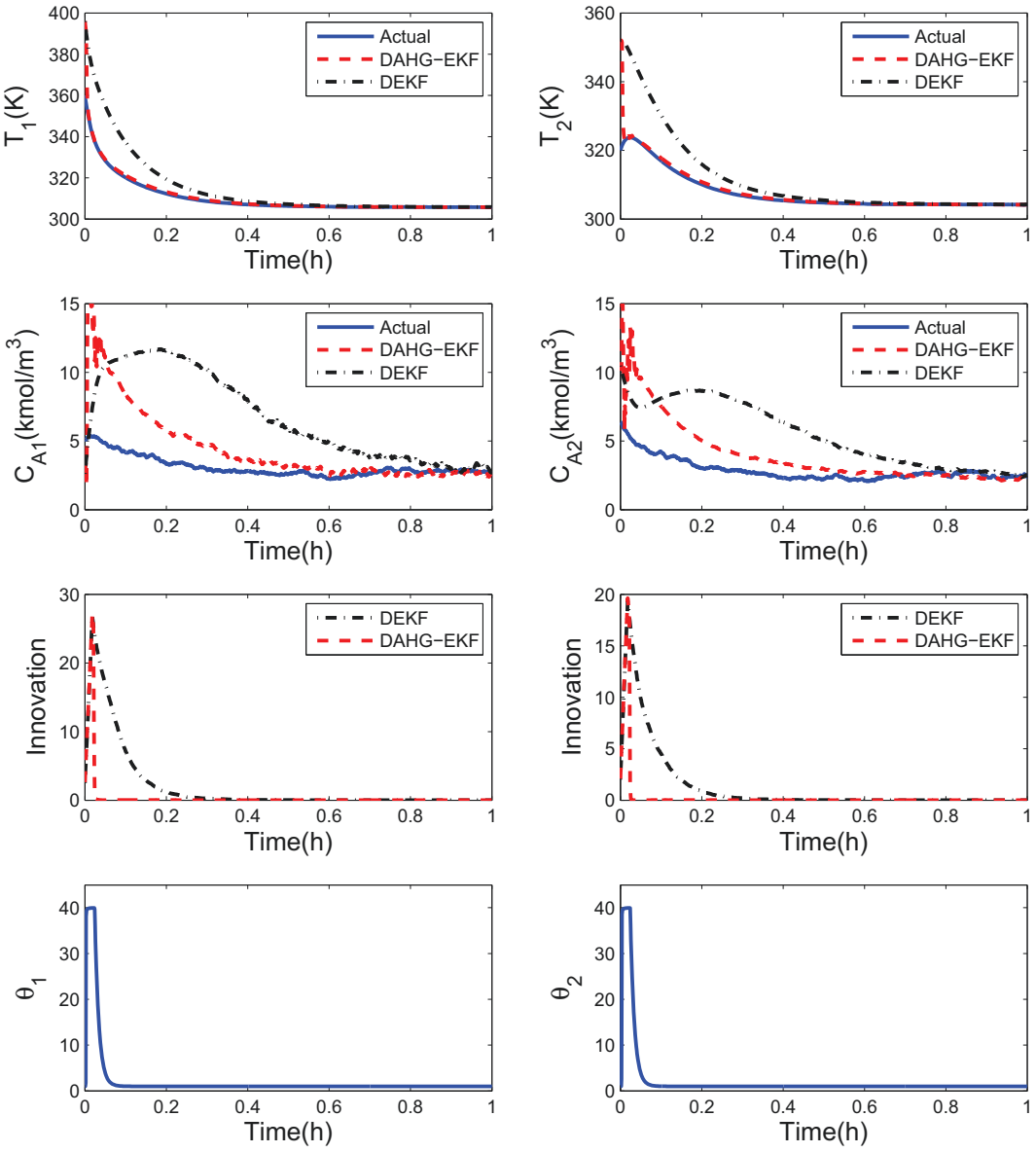

Figure 2.4: Trajectories of T_1, C_{A1}, T_2, C_{A2} and their estimates, and trajectories of θ and the corresponding innovation under the proposed distributed AHG-EKF and a regular distributed EKF with continuous communication.

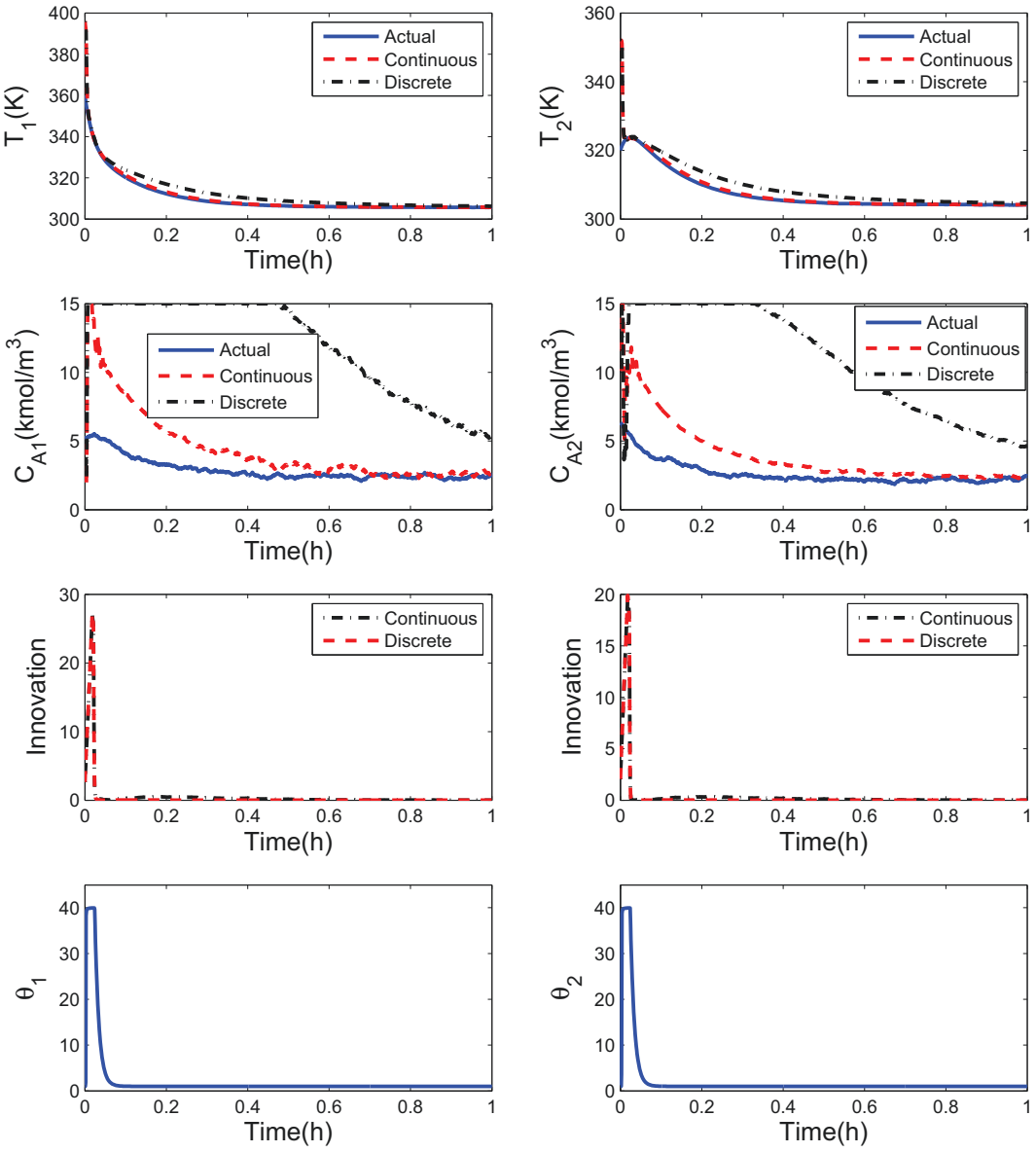

Figure 2.5: Trajectories of T_1, C_{A1}, T_2, C_{A2} and their estimates and the corresponding innovation under the proposed distributed AHG-EKF with continuous and discrete communication.

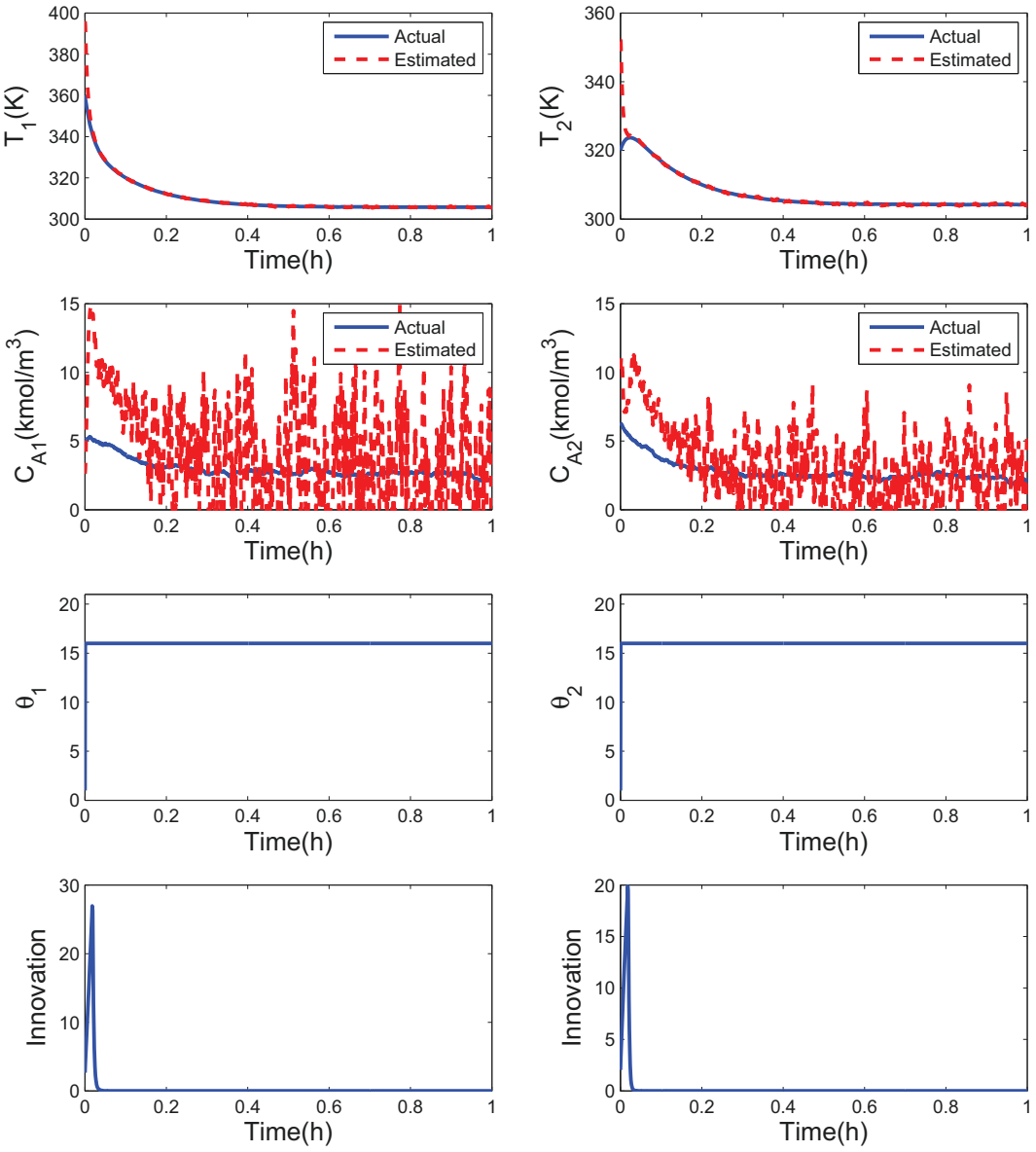

Figure 2.6: Trajectories of T_1, C_{A1}, T_2, C_{A2} and their estimates, and the corresponding innovations under the proposed distributed AHG-EKF with permanent high gain and continuous communication.

constant. Figure 2.5 shows the simulation results. From this figure, it can be seen that the distributed filters are still able to track the actual process states. But the speed of convergence especially for the concentration estimates is slower compared with the case of continuous communication.

Finally, we evaluate the performance of DAHGEKF with permanent high gain. In this set of simulations, we take measurement and process noise into account and show the effect of high gain on process uncertainties. As shown in Figure 2.6, the trajectories of concentrations represent severe sensitivity to the high gain. This sensitivity is illustrated by high amplitude fluctuations in the estimates of concentrations. Also, Figure 2.6 shows that within the high gain, the mean of the estimates matches the mean of actual states very fast, however the estimation error covariance is highly increased by noise. Due to the high sensitivity of uncertainties to the high gain, we used a smaller bound for the high-gain parameter ($\theta_m = 8$) in this set of simulations.

2.7 Conclusions

In this chapter, a distributed adaptive high-gain extended Kalman filtering approach was developed for a class of nonlinear systems composed of interacted subsystems. An adaptive high-gain extended Kalman filter is designed for each subsystem with the consideration of continuous measurement. The distributed filters communicate with each other to exchange subsystem state estimates which are used to compensate for interactions between subsystems. Based on the proposed design, an algorithm is developed to describe how the distributed filters should communicate continuously. Sufficient conditions were derived under which the convergence of the proposed distributed filtering approach is ensured. By the stability analysis, it is shown that the entire system estimate converges to the actual state exponentially. The proposed approach was simulated on a chemical process example to illustrate its applicability and effectiveness.

As a final remark, we would like to note that one particular limitation of the proposed approach is in the calculation of the new coordinates for subsystems. In the calculation of the coordinates, Lie derivatives are needed. Analytical calculation of these Lie derivatives is prohibitive and numerical methods need to be used. Algorithmic differentiation [56] is one of the numerical approaches and is able to handle Lie derivatives up to an order greater than 10.

Chapter 3

Distributed adaptive high-gain extended Kalman filters with discrete communication

3.1 Introduction

In this chapter, we consider distributed state estimation in which the distributed filters communicate with each other only at discrete time instants. As described in Chapter 2, the design of high-gain extended Kalman filter requires the transformation of the system into the normal form. A schematic of the proposed design is shown in Figure 3.1. In this design, for each subsystem, a state predictor is designed to provide state predictions between two communication instants to the subsystem filter. We assume that the distributed filters communicate with each other at discrete time instants $t_{k\geq 0}$ where $t_k = t_0 + k\Delta$ with $t_0 = 0$ being the initial time, k being positive integers and Δ being a positive constant. In this chapter we will consider both deterministic and stochastic form of nonlinear systems and sufficient conditions will be derived under which the stability of the proposed approach is guaranteed.

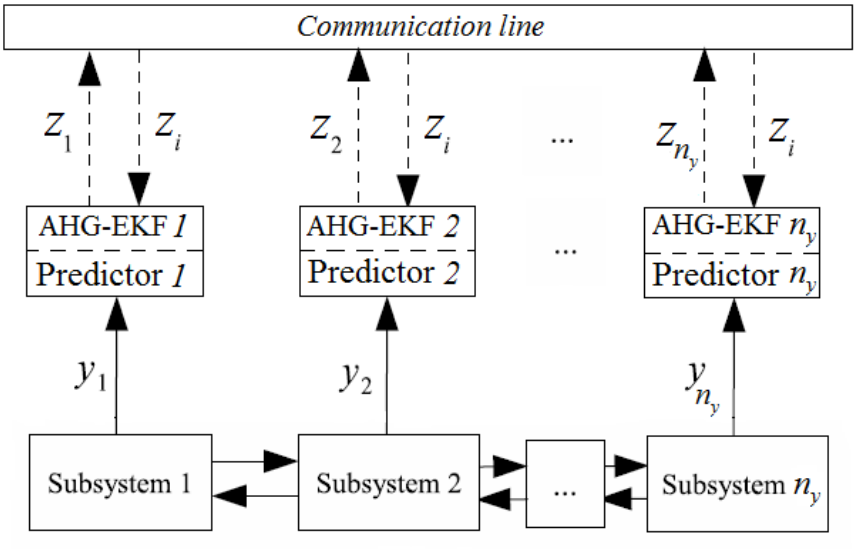

Figure 3.1: Proposed distributed state estimation design with discrete communiation and state predictors.

3.2 Problem formulation

In this section, we consider the interconnected nonlinear subsystem structure in the form of (2.1) but in a more general form including external inputs and uncertainties. Indeed, the entire system is composed of n_y interconnected subsystems with the dynamics of subsystem i described as follows:

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i(t) + b_i(x(t), w(t)), \\ y_i &= C_i x_i(t) + v_i(t) \end{aligned} \tag{3.1}$$

where $i \in \mathbb{I}$ with $\mathbb{I} = \{1, \ldots, n_y\}$, $x_i(t) \in \mathbb{R}^{n_{x_i}}$ is the state vector of subsystem i, $x = [x_1^T \cdots x_i^T \cdots x_{n_y}^T]^T \in \mathbb{R}^n$ is the entire system state vector, $w(t) \in \mathbb{R}^n$ is the vector of entire process noise (or input), b_i represents the nonlinearities of subsystem i describing interaction between subsystem i and other subsystems, and $y_i \in \mathbb{R}$ denotes the measured output of subsystem i with $v_i \in \mathbb{R}$ the associated measurement noise. Further, in (3.1), it is assumed

that $C_i = [1, 0, \ldots, 0]_{1\times n_{x_i}}$ and

$$A_i = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ & 0 & 1 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & 0 & 1 \\ 0 & & \ldots & & 0 \end{bmatrix}_{n_{x_i}\times n_{x_i}}, \; b_i(x,w) = \begin{bmatrix} b_{i,1}(x_{i,1}, w) \\ b_{i,2}(x_{i,1}, x_{i,2}, w) \\ \vdots \\ b_{i,n_{x_i}-1}(x_{i,1}, x_{i,2}, \ldots, x_{i,n_{x,i}-1}, w) \\ b_{i,n_{x_i}}(x, w) \end{bmatrix} \tag{3.2}$$

Note that to simplify the analysis without loss of generality, we assume that each subsystem has only one measured output. The entire system measured output vector y is defined as follows: $y = [y_1 \cdots y_i \cdots y_{n_y}]^T \in \mathbb{R}^{n_y}$. The dynamics of the entire system can be described by the following state-space model:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x(t), w(t)) \\ y(t) &= Cx(t) + v(t) \end{aligned} \tag{3.3}$$

where A is a diagonal composition of A_i and $b(x,w)$ is a composition of $b_i(x,w)$ with $i \in \mathbb{I}$ as follows:

$$A = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & A_{n_y} \end{bmatrix}, \; b(x,w) = \begin{bmatrix} b_1(x,w) \\ b_2(x,w) \\ \vdots \\ b_{n_y}(x,w) \end{bmatrix}, \; C = \begin{bmatrix} C_1 & \mathbf{0}_{1\times n_{x_2}} & \ldots & \mathbf{0}_{1\times n_{x_{n_y}}} \\ \mathbf{0}_{1\times n_{x_1}} & C_2 & \ldots & \mathbf{0}_{1\times n_{x_{n_y}}} \\ \vdots & \ddots & \ddots & \\ \mathbf{0}_{1\times n_{x_1}} & \ldots & \mathbf{0}_{1\times n_{x_{n_y-1}}} & C_{n_y} \end{bmatrix} \tag{3.4}$$

and $v = [v_1 \cdots v_i \cdots v_{n_y}]^T \in \mathbb{R}^{n_y}$. It is assumed that the process and measurement noises are bounded such that $\|w\| \leq \phi_w$ and $\|v\| \leq \phi_v$. It is also assumed that the subsystem state x_i evolves following the constraint:

$$x_i \in \mathbb{X}_i \tag{3.5}$$

for all $i \in \mathbb{I}$ and $\mathbb{X}_i$ should be known. It is further assumed that measurements of the outputs of the subsystems are available continuously. Note that the assumption of boundedness of subsystem states is based on the fact that many systems (e.g., chemical processes) are regulated by control systems and are operated within specified bounded regions. The boundedness of subsystem states will be used in the stability analysis of the proposed method.

Remark 3. *Note that in order to simplify the discussion without loss of generality, we assume that each subsystem has only one measured output. The proposed approach can be extended to consider subsystems containing more than one measured outputs. In the case of multiple outputs, the subsystem dynamics would be in the form of (3.1), with the matrices* A_i *and* $b_i(x, w)$ *being defined as*

$$A_i = \begin{bmatrix} A_{i,1} & 0 & \dots & 0 \\ 0 & A_{i,2} & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & A_{i,n_{y_i}} \end{bmatrix}, \quad A_{i,j} = \begin{bmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & 0 & 0 \end{bmatrix}, \; b_i(x,w) = \begin{bmatrix} b_{i,1}(x,w) \\ \vdots \\ b_{i,n_{y_i}}(x,w) \end{bmatrix} \tag{3.6}$$

where $j \in \{1, \dots, n_{y_i}\}$*. In other words, each subsystem should be able to be decomposed into several smaller single-output subsystems. Please refer to [57, 44] for more discussion on the extension to the multiple outputs case.*

3.3 Predictor design

In this section, the predictors based on the system model in (3.3) will be designed and an algorithm for the implementation of the proposed distributed estimation with discrete

communication will be described. We consider subsystem filter i ($i \in \mathbb{I}$) at time t_k. It communicates with other subsystem filters and will receive state estimates from other subsystems. The next time instant that filter i will communicate with other subsystem filters is t_{k+1}. In order for the filter i to get more accurate state estimates for $t \in [t_k, t_{k+1})$, we use a state predictor to predict the states of other subsystems and use the predictions in the filter. The design of the predictor associated with subsystem i for $t \in [t_k, t_{k+1})$ is as follows:

$$\dot{x}^{p,i}(t) = Ax^{p,i}(t) + b(x^{p,i}(t), 0) \tag{3.7}$$

where $x^{p,i}(t)$ includes the estimates of subsystem i, $z_i(t)$, and the prediction of the states of subsystems which interact with subsystem i, $x_j^{p,i}(t)$, $j \in \mathbb{I}\backslash\{i\}$ within the communication interval Δ, and

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & A_{n_y} \end{bmatrix}, \quad b(x^p(t), 0) = \begin{bmatrix} b_1(x^{p,1}(t), 0) \\ b_2(x^{p,2}(t), 0) \\ \vdots \\ b_{n_y}(x^{p,n_y}(t), 0) \end{bmatrix} \tag{3.8}$$

Note that $z(t_k)$ is the state estimate of the entire system at time t_k and $x^{p,i}(t_k) = z(t_k)$. Note also that for all the subsystems, the predictors have the same design.

3.4 Implementation algorithm

The implementation strategy of DAHG-EKF with discrete communication is as follows:

1. At $t_0 = 0$, filter i, $i \in \mathbb{I}$, is initialized with subsystem measurement at the initial time, $y_i(t_0)$, and the initial guess for the states of overall system $z(t_0)$.

2. At t_k with $k \geq 0$, each filter carries out the following steps:

   (a) Filter i ($i \in \mathbb{I}$) receives the local output measurement $y_i(t_k)$.

(b) Filter i ($i \in \mathbb{I}$) calculates the local estimate $z_i(t_k)$ and updates its innovation term J_i.

(c) Filter i ($i \in \mathbb{I}$) sends z_i to other subsystems.

(d) Filter i ($i \in \mathbb{I}$) updates the adaptive gain θ_i based on its J_i.

(e) Predictor i ($i \in \mathbb{I}$) is updated with $x^{p,i}(t_k) = z(t_k)$.

3. Between t_k and t_{k+1}, filter i and predictor i perform state estimation and state prediction continuously. When $t = t_{k+1}$, go to Step 2 ($k \to k+1$).

Remark 4. *Regarding the design of the predictors, the use of the centralized model is necessary since the subsystems are fully coupled. In the predictor of subsystem i, it uses received subsystem state estimates from other subsystems at a communication instant (i.e., t_k) as the initial condition and uses the centralized system nominal model to predict the states of other subsystems between two consecutive communication instants (i.e., from t_k to t_{k+1}). The subsystem filter i uses the predictions of other subsystems' states to compensate for the interactions between subsystem i and other subsystems. The main purpose of the predictors is to improve the estimation performance of the distributed filters with discrete communication. It will be demonstrated in the simulations that the predictors can significantly improve the estimation performance. If the subsystems are connected in some special patterns (like in series), it is possible to simplify the design of the predictors without reducing the performance. Note that the use of a centralized model in subsystem controller or estimator or predictor designs is not uncommon. This type of strategy has been used in distributed model predictive control (e.g., [58, 59]), quasi-decentralized control [17], etc. The predictors of the subsystems are independent and are computed in a distributed fashion. One important advantage of distributed design over a centralized design is its improved fault tolerance.*

3.5 Deterministic systems

In this section, we investigate the stability properties of the proposed distributed estimation approach in the deterministic system structure, i.e. the specific form of system (2.3) without process and measurement noises terms being described as follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x(t), u(t)) \\ y(t) &= Cx(t) \end{aligned} \tag{3.9}$$

where A, b and C are the same as mentioned in (3.4). In this section we assume the nonlinear function $b(x(t), u(t))$ is Lipschitz with respect to $x(t)$ and uniform with respect to $u(t)$. In the following, based on the system structure in (3.9), the filters are designed and the stability of the distributed filters is analyzed.

3.5.1 Design of subsystem filters

Similar to the previous section, an AHGEKF is designed for each subsystem. Since the communication between the filters is discrete, the proposed design of filter i ($i \in \mathbb{I}$) would be formulated as follows:

$$\dot{z}_i = A_i z_i + b_i(x^{p,i}, u(t)) - S_i^{-1} C_i^T R_{\theta_i}^{-1} (C_i z_i - y_i) \tag{3.10}$$

Also, the Riccati equation is changed to

$$\dot{S}_i = -\left(A_i + b_i^*(x^{p,i}, u)\right)^T S_i - S_i \left(A_i + b_i^*(x^{p,i}, u)\right) + C_i^T R_{\theta_i}^{-1} C_i - S_i Q_{\theta_i} S_i \tag{3.11}$$

where $b_i^*(x^{p,i}, u)$ denotes the Jacobian of $b_i(x^{p,i}, u)$ with respect to z_i (i.e. $b_i^*(x^{p,i}, u) = \frac{\partial b_i(x^{p,i}, u)}{\partial z_i}$), $R_{\theta_i} = \dfrac{R_i}{\theta_i}$ and $Q_{\theta_i} = \theta_i Q_i$. It should be noted that in (3.11), the Jacobian of $b_i(x^{p,i}, 0)$ with respect to z_j, $j \neq i$, is not considered. Note that this design significantly simplifies the design of the Riccati equation while not significantly reduces the performance

nor loses the stability. In the proposed design, the gain in each filter is adaptive and could be high which is able to compensate for uncertainty introduced by the simplified Riccati equation.

In the design of (3.10)-(3.11), the parameter θ_i involved in R_{θ_i} and Q_{θ_i} is an adaptive parameter for subsystem i whose adaptation depends on the innovation information which will be defined. When the innovation of subsystem i indicates that the corresponding state estimate is close to the actual subsystem state, θ_i will be small (i.e., close to 1). When $\theta_i = 1$, the design of (3.10)-(3.11) essentially reduces to the standard extended Kalman filter (with interactions taken into account). When the innovation indicates that the local state estimate is far away from the actual subsystem state, θ_i will evolve to large values to ensure convergence of the estimate to the actual subsystem state. Specifically, the equation governing the adaptation of θ_i is as follows:

$$\begin{aligned} \dot{\theta}_i &= \mu(J_i)D(\theta_i) + (1-\mu(J_i))\lambda_i(1-\theta_i) \\ &\triangleq F(\theta_i, J_i) \end{aligned} \tag{3.12}$$

where $\mu_i(s_i) = \left[1 + e^{-\beta_i(s_i - m_i)}\right]^{-1}$ is a β_i and m_i parameterized sigmoid function, λ_i is a positive constant, and

$$D(\theta_i) = \begin{cases} \dfrac{1}{\Delta T_i}\theta_i^2, & \text{if} \quad \theta_i \le \theta_m \\ \dfrac{1}{\Delta T_i}(\theta_i - 2\theta_m)^2, & \text{if} \quad \theta_i > \theta_m \end{cases} \tag{3.13}$$

with $\Delta T_i > 0$, $\theta_m > 1$. In (3.12), the variable J_i is the innovation for subsystem i. Considering a forgetting horizon for subsystem i, d_i, the innovation for each subsystem can be defined as

$$J_i(t) = \int_{t-d_i}^{t} \|y_i(t-d, x(t-d), \tau) - y_i(t-d, x^{p,i}(t-d), \tau)\|^2 d\tau \tag{3.14}$$

where $y_i(t_0, x_0, \tau)$ is the output of subsystem (3.1) at time τ with $x(t_0) = x_0$. Again we note that $y_i(t-d, x^{p,i}(t-d), \tau)$ is an open-loop predicted filter output and is not the actual

filter output. In the calculation of J_i, $y_i(t-d, x(t-d), \tau)$ for $\tau \in [t-d, t]$ is measured and $y_i(t-d, x^{p,i}(t-d), \tau)$ for $\tau \in [t-d, t]$ needs to be evaluated. In the evaluation of $y_i(t-d, x^{p,i}(t-d), \tau)$, the entire system state prediction $x^{p,i}(t-d)$ is needed and the entire system model should be used due to the interaction between subsystems.

3.5.2 Stability analysis

In this section, the stability properties of the proposed distributed state estimation approach for systems with discrete communications are investigated. First the evolution of prediction error with time is evaluated. An upper bound on the deviation of the predicted system state given by state predictor from the actual system state is provided by the following Proposition 1.

Proposition 1. *Consider the following centralized system and state predictor:*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x(t), u(t)) \\ \dot{x}^p(t) &= Ax^p(t) + b(x^p(t), u(t)) \end{aligned} \tag{3.15}$$

where $x^p(t)$ is the prediction of $x(t)$, $A = diag\{[A_1, A_2, \ldots, A_{n_y}]\}$ and $b(x,u) = [b_1^T(x,u), \ldots, b_{n_y}^T(x,u)]^T$. The system prediction error $e = x^p - x$ satisfies

$$\|e(t)\| \leq f_e(t - t_0, \|e(t_0)\|) \tag{3.16}$$

for all $x(t), x^p(t) \in \mathbb{X} \subset \mathbb{R}^n$, where $f_e(t-t_0, \|e(t_0)\|) = \sqrt{n}\|e(t_0)\|\exp[(L_b+1)(t-t_0)]$ with L_b being the Lipschitz constant of $b(x,u)$ with respect to x.

Proof: The time derivative of the error is,

$$\dot{x}^p - \dot{x} = \dot{e} = Ae + b(x^p, u) - b(x, u) \tag{3.17}$$

According to Lipschitz property of b and the continuity of x and x^p, there exists a constant L_b such that,

$$\left\| \begin{bmatrix} b_1(x^{p,1},u) - b_1(x,u) \\ \vdots \\ b_{n_y}(x^{p,n_y},u) - b_{n_y}(x,u) \end{bmatrix} \right\|_1 = \|b(x^p,u) - b(x,u)\|_1 \leq L_b \|x^p - x\|_1 = L_b \|e\|_1 \tag{3.18}$$

Based on the structure of matrix A in Brunovsky canonical form, we have $\|A\|_1 = 1$. Also from (3.17) we obtain

$$\frac{d}{dt}\|e\|_1 \leq \|\dot{e}\|_1 \leq \|A\|_1\|e\|_1 + L_b\|e\|_1 = (L_b+1)\|e\|_1 \tag{3.19}$$

The solution to (3.19) would be

$$\|e(t)\|_1 \leq \|e(t_0)\|_1 \exp[(L_b+1)(t-t_0)] \tag{3.20}$$

and since $\|e\| \leq \|e\|_1 \leq \sqrt{n}\|e\|$,

$$\|e(t)\| \leq \sqrt{n}\|e(t_0)\| \exp[(L_b+1)(t-t_0)] \tag{3.21}$$

and this proves Proposition 1. □

From Proposition 1, it can be inferred that the prediction error given by the open-loop predictor increases with time between two state updates. In (3.21), t_0 is considered as the initial time in each communication interval. Since the information from other observers is received discretely, the predicted system states will be used in the filters between two communication time instants.

Using (3.10), the dynamics of estimation error, $\epsilon_i = z_i - x_i$, would be:

$$\dot{\epsilon}_i = \dot{z}_i - \dot{x}_i = (A_i - S_i^{-1}C_i^T R_{\theta_i}^{-1} C_i)\epsilon_i + b_i(x^{p,i},u) - b_i(x,u) \tag{3.22}$$

Let us consider the Lyapunov function $V_i = \epsilon_i^T S_i \epsilon_i$ for subsystem i, and use (3.22) and (3.11) to obtain its time derivative:

$$\begin{aligned}\frac{d(\epsilon_i^T S_i \epsilon_i)}{dt} &= \dot{\epsilon}_i^T S_i \epsilon_i + \epsilon_i^T \dot{S}_i \epsilon_i + \epsilon_i^T S_i \dot{\epsilon}_i \\ &= -\theta_i \epsilon_i S_i Q_i S_i \epsilon_i - \theta_i \epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i + 2\epsilon_i^T S_i [b_i(x^{p,i}, u) - b_i(x, u) - b_i^*(x^{p,i}, u)\epsilon_i]\end{aligned} \tag{3.23}$$

Theorem 2. *Considering system (3.3) defined by subsystems in the form of (3.1), if the following conditions are held,*

1. *the nonlinear function, $b(x, u)$, has the Lipschitz property with respect to $x \in \mathbb{R}^n$,*
2. *the subsystem nonlinearity Jacobian, $b_i^*(x^{p,i}, u)$ is bounded,*
3. *the conditions in Theorem 2.18 of [60] are satisfied*
4. *all the estimators make the communications at time instants $\{t_{k\geq 0}\}$,*
5. *the following conditions are satisfied:*

$$q_{m_{min}} > 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta^2_{min}} \tag{3.24}$$

and

$$\frac{2\sqrt{n} L_b e^{(L_b+1)\Delta}}{\min_i\{\theta_i\} q_{m_{min}} \delta_{min} - 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}}} < \frac{\delta_{min}}{\delta_{max}} \left[1 - \left(\frac{2\theta_m}{R_{min}\delta_{min}} + \sqrt{n} L_b e^{(L_b+1)\Delta}\right)\Delta\right]^2 \tag{3.25}$$

with Δ being the discrete communication interval, $\delta_{min} = \min_i\{\delta_{min_i}\}$, $\delta_{max} = \max_i\{\delta_{max_i}\}$ and $R_{min} = \min_i\{R_i\}$,

then for any initial condition of subsystems and observers in $\mathbb{X}_i$ (i.e. $x_i(t_k) \in \mathbb{X}_i$ and $z_i(t_k) \in \mathbb{X}_i$) for all $i \in \mathbb{I}$, the error dynamics asymptotically converges to zero.

Proof: It can be verified that $\theta_i \epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i \geq 0$. If the positive definite matrix Q_i is chosen in a way that $Q_i \geq q_{m_i} I$ for a $q_{m_i} > 0$, then from (3.23) the following inequality can

be obtained:

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \epsilon_i^T S_i^2 \epsilon_i + 2\epsilon_i^T S_i \left(b_i(x^{p,i}, u) - b_i(x, u) - b_i^*(x^{p,i}, u)\epsilon_i\right) \tag{3.26}$$

From Theorem 2.18 in [60], there exist scalars $\delta_{min_i} > 0$ and $\delta_{max_i} > 0$ such that $\delta_{min_i} I \leq S_i \leq \delta_{max_i} I$. This means that (3.26) can be further written as:

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \delta_{min_i} \epsilon_i^T S_i \epsilon_i + 2\epsilon_i^T S_i \left(b_i(x^{p,i}, u) - b_i(x, u) - b_i^*(x^{p,i}, u)\epsilon_i\right) \tag{3.27}$$

If we add (3.27) for all $i \in \mathbb{I}$, the time derivative of the overall system's Lyapunov function satisfies

$$\begin{aligned}
&\frac{d}{dt}\sum_{i=1}^{n_y} \epsilon_i^T S_i \epsilon_i \leq -\theta_i \min_i\{q_{m_i}\delta_{\min_i}\}\epsilon^T S \epsilon + 2\epsilon^T S \\
&\times \left(\begin{bmatrix} b_1(x^{p,1}, u) - b_1(x, u) \\ b_2(x^{p,2}, u) - b_2(x, u) \\ \vdots \\ b_{n_y}(x^{p,n_y}, u) - b_{n_y}(x, u) \end{bmatrix} - \begin{bmatrix} b_1^*(x^{p,1}, u) & 0 & \dots & 0 \\ 0 & b_2^*(x^{p,2}, u) & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y}, u) \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{n_y} \end{bmatrix} \right)
\end{aligned} \tag{3.28}$$

Let us consider the time interval $t \in [t_k, t_{k+1})$. Since at t_k, the filters communicate and exchange information and function b is Lipschitz, based on Proposition 1 for $t \in [t_k,\ t_{k+1})$, we would have

$$\|b(x^p, u) - b(x, u)\| \leq L_b \|x^p - x\| \leq L_b \sqrt{n} \|e(t_k)\| \exp[(L_b + 1)(t - t_k)] \tag{3.29}$$

Noting that $\epsilon(t_k) = e(t_k)$, using the Lipschitz property of nonlinear function b, for the last

term of (3.28), we have:

$$\begin{aligned}
&\left\|\begin{bmatrix} b_1(x^{p,1},u)-b_1(x,u) \\ b_2(x^{p,2},u)-b_2(x,u) \\ \vdots \\ b_{n_y}(x^{p,n_y},u)-b_{n_y}(x,u)\end{bmatrix} - \begin{bmatrix} b_1^*(x^{p,1},u) & 0 & \dots & 0 \\ 0 & b_2^*(x^{p,2},u) & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y},u)\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}\right\| \\
&\leq \left\|\begin{bmatrix} b_1(x^{p,1},u)-b_1(x,u) \\ b_2(x^{p,2},u)-b_2(x,u) \\ \vdots \\ b_{n_y}(x^{p,n_y},u)-b_{n_y}(x,u)\end{bmatrix}\right\| + \left\|\begin{bmatrix} b_1^*(x^{p,1},u) & 0 & \dots & 0 \\ 0 & b_2^*(x^{p,2},u) & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y},u)\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}\right\| \\
&\leq L_b\sqrt{n}\|\epsilon(t_k)\|\exp[(L_b+1)(t-t_k)] + n_y L_{b^*_{max}}\|\epsilon(t)\|
\end{aligned} \tag{3.30}$$

where $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$. From (3.28) and (3.30), it can be obtained that

$$\begin{aligned}
\frac{d}{dt}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) &\leq -\theta_i \min_i\{q_{m_i}\delta_{\min_i}\}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) + 2n_y L_{b^*_{max}}\frac{\delta_{\max}}{\delta_{\min}}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) \\
&+2L_b\sqrt{n}\delta_{\max}\exp[(L_b+1)(t-t_k)]\|\epsilon^T(t)\|\|\epsilon(t_k)\|
\end{aligned} \tag{3.31}$$

where $\delta_{max} = \max_i\{\delta_{max_i}\}$, and $\delta_{min} = \min_i\{\delta_{min_i}\}$. We assume the estimation error takes the maximum value at time instant $\eta \in [t_k, t_{k+1})$ within the communication interval, i.e. $\|\epsilon(\eta)\| \geq \|\epsilon(t)\|$ for any $t \in [t_k, t_{k+1})$ which results in

$$\|\epsilon(t)\|\|\epsilon(t_k)\| \leq \|\epsilon(\eta)\|^2 \tag{3.32}$$

Based on (3.32), it can be inferred that

$$\sum_{i=1}^{n_y}\epsilon_i^T(t)S_i(t)\epsilon_i(t) < \sum_{i=1}^{n_y}\epsilon_i^T(\eta)S_i(t)\epsilon_i(\eta) < \delta_{max}\|\epsilon(\eta)\|^2 := \Lambda\left(\epsilon(\eta)\right) \tag{3.33}$$

We pick $V(t) = \sum_{i=1}^{n_y} V_i(t)$ as the Lyapunov candidate function for the overall system and use

(3.33) in (3.31) to obtain

$$\frac{d}{dt}V(t) \leq \left(-\min_i\{\theta_i q_{m_i}\alpha_{min_i}\} + 2n_y L_{b^*_{max}}\frac{\delta_{max}}{\delta_{min}} \right) V(t) + (2L_b\sqrt{n}\Lambda(\epsilon(\eta)))\exp[(L_b+1)(t-t_k)] \tag{3.34}$$

In addition, (3.22) for the overall system can be written as

$$\dot{\epsilon}(t) = A\epsilon(t) - H_t\epsilon(t) + b(x^p(t)) - b(x(t)) \tag{3.35}$$

where $H_t = \text{diag}\{[S_1^{-1}C_1^T R_{\theta_1}^{-1}C_1, \ldots, S_{n_y}^{-1}C_{n_y}^T R_{\theta_{n_y}}^{-1}C_{n_y}]\}$. Following the boundedness of the Riccati matrix and the high-gain parameter, we can obtain that

$$\|S_i^{-1}C_i^T R_{\theta_i}^{-1}C_i\| \leq \|S_i^{-1}\|\|C_i^T\|\|R_{\theta_i}^{-1}\|\|C_i\| \leq \frac{2\theta_m}{\delta_{min_i}R_i} \tag{3.36}$$

Consequently, the matrix H_t will have the following upper bound:

$$\|H_t\| \leq \frac{2\theta_m}{\delta_{min}R_{min}} \tag{3.37}$$

Since the distributed filters communicate at a fixed interval, Δ, and noting that $\|\epsilon(\eta)\| \geq \|\epsilon(t)\|$, by taking the norm of both sides of (3.35) and using (3.37) and (3.36) we can obtain that,

$$\|\dot{\epsilon}(t)\| \leq \left(\frac{2\theta_m}{\delta_{min}R_{min}} + 1 + L_b\sqrt{n}e^{(L_b+1)\Delta} \right) \|\epsilon(\eta)\| \tag{3.38}$$

On the other hand, within the time period $t \in [t_k, t_{k+1})$ we have

$$\|\epsilon(t)\| \geq \|\epsilon(\eta)\| - \max\{\|\dot{\epsilon}(t)\|\}\Delta \tag{3.39}$$

Also, due to the boundedness of the overall Riccati matrix S, it can be obtained that

$$V(t) \geq \alpha_{min}\min\{\|\epsilon\|^2\} \tag{3.40}$$

Using (3.38) and (3.39) in (3.40) leads to

$$\frac{V(t)}{\Lambda(\epsilon(\eta))} \geq \frac{\delta_{min}}{\delta_{max}} \left[1 - \left(\frac{2\theta_m}{\delta_{min} R_{min}} + 1 + L_b \sqrt{n} e^{(L_b+1)\Delta}\right) \Delta\right]^2 \tag{3.41}$$

From the conditions (3.24) and (3.25) in Theorem 2, and (3.41) it can be verified that

$$\frac{V(t)}{\Lambda(\epsilon(\eta))} > \frac{2L_b \sqrt{n} e^{(L_b+1)\Delta}}{\min_i\{\theta_i\} q_{m_{min}} \delta_{min} - 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}}} \tag{3.42}$$

and it can be concluded that

$$2\Lambda(\epsilon(\eta))\sqrt{n} L_b e^{(L_b+1)\Delta} - \left(\min_i\{\theta_i\} q_{m_{min}} \delta_{min} - 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}}\right) V(t) < 0 \tag{3.43}$$

By comparing (3.43) and (3.34), it can be inferred that $\frac{dV}{dt} < 0$ which implies asymptotic convergence of the error dynamics. This proves Theorem 2.

Remark 5. *Note that in Theorem 1 in Chapter 2, we prove that the estimation error decays exponentially with continuous communication. In the case of discrete-time communication, it is difficult to establish conditions for exponential decay of the error. Instead, in Theorem 2, we provide conditions that ensure asymptotic convergence of the estimation error. We also note that in Theorem 2, the results also apply to EKF (i.e., $\theta = 1$). However, we would like to emphasize that as pointed out in the proof of Theorem 2, the proposed distributed state estimation scheme with an adaptive gain has much faster convergence rate when error is large. The fast convergence rate ensures that the proposed scheme can tolerate larger process disturbances/noise and may have enlarged operating region.*

3.6 Stochastic systems

In this section, we consider the more general form of subsystems in (3.1) and design distributed adaptive high-gain EKF with the effect of noise and disturbance. The schematic of

the proposed design is the same as shown in Figure 3.1. The distributed AHG-EKFs communicate with each other to exchange information and the predictors are designed based on deterministic form of system (2.3) to provide state predictions between two communication instants to the subsystem filters.

3.6.1 Design of subsystem filters

For each subsystem, an AHG-EKF is designed. The subsystem AHG-EKF design is based on the centralized AHG-EKF presented in [52, 44] with appropriate modifications to account for interactions between subsystems. Since the communication between the filters is discrete, the proposed design of filter i ($i \in \mathbb{I}$) is formulated as follows:

$$\dot{z}_i = A_i z_i + b_i(x^{p,i}, 0) - S_i^{-1} C_i^T R_{\theta_i}^{-1}(C_i z_i - y_i) \tag{3.44}$$

The associated Riccati equation is:

$$\dot{S}_i = -\left(A_i + b_i^*(x^{p,i})\right)^T S_i - S_i\left(A_i + b_i^*(x^{p,i})\right) + C_i^T R_{\theta_i}^{-1} C_i - S_i Q_{\theta_i} S_i \tag{3.45}$$

where $b_i^*(x^{p,i})$ denotes the Jacobian of $b_i(x^{p,i}, 0)$ with respect to z_i (i.e. $b_i^*(x^{p,i}) = \dfrac{\partial b_i(x^{p,i}, 0)}{\partial z_i}$), $R_{\theta_i} = \dfrac{R_i}{\theta_i}$ and $Q_{\theta_i} = \theta_i Q_i$. Note that in the deterministic system (3.9), although the inputs ($u(t)$) are time varying, they are assumed to be known and the equations (3.10) and (3.11) are designed based on the nonlinear term $b_i(x, u)$ with the uniformity with respect to $u(t)$. On the other hand, since the inputs ($w(t)$) are unknown, the filter and Riccati equations in (3.44) and (3.45) are designed without the consideration of $w(t)$, and as a result, they are replaced by zero.

Remark 6. *We note that the proposed distributed state estimation method is not an optimal estimation scheme. The proposed method has the following features: (a) the convergence rate is tunable via the tuning of the upper limit of the gain (i.e., θ_m) and the convergence*

rate is potentially fast due to the use of high gains; and (b) the proposed design is less sensitive to noise compared with designs with high gains. These features that render the proposed design are very appealing for distributed output feedback control. However, the proposed design does pose a limitation on the type of systems that can be handled. This limitation is common for observers/estimators using high gains. We also note that in order to have deterministic results, we consider bounded process and measurement noises. This is different from the typical noise realizations in Bayesian based approaches (e.g., [61]). One important application/motivation of the proposed design is distributed output feedback control of nonlinear systems.

3.6.2 Stability analysis

In this section, the stability properties of the proposed distributed state estimation approach with discrete communications are investigated. First the evolution of prediction error with time is evaluated. An upper bound on the deviation of the predicted system state given by the state predictor from the actual system state is provided by the following Proposition 2.

Proposition 2. *Consider the following centralized system and state predictor:*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x(t), w(t)) \\ \dot{x}^p(t) &= Ax^p(t) + b(x^p(t), 0) \end{aligned} \tag{3.46}$$

where $A = diag\{[A_1, A_2, \ldots, A_{n_y}]\}$, $b(x, w) = [b_1^T(x, w), \ldots, b_{n_y}^T(x, w)]^T$ *and* $\|w\|_1 \leq \phi_w$. *The system prediction error* $e = x^p - x$ *satisfies*

$$\|e(t)\| \leq f_e(t - t_0, \|e(t_0)\|) \tag{3.47}$$

for all $x(t), x^p(t) \in \mathbb{X}$, *where* $f_e(t - t_0, \|e(t_0)\|) = (\sqrt{n}\|e(t_0)\| + 2\frac{L_b^w \phi_w}{L_b^x + 1}) \exp[(L_b^x + 1)(t - t_0)] - \frac{L_b^w \phi_w}{L_b^x + 1}$ *with* t_0 *being the initial time instant and* L_b^x *and* L_b^w *being the Lipschitz constant of* $b(x, w)$ *with respect to* x *and* w, *respectively.*

Proof: The time derivative of the error is,

$$\dot{e} = \dot{x}^p - \dot{x} = Ae + b(x^p, 0) - b(x, w) \tag{3.48}$$

According to locally Lipschitz property of b and the continuity of x and x^p, there exists a constants L_b^x, L_b^w such that,

$$\|b(x^p, 0) - b(x, w)\|_1 \leq L_b^x \|x^p - x\|_1 + L_b^w \|w\|_1 \leq L_b^x \|e\|_1 + L_b^w \phi_w \tag{3.49}$$

Based on the structure of matrix A, we have $\|A\|_1 = 1$. From the above equations, we obtain

$$\frac{d}{dt}\|e\|_1 \leq \|\dot{e}\|_1 \leq \|A\|_1\|e\|_1 + L_b^x\|e\|_1 + L_b^w\|w\|_1 \leq (L_b^x + 1)\|e\|_1 + L_b^w\phi_w \tag{3.50}$$

Solving the above inequality with the initial condition $\|e(t_0)\|$, it is obtained that

$$\|e(t)\|_1 \leq (\|e(t_0)\|_1 + 2\frac{L_b^w\phi_w}{L_b^x + 1}) \exp[(L_b^x + 1)(t - t_0)] - \frac{L_b^w\phi_w}{L_b^x + 1}. \tag{3.51}$$

Since $\|e\| \leq \|e\|_1 \leq \sqrt{n}\|e\|$, from (3.51) we obtain

$$\|e(t)\| \leq (\sqrt{n}\|e(t_0)\| + 2\frac{L_b^w\phi_w}{L_b^x + 1}) \exp[(L_b^x + 1)(t - t_0)] - \frac{L_b^w\phi_w}{L_b^x + 1}. \tag{3.52}$$

Given the definition of f_e, this proves Proposition 2. □

Using (3.44), the dynamics of estimation error of subsystem i, $\epsilon_i = z_i - x_i$, is

$$\dot{\epsilon}_i = \dot{z}_i - \dot{x}_i = (A_i - S_i^{-1}C_i^T R_{\theta_i}^{-1} C_i)\epsilon_i + b_i(x^{p,i}, 0) - b_i(x, w) + S_i^{-1}C_i^T R_{\theta_i}^{-1} v_i \tag{3.53}$$

Let us consider the Lyapunov function $V_i = \epsilon_i^T S_i \epsilon_i$ for subsystem i. Its time derivative can

be obtained based on (3.53) and (3.45) as follows:

$$\begin{aligned}\frac{d(\epsilon_i^T S_i \epsilon_i)}{dt} &= \dot{\epsilon}_i^T S_i \epsilon_i + \epsilon_i^T \dot{S}_i \epsilon_i + \epsilon_i^T S_i \dot{\epsilon}_i = -\theta_i \epsilon_i S_i Q_i S_i \epsilon_i - \theta_i \epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i \\ &\quad + 2\epsilon_i^T S_i [b_i(x^{p,i}, 0) - b_i(x, w) - b_i^*(x^{p,i})\epsilon_i] + 2\theta_i \epsilon_i^T C_i^T R_i v_i\end{aligned} \tag{3.54}$$

The stability of the proposed design will be analyzed based on the above subsystem Lyapunov function and is summarized in below in Theorem 3.

Theorem 3. *Consider system (3.3) with subsystems described by (3.1) with the subsystem AHG-EKFs designed following (3.44)-(3.45). If the following assumptions are satisfied:*

1. *function $b(\cdot, \cdot)$ is locally Lipschitz with respect to its arguments,*
2. *the Jacobian $b_i^*(x^{p,i}(t))$ is bounded in the entire operating region of the distributed filters,*
3. *the conditions in Theorem 2.18 of [60] are satisfied (i.e., the Riccati matrices S_i with $i \in \mathbb{I}$ are bounded),*
4. *the positive definite matrix Q_i is chosen in a way that $Q_i \geq q_{m_i} I$ for certain $q_{m_i} > 0$ for $i \in \mathbb{I}$,*
5. *the following condition is held:*

$$q_{m_{min}} > 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta^2_{min}} \tag{3.55}$$

   where $q_{m_{min}} = \min\{q_{m_i}\}$, $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$ with $L_{b_i^*}$ being the Lipschitz constant associated with b_i^*, and δ_{max} and δ_{min} are the maximum and minimum values of the upper and lower bounds of all the subsystem Riccati matrices S_i,
6. *the parameter d_m is the maximum distance between any two points in $\mathbb{X}$ and $R_{min} = \min\{R_i\}$,*

then for any initial condition of subsystems and observers in $\mathbb{X}_i$ (i.e. $x_i(t_0) \in \mathbb{X}_i$ and $z_i(t_0) \in \mathbb{X}_i$) for all $i \in \mathbb{I}$, the norm of the overall system's estimation error $\|\epsilon(t)\|$ is descending and

ultimately smaller than ϵ^*; that is,

$$\lim_{t\to\infty}\|\epsilon(t)\|\leq \epsilon^* \tag{3.56}$$

where $\epsilon^* = \frac{2\delta_{max}L_b^x e^{(L_b^x+1)\Delta}\left(\sqrt{n}d_m+\frac{2L_b^w\phi_w}{L_b^x+1}\right)+\frac{2L_b^w\phi_w}{L_b^x+1}\delta_{max}+\frac{4\theta_m n_y\phi_v}{R_{min}}}{\min\{\theta_i q_{m_i}\}\alpha_{min}^2-2\delta_{max}n_y L_{b_{max}}^*}$.

Proof: It can be verified that $\theta_i\epsilon_i^T C_i^T R_i^{-1}C_i\epsilon_i \geq 0$. If the positive definite matrix Q_i is chosen in a way that $Q_i \geq q_{m_i}I$ for a $q_{m_i}>0$, then from (3.54) the following inequality can be obtained:

$$\frac{d\left(\epsilon_i^T S_i\epsilon_i\right)}{dt} \leq -\theta_i q_{m_i}\epsilon_i^T S_i^2\epsilon_i + 2\theta_i\epsilon_i^T C_i^T R_i v_i + 2\epsilon_i^T S_i\left(b_i(x^{p,i},0)-b_i(x,w)-b_i^*(x^{p,i})\epsilon_i\right) \tag{3.57}$$

From Theorem 2.18 in [60], there exist scalars $\delta_{min_i}>0$ and $\delta_{max_i}>0$ such that $\delta_{min_i}\leq S_i\leq\delta_{max_i}$. This means that (3.57) can be further written as:

$$\frac{d\left(\epsilon_i^T S_i\epsilon_i\right)}{dt} \leq -\theta_i q_{m_i}\delta_{min_i}\epsilon_i^T S_i\epsilon_i + 2\theta_i\epsilon_i^T C_i^T R_i v_i + 2\epsilon_i^T S_i\left(b_i(x^{p,i},0)-b_i(x,w)-b_i^*(x^{p,i})\epsilon_i\right) \tag{3.58}$$

If we add (3.58) for all $i\in\mathbb{I}$, the time derivative of the overall system's Lyapunov function $V=\sum_{i=1}^{n_y}V_i=\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i$ satisfies:

$$\begin{aligned}\frac{dV}{dt} \leq & -\min_i\{\theta_i q_{m_i}\delta_{\min_i}\}\epsilon^T S\epsilon + \sum_{i=1}^{n_y}2\theta_i\epsilon_i^T C_i^T R_i^{-1}v_i \\ & +2\epsilon^T S\begin{bmatrix} b_1(x^{p,1},0)-b_1(x,w)\\ b_2(x^{p,2},0)-b_2(x,w)\\ \vdots \\ b_{n_y}(x^{p,n_y},0)-b_{n_y}(x,w)\end{bmatrix} - 2\epsilon^T S\begin{bmatrix} b_1^*(x^{p,1}) & 0 & \dots & 0\\ 0 & b_2^*(x^{p,2}) & 0 & \vdots\\ \vdots & & \ddots & 0\\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y})\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}\end{aligned} \tag{3.59}$$

Let us consider the time interval $t\in[t_k,t_{k+1})$. Since at t_k, the filters communicate and exchange information and function b is Lipschitz in the entire operating region of the dis-

tributed filters, based on Proposition 2 for $t \in [t_k,\ t_{k+1})$, we would have

$$\|b(x^p,0)-b(x,w)\| \leq (L_b^x+1)\left((\sqrt{n}\|e(t_k)\|+2\frac{L_b^w\phi_w}{L_b^x+1})\exp[(L_b^x+1)(t-t_k)]-\frac{L_b^w\phi_w}{L_b^x+1}\right)+L_b^w\phi_w \tag{3.60}$$

At time instant t_k, since the filters communicate and the states of the predictors are reset to the state estimates, we have $\epsilon(t_k)=e(t_k)$. From (3.59) and (3.60), it can be obtained that:

$$\begin{aligned}\frac{dV(t)}{dt} \leq & -\min\{\theta_i q_{m_i}\delta_{min_i}\}V(t)+2\|\epsilon(t)\|^2\|S(t)\|n_y L_{b^*_{max}}+2\sqrt{n}\|\epsilon(t)\|\|\epsilon(t_k)\|\|S(t)\|L_b^x e^{(L_b^x+1)(t-t_k)} \\ & +\frac{4L_b^xL_b^w\phi_w}{L_b^x+1}e^{(L_b^x+1)(t-t_k)}\|S(t)\|\|\epsilon(t)\|+2(L_b^w\phi_w-\frac{L_b^wL_b^x\phi_w}{L_b^x+1})\|\epsilon(t)\|\|S(t)\|+\frac{4\theta_m n_y\phi_v}{R_{min}}\|\epsilon(t)\|\end{aligned} \tag{3.61}$$

where $\delta_{max}=\max_i\{\delta_{max_i}\}$, and $\delta_{min}=\min_i\{\delta_{min_i}\}$, $L_{b^*_{max}}=\max_i\{L_{b_i^*}\}$, $R_{min}=\min\{R_i\}$. Based on the definition of V and the boundedness of S such that $\delta_{min}\leq\|S\|\leq\delta_{max}$ (according to Theorem 2.18 of [60]), it is derived that:

$$\delta_{min}\|\epsilon(t)\|^2 \leq V(t) \leq \delta_{max}\|\epsilon(t)\|^2 \tag{3.62}$$

From (3.61) and (3.62) and the boundedness of S, it is obtained that:

$$\begin{aligned}\frac{dV(t)}{dt} \leq & -\min\{\theta_i q_{m_i}\delta_{min_i}\}\delta_{min}\|\epsilon(t)\|^2+2\|\epsilon(t)\|^2\delta_{max}n_y L_{b^*_{max}}+2\sqrt{n}\|\epsilon(t)\|\|\epsilon(t_k)\|\delta_{max}L_b^x e^{(L_b^x+1)(t-t_k)} \\ & +\frac{4L_b^xL_b^w\phi_w}{L_b^x+1}e^{(L_b^x+1)(t-t_k)}\delta_{max}\|\epsilon(t)\|+2(L_b^w\phi_w-\frac{L_b^wL_b^x\phi_w}{L_b^x+1})\|\epsilon(t)\|\delta_{max}+\frac{4\theta_m n_y\phi_v}{R_{min}}\|\epsilon(t)\|\end{aligned} \tag{3.63}$$

Since $\|\epsilon(t_k)\|\leq d_m$ from the definition of d_m, $\|t-t_k\|<\Delta$ for $t\in[t_k,t_{k+1})$, and based on the definition of ϵ^*, it can be verified that if condition (3.55) is satisfied, the time derivative of the Lyapunov function V is negative for all $\|\epsilon(t)\|>\epsilon^*$ and $t\in[t_k,t_{k+1})$; that is,

$$\frac{dV}{dt}<0 \tag{3.64}$$

for all $\|\epsilon(t)\|>\epsilon^*$ and $t\in[t_k,t_{k+1})$. Since the distributed filters communicate with each other

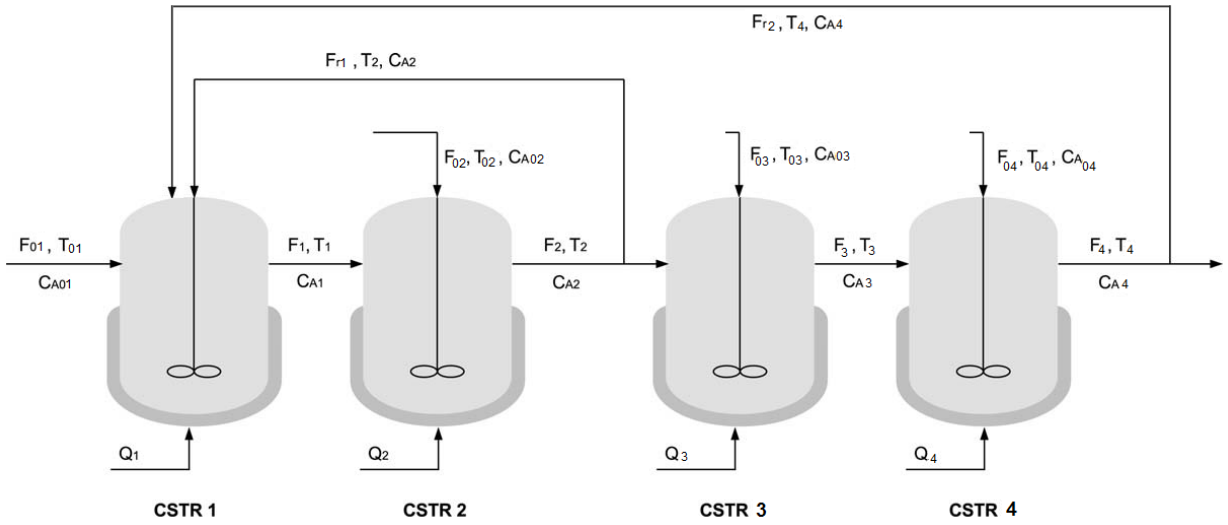

Figure 3.2: Four connected CSTRs with recycle streams.

every sampling time, ϵ is reset to be the same as e every sampling time, using the above result recursively, it is proved that $\frac{dV}{dt} < 0$ for all time as long as $\|\epsilon(t)\| > \epsilon^*$. This implies that within finite time duration, the estimation error $\|\epsilon(t)\|$ will be reduced to be smaller than ϵ^*. Once $\|\epsilon(t)\| < \epsilon^*$, the time derivative of V may be positive and $\|\epsilon(t)\|$ may increase, however, $\|\epsilon(t)\|$ will be always smaller than ϵ^* because $\frac{dV}{dt} < 0$ when $\|\epsilon(t)\| > \epsilon^*$. This proves the theorem. □

3.7 Simulations on a chemical process

In this section, the proposed distributed AHG-EKF is applied to a chemical process composed of four connected continuous-stirred tank reactors (CSTRs) with recycle streams as shown in Figure 3.2. Pure reactant A at flow rate F_{01}, molar concentration C_{A01} and temperature T_{01} is fed into the first reactor. The effluent of the first reactor enters the second reactor at flow rate F_1, molar concentration C_{A1}, and temperature T_1. Additional pure A at flow rate F_{02}, molar concentration C_{A02}, and temperature T_{02} is also fed into CSTR 2. A portion of the effluent of the second reactor is recycled back to the first reactor at flow rate F_{r1} and the rest is directed to the third reactor. Additional pure reactant A is fed into the third and fourth

tanks at flow rates F_{03} and F_{04}, concentrations C_{A03} and C_{A04} and temperatures T_{03} and T_{04}, respectively. There is another recycle stream from CSTR 4 to CSTR 1 at flow rate F_{r2}. Three parallel irreversible exothermal reactions take place in the reactors: $A \to B$, $A \to C$, and $A \to D$. Each reactor is equipped with a jacket to provide/remove heat to/from the reactor. The dynamic equations describing temperatures and concentrations of A derived based on energy and material balances are as follows:

$$\frac{dT_1}{dt} = \frac{F_{01}}{V_1}(T_{01} - T_1) + \frac{F_{r1}}{V_1}(T_2 - T_1) + \frac{F_{r2}}{V_1}(T_4 - T_1) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A1} + \frac{Q_{h_1}}{\rho c_p V_1} \tag{3.65a}$$

$$\frac{dC_{A1}}{dt} = \frac{F_{01}}{V_1}(C_{A01} - C_{A1}) + \frac{F_{r1}}{V_1}(C_{A2} - C_{A1}) + \frac{F_{r2}}{V_1}(C_{A4} - C_{A1}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_1}} C_{A1} \tag{3.65b}$$

$$\frac{dT_2}{dt} = \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_{02}}{V_2}(T_{02} - T_2) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A2} + \frac{Q_{h_2}}{\rho c_p V_2} \tag{3.65c}$$

$$\frac{dC_{A2}}{dt} = \frac{F_1}{V_2}(C_{A1} - C_{A2}) + \frac{F_{02}}{V_2}(C_{A02} - C_{A2}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_2}} C_{A2} \tag{3.65d}$$

$$\frac{dT_3}{dt} = \frac{F_2 - F_{r1}}{V_3}(T_2 - T_3) + \frac{F_{03}}{V_3}(T_{03} - T_3) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_3}} C_{A3} + \frac{Q_{h_3}}{\rho c_p V_3} \tag{3.65e}$$

$$\frac{dC_{A3}}{dt} = \frac{F_2 - F_{r1}}{V_3}(C_{A2} - C_{A3}) + \frac{F_{03}}{V_3}(C_{A03} - C_{A3}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_3}} C_{A3} \tag{3.65f}$$

$$\frac{dT_4}{dt} = \frac{F_3}{V_4}(T_3 - T_4) + \frac{F_{04}}{V_4}(T_{04} - T_4) - \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_{i0} e^{\frac{-E_i}{RT_4}} C_{A4} + \frac{Q_{h_4}}{\rho c_p V_4} \tag{3.65g}$$

$$\frac{dC_{A4}}{dt} = \frac{F_3}{V_4}(C_{A3} - C_{A4}) + \frac{F_{04}}{V_4}(C_{A04} - C_{A4}) - \sum_{i=1}^{3} k_{i0} e^{\frac{-E_i}{RT_4}} C_{A4} \tag{3.65h}$$

where $T_j,\ C_{Aj},\ Q_j,\ V_j,\ j = 1, 2, 3, 4$ denote the temperatures, the concentrations of A, the rates of heat input/removal to/from the reactors and the reactor volumes respectively, c_p and ρ denote the heat capacity and density of the mixture in the reactors, $\Delta H_i,\ k_i,\ E_i, i = 1, 2, 3$ denote the enthalpies, pre-exponential constants and activation energies of the reactions respectively. The values of model parameters are given in Table 3.1 and the heat inputs to the reactors are $Q_{h_1} = 10^4 kJ/h$, $Q_{h_2} = 2 \times 10^4$, $Q_{h_3} = 2.5 \times 10^4 kJ/h$ and $Q_{h_4} = 10^4 kJ/h$.

Table 3.1: Process parameters for the reactors.

$F_{01} = 5\ m^3/h$	$\Delta H_1 = -5.0 \times 10^4\ KJ/kmol$
$F_1 = 35\ m^3/h$	$\Delta H_2 = -5.2 \times 10^4\ KJ/kmol$
$F_2 = 45\ m^3/h$	$\Delta H_3 = -5.04 \times 10^4\ KJ/kmol$
$F_3 = 33\ m^3/h$	$k_{10} = 3.0 \times 10^6\ h^{-1}$
$F_{02} = 10\ m^3/h$	$k_{20} = 3.0 \times 10^5\ h^{-1}$
$F_{03} = 8\ m^3/h$	$k_{30} = 3.0 \times 10^5\ h^{-1}$
$F_{04} = 12\ m^3/h$	$E_1 = 5.0 \times 10^4\ KJ/kmol$
$F_{r1} = 20\ m^3/h$	$E_2 = 7.53 \times 10^4\ KJ/kmol$
$F_{r2} = 10\ m^3/h$	$E_3 = 7.53 \times 10^4\ KJ/kmol$
$\rho = 1000.0\ kg/m^3$	$C_{A01} = 4.0\ kmol/m^3$
$c_p = 0.231\ KJ/kg$	$C_{A02} = 2.0\ kmol/m^3$
$V_1 = 1.0\ m^3/h$	$C_{A03} = 3.0\ kmol/m^3$
$V_2 = 3.0\ m^3/h$	$C_{A04} = 3.5\ kmol/m^3$
$V_3 = 4.0\ m^3/h$	$T_{03} = 300\ K$
$V_4 = 6.0\ m^3/h$	$T_{04} = 300\ K$
$T_{01} = 300\ K$	$R = 8.314\ KJ/kmol \cdot K$
$T_{02} = 300\ K$	

It is assumed that the four temperatures T_1, T_2, T_3 and T_4 are the continuously measured outputs. The objective is to estimate the entire system state based on these measurements in a distributed manner. The entire process is decomposed into four subsystems with respect to the four reactors. For each reactor, an AHG-EKF is designed. In the design of the AHG-EKFs, $Q_1 = Q_2 = Q_3 = Q_4 = \text{diag}([5, 5])$ and $R_1 = R_2 = R_3 = R_4 = 1$, respectively. The parameter θ_m is uniquely defined for all the subsystem estimators and it is selected to be $\theta_m = 20$. In the calculation of the innovation terms, the filters use $d = 0.03h$. The parameters in the adaptation functions are defined as follows: $\Delta T_i = 0.001h$, $\beta_i = 150$, $m_i = 0.1$, $\lambda_i = 100$ for $i = 1, \ldots, 4$. Random measurement noise with an upper bound $1°K$ is added to the measurements. Random additive process noise with upper bounds $1°K/sec$ and $30mol/m^3/sec$ for temperature and concentration dynamics, respectively, is also considered. The discrete communication interval among filters is $\Delta = 36sec$. In the simulations, the initial state of the process is $x(0) = [340, 2, 350, 3, 345, 2.5, 360, 4]^T$ and the initial guesses in the four distributed filters are $z_1(0) = [370, 2.5]^T$, $z_2(0) = [335, 2.7]^T$, $z_3(0) = [360, 2.65]^T$

and $z_4(0) = [390, 3.6]^T$.

First, we compare the performance of the proposed distributed AHG-EKF with a regular distributed EKF design (the proposed design with $\theta_i = 1$, $i = 1, \ldots, 4$ at all time) in the presence of measurement and process noise. In both designs, the same design parameters (i.e., Q_i and R_i with $i = 1, \ldots, 4$) are used and both designs take advantage of the predictors and exchange information as described in the proposed design. Figures 3.3 and 3.4 show the trajectories of the actual process states and the estimated values as well as the trajectories of the corresponding adaptive gains and the innovations. From Figures 3.3 and 3.4, it can be seen that both the proposed distributed AHG-EKF and the regular distributed EKF are able to track the actual process states; however, the proposed design converges to the actual states much faster. In addition, it is clear from the figures that the gains of the distributed filters in the proposed design change adaptively with the innovations. Also, it can be seen that when the estimation errors are large (at the initial period), the innovation values increase quickly which renders the four adaptive gains to increase quickly. When the estimation errors become small, the innovation values decrease and the gains decrease to one.

Next, we study the effect of adaptive gain in reducing the sensitivity of the estimates to measurement and process noise. Particularly, in this set of simulations, we fix the gains of the four distributed filters at high values (i.e., $\theta_i = 40$) and the gains do not change over the simulations. The simulation results are illustrated in Figures 3.5 and 3.6. It can be seen from Figures 3.5 and 3.6 that the estimates of concentrations are very noisy. This is expected since the high gains increase the sensitivity of the filters to uncertainties. By comparing the results shown in Figures 3.5 and 3.6 with results in Figures 3.3 and 3.4, it can be inferred that the proposed distributed AHG-EKF not only gives fast convergence, but also significantly reduces the sensitivity of the state estimates to process and measurement noises.

Finally, we evaluate the effect of the predictors in improving the performance of the

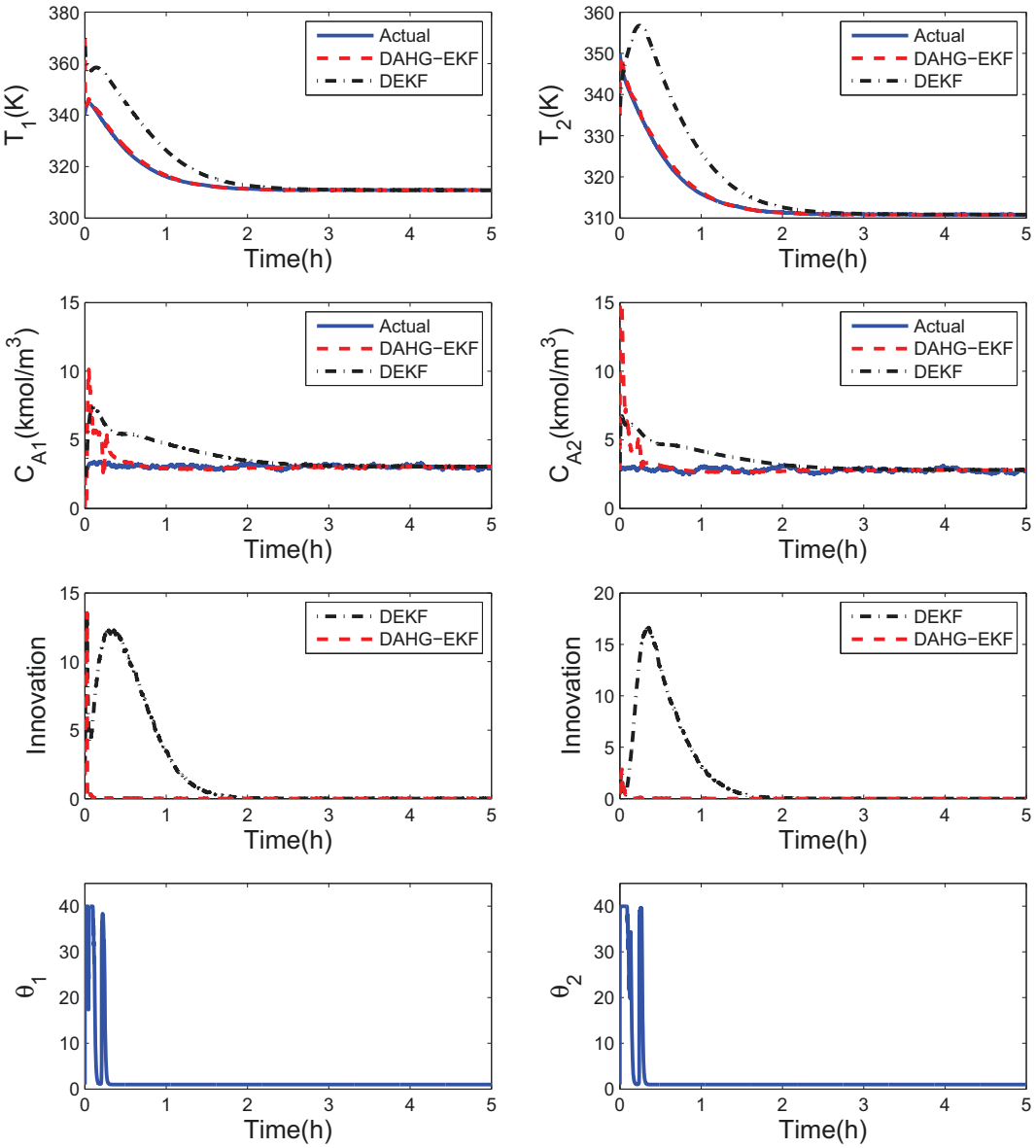

Figure 3.3: The trajectories of the states of CSTR 1 and CSTR 2, their gains and innovations under the proposed distributed AHG-EKF and regular distributed EKF.

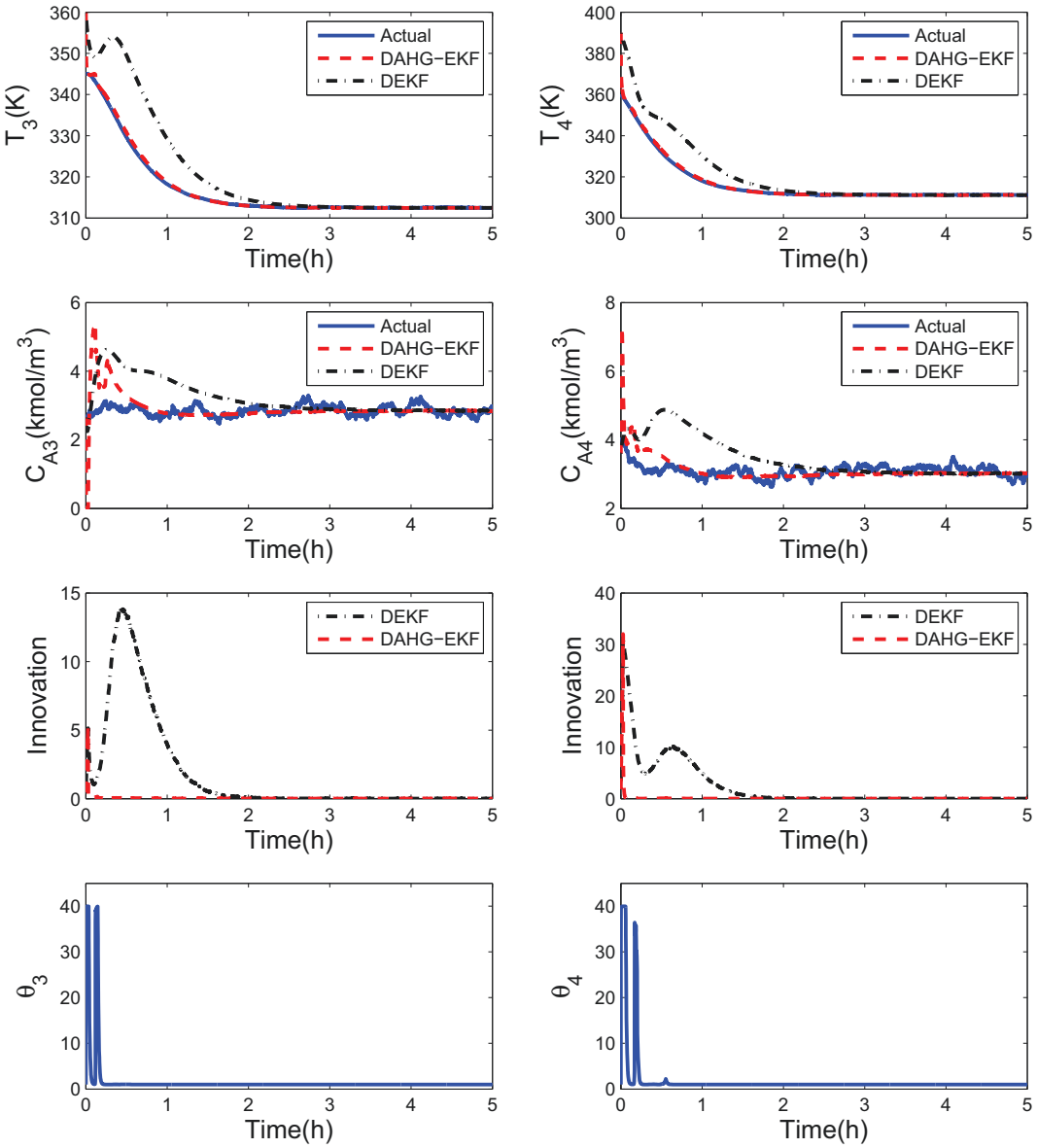

Figure 3.4: The trajectories of the states of CSTR 3 and CSTR 4, their gains and innovations under the proposed distributed AHG-EKF and regular distributed EKF.

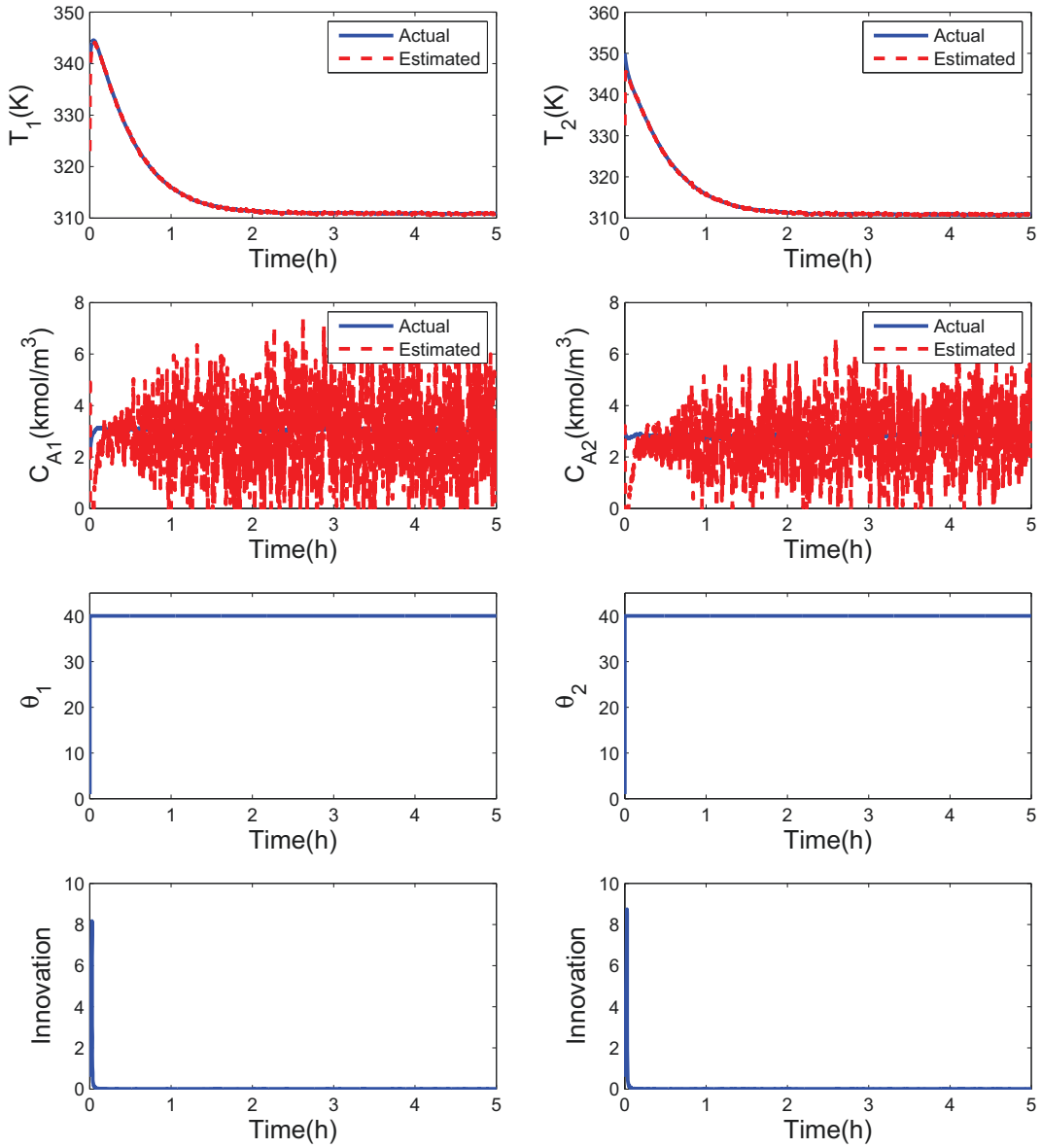

Figure 3.5: The trajectories of the states of CSTR 1 and CSTR 2 and innovations under the distributed high-gain EKF.

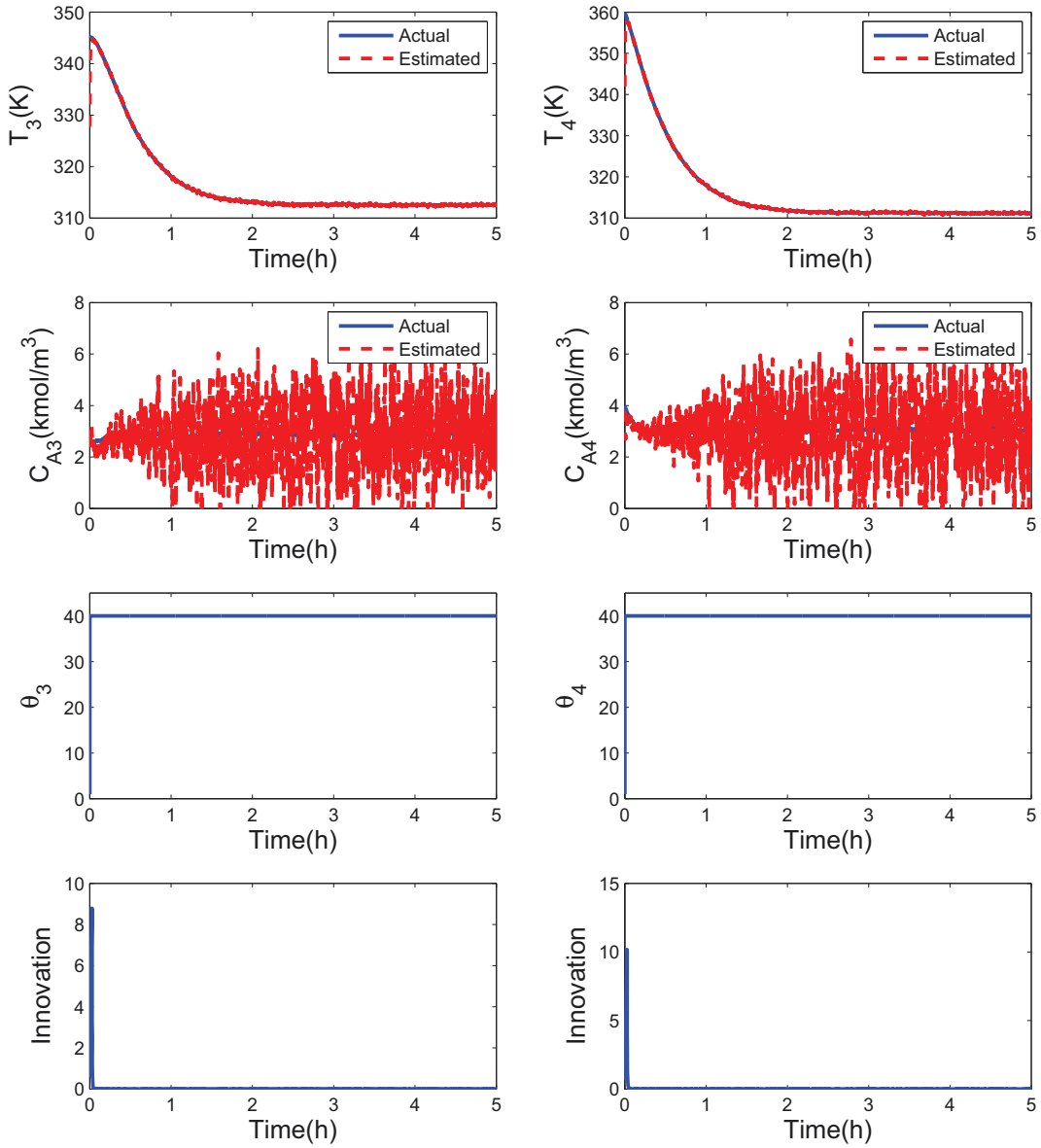

Figure 3.6: The trajectories of the states of CSTR 3 and CSTR 4 and innovations under the distributed high-gain EKF.

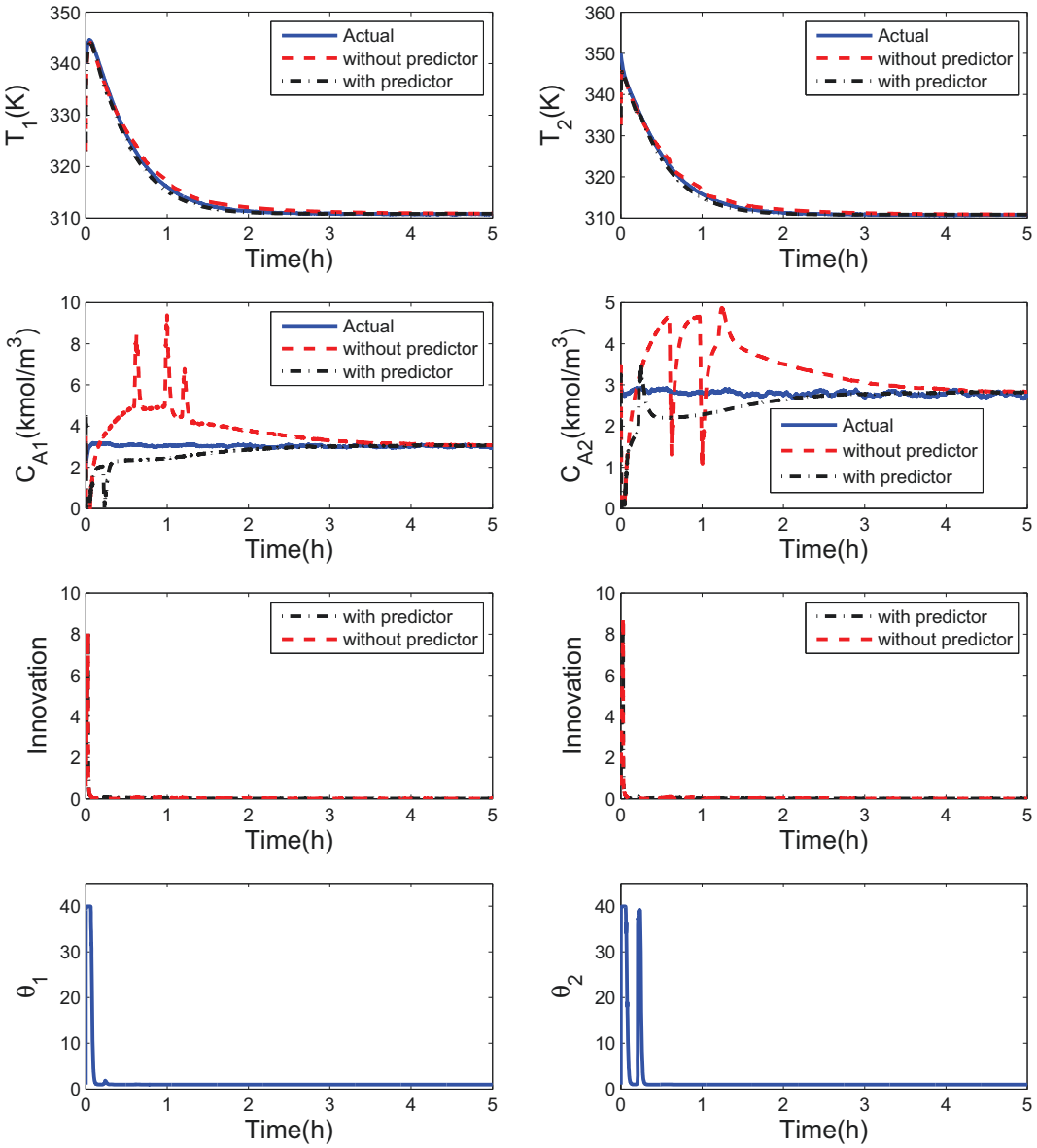

Figure 3.7: The trajectories of the states of CSTR 1 and CSTR 2, their gains and innovations under the proposed distributed AHG-EKF with and without predictors.

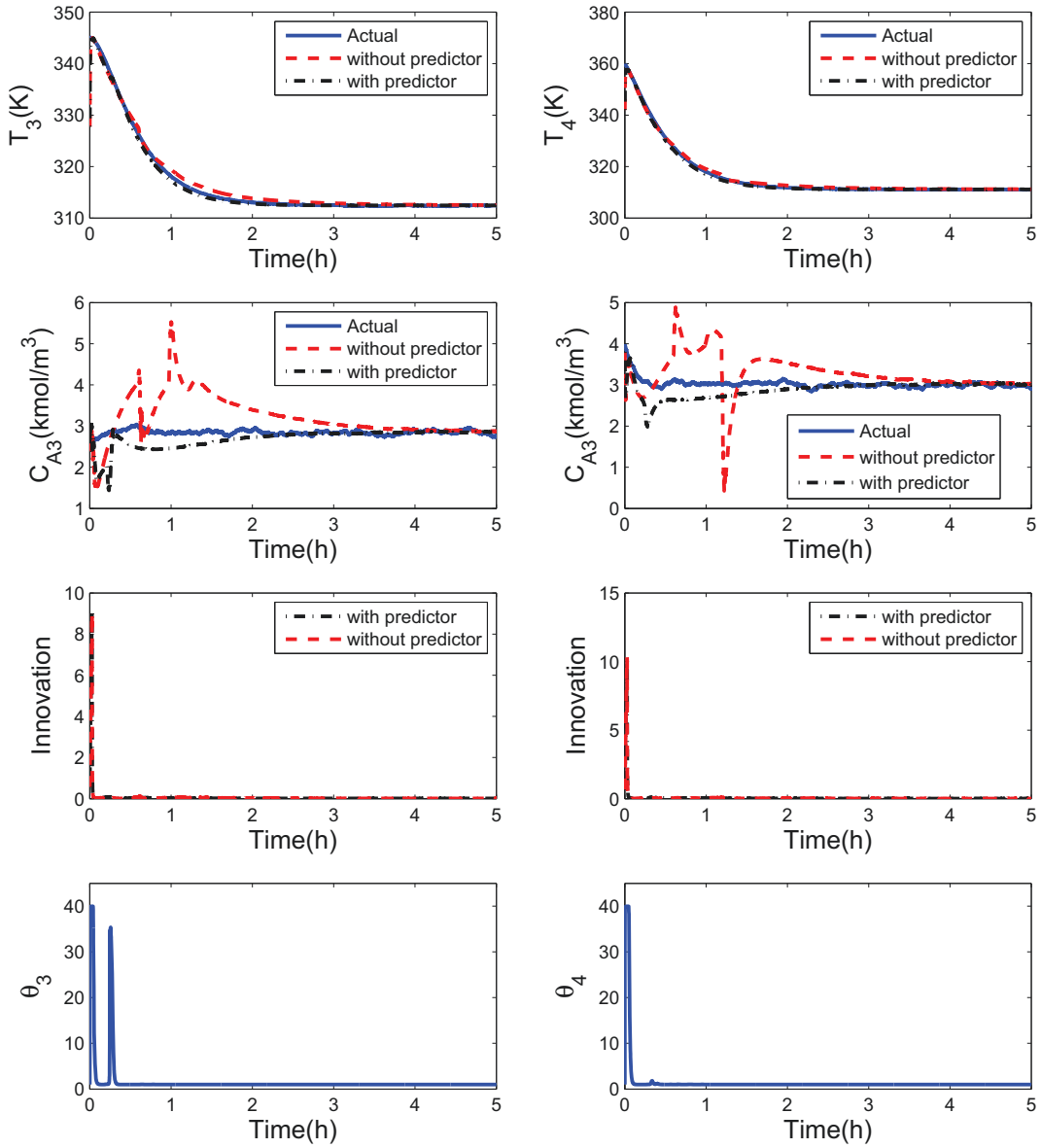

Figure 3.8: The trajectories of the states of CSTR 3 and CSTR 4, their gains and innovations under the proposed distributed AHG-EKF with and without predictors.

proposed distributed AHG-EKF. The objective of the predictors in the proposed design is to predict the states of other subsystems during communication intervals based on latest received information. In this set of simulations, we carry out simulations with and without the predictors. Figures 3.7 and 3.8 show the simulation results. It should be noted that, when there is no predictor, the interactions between the subsystems are approximated using the latest available information. From Figures 3.7 and 3.8, it can be seen that the estimates given by the proposed design with predictors converge faster to the actual states. Also, the trajectories of the estimates given by the proposed design with predictors are much smoother than the ones obtained without predictors. In other words, the use of the state predictors contributes to the significant improvement of the estimation performance.

3.8 Conclusions

In this chapter, we took a more realistic consideration into account and designed a distributed adaptive high-gain EKF with discrete communication for a type of nonlinear systems composed of interacted subsystems. Considering continuous measurements, the distributed filters are designed to compensate for interactions between subsystems by exchanging information at discrete time instants. An implementation strategy is designed based on discrete communication among distributed filters which describes how the filters should communicate. Then, to enhance the performance of the local estimators, state predictors are designed for each subsystem based on the deterministic model of the centralized system. Sufficient conditions under which the proposed distributed filtering approach is stable were derived. The proposed approach was applied to a simulated chemical process example to illustrate its applicability and effectiveness.

Chapter 4

Distributed adaptive high-gain extended Kalman filters with communication delays and data losses

4.1 Introduction

In this section, we introduce the proposed distributed state estimation design taking into account communication delays and data losses. A schematic view of the proposed design is shown in Figure 4.1. In this design, each subsystem is assigned to a local estimator and each estimator contains an AHG-EKF and a state predictor. The distributed estimators transmit the information through a communication network. The communication between subsystem estimators is subject to time-varying delays and data losses. The filter of a subsystem estimates the subsystem state based on the subsystem output measurements, the information received from other filters and the predictions provided by the associated predictor.

In this chapter, the system under consideration is similar to the general form of the described system in (3.3), and the distributed filters will be designed based on both the

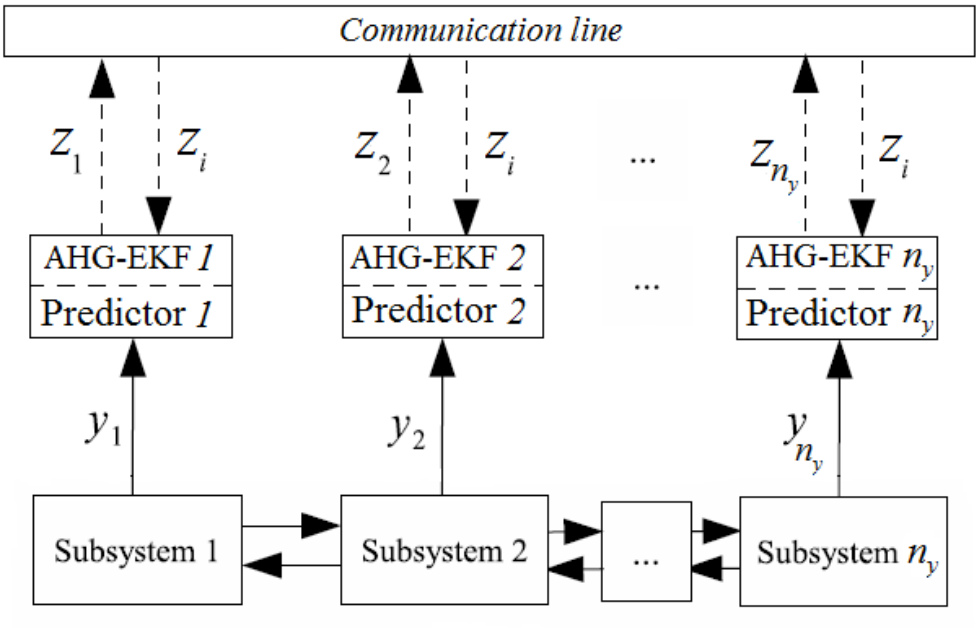

Figure 4.1: Proposed distributed state estimation design subject to communication delays and losses.

deterministic and stochastic forms of subsystems in (3.1).

4.2 Modeling of communication network and measurements

In this chapter, each subsystem estimator is assumed to have immediate and direct access to the measurements of its corresponding subsystem continuously. The subsystem estimators are also assumed to have the capability of information transmission in a mutual communication network and the exchange of information is subject to time-varying delays and data losses at discrete time instants $\{t_{k\geq 0}\}$ such that $t_k = t_0 + k\Delta$ with the initial time $t_0 = 0$, a fixed time interval Δ and a positive integer k. An auxiliary variable $d_{i,j}(t_k)$ is incorporated to denote the associated delay for the information of subsystem j available to subsystem i at time instants t_k, and the possible values for the variable $d_{i,j}(t_k)$ are positive integers. It is assumed that there is a predetermined maximum allowable D on the communication delay

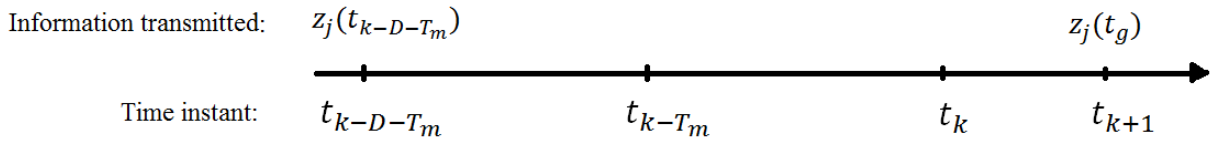

Figure 4.2: The worst case scenario of the available information of filter j to filter i.

$d_{i,j}(t_k)$, and the data received with time delay larger than D is taken as lost information. The data losses are also associated with an upper bound T_m on the number of consecutive sampling periods. Because of the existence of D and T_m, a subsystem will receive at least once information from another subsystem within $D + T_m$ communication periods (i.e., $(D + T_m)\Delta$). Note that both the upper bounds D and T_m are critical parameters to obtain deterministic results for the proposed distributed state estimation scheme.

In distributed filtering with communication delays and data losses, a worst case scenario may occur in which none of the communicating filters transmit information within the period $(D+T_m)\Delta$. To elaborate this scenario, let us consider two filters i and j, where filter j sends information (z_j) to filter i at time t_{k-D-T_m}. As shown in Figure 4.2, the worst case scenario occurs when no information is received by filter i up to time t_k. Then, at the next sampling time t_{k+1}, filter i must receive another update of $z_j(t_g)$ with $t_g > t_{k-D}$.

4.3 Distributed state estimation algorithm

The following algorithm describes the information flow between the distributed estimators and the roles of the predictors and the filters:

1. At initial time instant $t = 0$, all of the subsystem filters are initialized with guesses for subsystem state (i.e. $z_j(0)$, $j \in \mathbb{I}$), the actual subsystem output measurement (i.e. $y_i(0)$), and the initial value of each filter's adaptive gain (i.e. $\theta_i(0)$).

2. At time instant $t_k \geq 0$ the following steps should be taken:

(a) AHGEKF i receives the local output measurement $y_i(t_k)$.

(b) If any new data packages are received by AHGEKF i between the time interval t_{k-1} and t_k, the data will be stored if time delay is less than the maximum allowable delay (i.e. $d_{i,j}(t_k) < D$); otherwise the data package is discarded. If the received data provides more recent information, $d_{i,j}(t_k)$ values are updated and the data packages are stored; else only the data packages are stored.

(c) If any $d_{i,j}(t_k)$ is greater than 1, then based on previously received information, the corresponding predictor of AHGEKF i predicts the system state $x^{p,i}(t_{k-1})$. Then the state estimate for subsystem i is calculated based on local measurement $y_i(t_k)$, the state prediction $x^{p,i}(t_{k-1})$ and the received information from other subsystems.

(d) AHGEKF i sends the current estimated state $z_i(t_k)$ to AHGEKF j, $j \in \mathbb{I}_i$, $j \neq i$.

3. Filter i and predictor i perform state estimation and state prediction continuously in the period t_k and t_{k+1}. Return to step 2 at the next sampling time (and $k \to k+1$) when $t = t_{k+1}$.

4.4 Design of subsystem predictors

At a discrete-time instant t_k, each filter in the distributed estimation system sends out its latest state estimate to other filters. However, due to communication delays and losses, the information sent by a filter may not be received by other filters immediately. Since a filter requires the state information of other filters to characterize its interaction of its corresponding subsystem with other subsystems, the unavailable information is predicted by an open loop predictor for each subsystem filter. In order to reduce the number of evaluations of the predictor associated with a subsystem, the predictor only calculates at the communication time instants (i.e. $\{t_{k\geq 0}\}$).

At time t_k, predictor i generates a prediction of the entire system state, denoted as $x^{p,i}(t_k)$, using the nominal centralized system model of (3.3). A two-step prediction-update

algorithm is used to calculate the predictions recursively from t_f to t_{k-1}, where the time t_f is determined as follows: (1) If filter i receives information between t_{k-1} and t_k, then t_f would be the time when the oldest state estimate was received. (2) If filter i receives no information between t_{k-1} and t_k, then $t_f = t_{k-1}$.

For each sampling time period, $[t_g, t_{g+1}]$, between t_f and t_{k-1}:

- Prediction step. The unreceived state estimates from other filters will be approximated by the integration of the nominal centralized system model within the period $t \in [t_g, t_{g+1})$:

$$\begin{aligned} \dot{x}_-^{p,i}(t) &= Ax_-^{p,i}(t) + b(x_-^{p,i}(t), 0) \\ x_-^{p,i}(t_g) &= x^{p,i}(t_g) \end{aligned} \tag{4.1}$$

  with $x_-^{p,i}$ denoting the prediction prior to the update step and $x^{p,i}(t_g)$ representing the prediction at t_g after the update step. Also, as denoted in (4.1), in this chapter we consider the predictor to be function of states only. It should be noted that the initial condition $x^{p,i}(t_f)$ should be updated first.

- Update step. To obtain the updated predictions $x^{p,i}(t_{g+1})$, the corresponding portions of $x_-^{p,i}(t_{g+1})$ are needed to be replaced with the available subsystem state estimates at t_{g+1} (including $z_i(t_{g+1})$ and the state estimates received from other filters). For instance, if filter i receives $z_j(t_{g+1})$ within $[t_{k-1}, t_k]$ or previously, $z_j(t_{g+1})$ replaces the portion of the prediction for subsystem j in $x_-^{p,i}(t_{g+1})$.

The mentioned methodology describes that all the received subsystem state estimates will be used to update the predictions (obtained at t_{k-1}). Also, it can be inferred that the filters need to store the received information and in the absence of delay or data losses, no prediction-update method is necessary.

4.5 Deterministic system

In this section, we design the distributed filters subject to communication delays and data losses based on the deterministic form of subsystems in (3.1). The deterministic form of subsystems is defined as:

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i(t) + b_i(x(t), 0) \\ y_i &= C_i x_i(t) \end{aligned} \tag{4.2}$$

with the same properties described in Section 3.2. In this section, we assume b_i is only a nonlinear function of overall states and from now on we denote $b_i(x(t), 0)$ as $b_i(x(t))$.

4.5.1 Design of subsystem filters

In the proposed design, filter i ($i \in \mathbb{I}$) is defined as follows:

$$\dot{z}_i = A_i z_i + b_i(x^{p,i}) - S_i^{-1} C_i^T R_{\theta_i}^{-1} (C_i z_i - y_i) \tag{4.3}$$

where z_i is the state of the filter, $x^{p,i}$ is composed of the corresponding filter state z_i and the predictions of the states of all other filters $x_j^{p,i}$, $j \in \mathbb{I}$, $j \neq i$, and $R_{\theta_i} = \dfrac{1}{\theta_i} R_i$ with θ_i being the adaptive gain parameter of filter i and R_i being a positive scalar. Also, S_i is the solution to the following matrix Riccati equation:

$$\dot{S}_i = -\left(A_i + b_i^*(x^{p,i})\right)^T S_i - S_i \left(A_i + b_i^*(x^{p,i})\right) + C_i^T R_{\theta_i}^{-1} C_i - S_i Q_{\theta_i} S_i \tag{4.4}$$

where $b_i^*(x^{p,i})$ denotes the Jacobian of $b_i(x^{p,i})$ with respect to z_i (i.e. $b_i^*(x^{p,i}) = \frac{\partial b_i(x^{p,i})}{\partial z_i}$) and $Q_{\theta_i} = \theta_i Q_i$ with Q_i being an n_{x_i} by n_{x_i} symmetric positive definite matrix.

On the right-hand-side of (4.3), the first two terms are from the subsystem model (4.2) with the interacting states approximated using the predicted states to account for communication delays and losses. In order to adapt the high-gain parameter, we follow the same strategy that is described in Section 3.5, since the state predictors in both designs operate

similarly.

4.5.2 Stability analysis

In this section, the stability of the proposed distributed AHG-EKF design affording communication delays and information losses is investigated. Defining the estimation error for each subsystem as $\epsilon_i = z_i - x_i$, $i \in \mathbb{I}$, the subsystem's error dynamics would be

$$\dot{\epsilon}_i = \dot{z}_i - \dot{x}_i = \left(A_i - S_i^{-1} C_i^T R_{\theta_i}^{-1} C_i\right) \epsilon_i + b_i(x^{p,i}) - b_i(x) \tag{4.5}$$

Picking $\epsilon_i^T S_i \epsilon_i$ as the Lyapunov function for subsystem i, $i \in \mathbb{I}$, we can obtain that:

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} = \theta_i \left[-\epsilon_i^T S_i Q_i S_i \epsilon_i - \epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i + \frac{2}{\theta_i}\Big(b_i^T(x^{p,i}) S_i \epsilon_i - b_i^T(x) S_i \epsilon_i - \epsilon_i^T S_i b_i^*(x^{p,i}) \epsilon_i \Big) \right] \tag{4.6}$$

The equation (4.6) plays an important role in establishing the convergence of the proposed distributed state estimation approach with communication delays and data losses. The properties of the predictors are provided in the following Proposition 3. In this proposition, the evolution of prediction error is evaluated and an upper bound is established between the states of the actual system and the predictions provided by the predictors.

Proposition 3. *Consider the following centralized system and state predictor:*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + b(x(t)) \\ \dot{x}^p(t) &= Ax^p(t) + b(x^p(t)) \end{aligned} \tag{4.7}$$

where $x^p(t)$ is the prediction of $x(t)$, $A = diag\{[A_1, A_2, \ldots, A_{n_y}]\}$ and $b(x) = [b_1^T(x), \ldots, b_{n_y}^T(x)]^T$. The system prediction error $e = x^p - x$ satisfies

$$\|e(t)\| \leq f_e(t - t_0, \|e(t_0)\|) \tag{4.8}$$

for all $x(t), x^p(t) \in \mathbb{X} \subset \mathbb{R}^n$, *where* $f_e(t-t_0, \|e(t_0)\|) = \sqrt{n}\|e(t_0)\|\exp[(L_b+1)(t-t_0)]$ *with* L_b *denoting the Lipschitz constant of* $b(x)$ *with respect to* x.

Proof: From (4.7), the time derivative of the prediction error can be obtained as

$$\dot{x}^p - \dot{x} = \dot{e} = Ae + b(x^p) - b(x) \tag{4.9}$$

Due to the continuity of x and x^p and Lipschitz property of b, there exists a constant L_b such that,

$$\|b(x^p) - b(x)\|_1 \leq L_b\|x^p - x\|_1 \tag{4.10}$$

and consequently we obtain,

$$\|b(x^p) - b(x)\|_1 = \left\| \begin{bmatrix} b_1(x^{p,1}) - b_1(x) \\ \vdots \\ b_{n_y}(x^{p,n_y}) - b_{n_y}(x) \end{bmatrix} \right\|_1 \leq L_b\|e\|_1 \tag{4.11}$$

Based on the structure of matrix A in Brunovsky canonical form, it can be obtained that $\|A\|_1 = 1$. Also from (4.9) we obtain

$$\frac{d}{dt}\|e\|_1 \leq \|\dot{e}\|_1 \leq \|A\|_1\|e\|_1 + L_b\|e\|_1 = (L_b+1)\|e\|_1 \tag{4.12}$$

Since $\|e\| \leq \|e\|_1 \leq \sqrt{n}\|e\|$, from the solution to (4.12), it can be obtained that

$$\|e(t)\| \leq \sqrt{n}\|e(t_0)\|\exp[(L_b+1)(t-t_0)] = f_e(t-t_0, \|x^p(t_0) - x(t_0)\|) \tag{4.13}$$

and this proves Proposition 3. □

Theorem 4. *Considering system (3.3) defined by subsystems in the form of (3.1), if the following conditions are satisfied,*

1. *the nonlinear function,* $b(x)$*, is Lipschitz with respect to* x*,*

2. *the Jacobian matrix, $b_i^*(x)$, is bounded,*

3. *the following conditions hold,*

$$q_{m_{min}} > 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta^2_{min}} \tag{4.14}$$

and

$$\frac{2\sqrt{n}L_b e^{(L_b+1)(D+T_m)\Delta}}{\min_i\{\theta_i\} q_{m_{min}}\delta_{min} - 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}}} < \frac{\delta_{min}}{\delta_{max}} \left[1 - \left(\frac{2\theta_m}{R_{min}\delta_{min}} + L_b\sqrt{n}e^{(L_b+1)(T_m+D)\Delta}\right)(T_m + D)\Delta\right]^2 \tag{4.15}$$

where $\delta_{min} = \min_i\{\delta_{min_i}\}$, $\delta_{max} = \max_i\{\delta_{max_i}\}$, $R_{min} = \min_i\{R_i\}$, Δ is the communication interval, D denotes the maximum allowable delay and T_m represents the maximum consecutive samples of data losses,

then for any initial condition of subsystems and observers in $\mathbb{X}_i$ for all $i \in \mathbb{I}$, the error dynamics is asymptotically stable and the error will eventually converge to zero.

Proof: Since $R_i \geq 0$, we can verify that $\epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i \geq 0$. Also, considering a positive definite matrix Q_i, there exists a q_{m_i} such that $Q_i \geq q_{m_i} I$. Using (4.6) it can be obtained that

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \epsilon_i^T S_i^2 \epsilon_i + 2\epsilon_i^T S_i \left(b_i(x^{p,i}) - b_i(x) - b_i^*(x^{p,i})\epsilon_i\right) \tag{4.16}$$

Based on Theorem 2.18 in [60], the Riccati matrix in each subsystem can be bounded such that $\delta_{min_i} I \leq S_i(t) \leq \delta_{max_i} I$ where $\delta_{min_i} > 0$ and $\delta_{max_i} > 0$ are scalars. So, from (4.16) we can obtain that

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \delta_{min_i} \epsilon_i^T S_i \epsilon_i + 2\epsilon_i^T S_i \left(b_i(x^{p,i}) - b_i(x) - b_i^*(x^{p,i})\epsilon_i\right) \tag{4.17}$$

In order for the evaluation of the overall system's stability, we add (4.17) for all $i \in \mathbb{I}$ together

to obtain,

$$\begin{aligned}
&\frac{d}{dt}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i \leq -\min_i\{\theta_i q_{m_i}\delta_{\min_i}\} \\
&\times\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}^T\begin{bmatrix}S_1 & 0 & \dots & 0\\ 0 & S_2 & 0 & \vdots\\ \vdots & & \ddots & \\ 0 & \dots & & S_{n_y}\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}+2\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}^T\begin{bmatrix}S_1 & 0 & \dots & 0\\ 0 & S_2 & 0 & \vdots\\ \vdots & & \ddots & \\ 0 & \dots & & S_{n_y}\end{bmatrix} \\
&\times\left(\begin{bmatrix}b_1(x^{p,1})-b_1(x)\\ b_2(x^{p,2})-b_2(x)\\ \vdots\\ b_{n_y}(x^{p,n_y})-b_{n_y}(x)\end{bmatrix}-\begin{bmatrix}b_1^*(x^{p,1}) & 0 & \dots & 0\\ 0 & b_2^*(x^{p,2}) & 0 & \vdots\\ \vdots & & \ddots & 0\\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y})\end{bmatrix}\begin{bmatrix}\epsilon_1\\ \epsilon_2\\ \vdots\\ \epsilon_{n_y}\end{bmatrix}\right)
\end{aligned} \tag{4.18}$$

As explained in Section 4.2, the worst case scenario of the communication between two filters occurs when there is no information transmission among them within $D + T_m$ consecutive samples.

Let us focus on the time interval $t \in [t_{k-D-T_m}, t_k]$ and assume that there is no information transmission between the filters in the time period. Based on Proposition 3 for $t \in [t_{k-D-T_m}, t_k)$, we would have

$$\|b(x^p) - b(x))\| \leq L_b\|e(t)\| \leq L_b\sqrt{n}\|e(t_{k-D-T_m})\|\exp[(L_b+1)(t-t_{k-D-T_m})] \tag{4.19}$$

Noting that $\epsilon(t_{k-D-T_m}) = e(t_{k-D-T_m})$, the Lipschitz property of nonlinear function b for the

last term of (4.18) results in,

$$\begin{aligned}
&\left\| \begin{bmatrix} b_1(x^{p,1}) - b_1(x) \\ b_2(x^{p,2}) - b_2(x) \\ \vdots \\ b_{n_y}(x^{p,n_y}) - b_{n_y}(x) \end{bmatrix} - \begin{bmatrix} b_1^*(x^{p,1}) & 0 & \dots & 0 \\ 0 & b_2^*(x^{p,2}) & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y}) \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{n_y} \end{bmatrix} \right\| \\
&\leq \left\| \begin{bmatrix} b_1(x^{p,1}) - b_1(x) \\ b_2(x^{p,2}) - b_2(x) \\ \vdots \\ b_{n_y}(x^{p,n_y}) - b_{n_y}(x) \end{bmatrix} \right\| + \left\| \begin{bmatrix} b_1^*(x^{p,1}) & 0 & \dots & 0 \\ 0 & b_2^*(x^{p,2}) & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & b_{n_y}^*(x^{p,n_y}) \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_{n_y} \end{bmatrix} \right\| \\
&\leq L_b\sqrt{n}\|\epsilon(t_{k-D-T_m})\|\exp[(L_b+1)(t-t_{k-D-T_m})] + n_y L_{b^*_{max}}\|\epsilon(t)\|
\end{aligned} \tag{4.20}$$

where $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$. From (4.20) and (4.18), the following inequality can be obtained

$$\begin{aligned}
&\frac{d}{dt}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) \leq -\min_i\{\theta_i q_{m_i}\delta_{\min_i}\}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) + 2n_y L_{b^*_{max}}\frac{\delta_{\max}}{\delta_{\min}}\sum_{i=1}^{n_y}\epsilon_i^T S_i\epsilon_i(t) \\
&+2\sqrt{n}L_b\delta_{\max}\exp[(L_b+1)(t-t_{k-D-T_m})]\|\epsilon^T(t)\|\|\epsilon(t_{k-D-T_m})\|
\end{aligned} \tag{4.21}$$

Let us assume that $\zeta \in [t_{k-D-T_m}, t_k)$ is the time at which the estimation error takes the maximum value over the time period, i.e. $\|\epsilon(\zeta)\| \geq \|\epsilon(t)\|$ for any $t \in [t_{k-D-T_m}, t_k)$. Then we have

$$\|\epsilon(t)\|\|\epsilon(t_{k-D-T_m})\| \leq \|\epsilon(\zeta)\|^2 \tag{4.22}$$

Also, we can further obtain that

$$\sum_{i=1}^{n_y}\epsilon_i^T(t)S_i(t)\epsilon_i(t) \leq \sum_{i=1}^{n_y}\epsilon_i^T(\zeta)S_i(t)\epsilon_i(\zeta) \leq \delta_{max}\|\epsilon(\zeta)\|^2 := \Omega\left(\epsilon(\zeta)\right) \tag{4.23}$$

Defining $V(t) = \sum_{i=1}^{n_y} \epsilon_i^T S_i \epsilon_i(t)$ and substituting (4.23) into (4.21) we get

$$\begin{aligned} \frac{d}{dt}V(t) \leq & \left(-\min_i\{\theta_i q_{m_i}\delta_{\min_i}\} + 2n_y L_{b^*_{max}} \frac{\delta_{\max}}{\delta_{\min}} \right) V(t) \\ & + (2\sqrt{n}L_b\Omega(\epsilon(\zeta)))\exp[(L_b+1)(t-t_{k-D-T_m})] \end{aligned} \tag{4.24}$$

On the other hand, from (4.5) for the overall system we obtain

$$\dot{\epsilon}(t) = A\epsilon(t) - \begin{bmatrix} S_1^{-1}C_1^T R_{\theta_1}^{-1} C_1 & 0 & \dots & 0 \\ 0 & S_2^{-1}C_2^T R_{\theta_2}^{-1} C_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & & S_{n_y}^{-1}C_{n_y}^T R_{\theta_{n_y}}^{-1} C_{n_y} \end{bmatrix} \epsilon(t) + b(x^p) - b(x) \tag{4.25}$$

Since $\delta_{min_i} I \leq S_i(t) \leq \delta_{max_i} I$ and $1 \leq \theta_i \leq 2\theta_m$ we can have,

$$\|S_i^{-1}C_i^T R_{\theta_i}^{-1} C_i\| \leq \|S_i^{-1}\|\|C_i^T\|\|R_{\theta_i}^{-1}\|\|C_i\| \leq \frac{2\theta_m}{\delta_{min_i} R_i} \tag{4.26}$$

Hence,

$$\left\| \begin{bmatrix} S_1^{-1}C_1^T R_{\theta_1}^{-1} C_1 & 0 & \dots & 0 \\ 0 & S_2^{-1}C_2^T R_{\theta_2}^{-1} C_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & & S_{n_y}^{-1}C_{n_y}^T R_{\theta_{n_y}}^{-1} C_{n_y} \end{bmatrix} \right\| \leq \frac{2\theta_m}{\delta_{min} R_{min}} \tag{4.27}$$

By considering a fixed communication interval $(T_m + D)\Delta$, from (4.25), (4.27) and (4.19) we can have,

$$\|\dot{\epsilon}(t)\| \leq (\frac{2\theta_m}{\delta_{min} R_{min}} + 1)\|\epsilon(t)\| + L_b\sqrt{n}e^{(L_b+1)(T_m+D)\Delta}\|\epsilon(t_{k-D-T_m})\| \tag{4.28}$$

and since $\|\epsilon(\zeta)\| \geq \|\epsilon(t)\|$ for any $t \in [t_{k-D-T_m}, t_k)$ we obtain,

$$\|\dot{\epsilon}(t)\| \leq \left(\frac{2\theta_m}{\delta_{min} R_{min}} + 1 + L_b \sqrt{n} e^{(L_b+1)(D+T_m)\Delta} \right) \|\epsilon(\zeta)\| \tag{4.29}$$

The norm of estimation error within the time period $t \in [t_{k-D-T_m}, t_k)$ satisfies

$$\|\epsilon(t)\| \geq \|\epsilon(\zeta)\| - \max\{\|\dot{\epsilon}(t)\|\}(T_m + D)\Delta \tag{4.30}$$

and since

$$V(t) = \epsilon^T(t) S(t) \epsilon(t) \geq \delta_{min} \min\{\|\epsilon(t)\|^2\} \tag{4.31}$$

plugging (4.29) and (4.30) into (4.31) results in

$$\frac{V(t)}{\Omega(\zeta)} \geq \frac{\delta_{min}}{\delta_{max}} \left[1 - \left(\frac{2\theta_m}{\delta_{min} R_{min}} + 1 + L_b \sqrt{n} e^{(L_b+1)\Delta(T_m+D)} \right) (T_m + D) \right]^2 \tag{4.32}$$

If (4.14) and (4.15) are satisfied, from (4.32) it can be obtained that

$$\frac{2\sqrt{n} L_b e^{(L_b+1)(D+T_m)\Delta}}{\min_i\{\theta_i\} q_{m_{min}} \delta_{min} - 2 n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}}} < \frac{V(t)}{\Omega(\epsilon(\zeta))} \tag{4.33}$$

and as a result,

$$0 > 2\sqrt{n} \Omega(\epsilon(\zeta)) L_b e^{(L_b+1)(D+T_m)\Delta} + \left(- \min_i\{\theta_i\} q_{m_{min}} \delta_{min} + 2 n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta_{min}} \right) V(t) \tag{4.34}$$

Plugging (4.34) into (4.24) leads to $\dot{V}(t) < 0$, which means that the error dynamics is asymptotically stable. This proves Theorem 4.

Remark 7. *Referring to condition (4.15) in Theorem 4 (as well as (4.33) in the proof), it can be seen that with increased gain θ_i, a smaller value can be obtained for the lower bound of $V(t)$ which implies that the error decreases faster. In the next section, we will demonstrate in simulations that high gains in the distributed filters ensure fast convergence and adaptation*

of the gains according to innovation can remove the sensitivity of high gains to noise.

4.6 Stochastic system

In this section, the general form of nonlinear systems including the uncertainties in the process and measurements is considered. The system is decomposed into the subsystems in the form of (3.1) with the same properties that are described in Section 3.2. In Section 3.6, the distributed filters were designed for the stochastic form of the interconnected subsystems based on discrete communications where the distributed filters communicated at discrete time instants simultaneously. In this section we will analyze the distributed filters for stochastic processes where they not only communicate discretely but also are subject to delays and data dropouts in information transmission. Also, the state predictors are designed for distributed filters to compensate for the missing information within delays and data losses.

4.6.1 Design of subsystem filters

The filter design in this section is similar to the that of section 3.6, however the information transmission time instants randomly change among different filters. The predictors are designed based on predict-update strategy for each filter to predict the delayed and lost information and update the predictions at the information arrival time.

4.6.2 Stability analysis

In order to have a comprehensive analysis of the proposed distributed filtering approach with process and measurement uncertainties, we consider the worst case scenario in communication of distributed filters. As described in Subsection (4.2), the communications among filters are allowed for having bounded delays and bounded consecutive dropouts in information transmission. Consequently, the distributed filters will experience the worst condition in information exchange when none of the distributed filters have received information from the

communication network after the maximum possible sampling time delays and consecutive data losses, i.e. $(D+T_m)\Delta$. This case of communication provides a discrete communication scenario in which the distributed filters communicate every $(D+T_m)$ sampling times.

Theorem 5. *Let us consider a network of interacting subsystems described by (3.1) with local AHGEKFs designed by (3.44)-(3.45). If the following conditions are satisfied:*

1. *subsystems' nonlinear function $b(\cdot,\cdot)$ is locally Lipschitz with respect to its arguments*
2. *the Jacobian $b_i^*(x^{p,i}(t))$ is bounded in the entire operating region of the distributed filters,*
3. *the Riccati matrices S_i with $i \in \mathbb{I}$ are bounded (i.e., the conditions in Theorem 2.18 of [60] are satisfied),*
4. *the elements of covariance matrix $Q_i \geq q_{m_i} I$ satisfy*

$$q_{m_{min}} > 2n_y L_{b^*_{max}} \frac{\delta_{max}}{\delta^2_{min}} \tag{4.35}$$

   where $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$ with $L_{b_i^*}$ being the Lipschitz constant associated with b_i^*, and δ_{max} and δ_{min} are the maximum and minimum values of the upper and lower bounds of all the subsystem Riccati matrices S_i,

5. *considering d_m as the maximum distance between any two points in $\mathbb{X}$, Δ as the maximum allowable sample delay and T_m as the maximum consecutive data loss in communication among AHGEKFs and $R_{min} = \min\{R_i\}$,*

then for any initial condition of subsystems and observers in $\mathbb{X}_i$ (i.e. $x_i(t_0) \in \mathbb{X}_i$ and $z_i(t_0) \in \mathbb{X}_i$) for all $i \in \mathbb{I}$, the norm of the overall system's estimation error $\|\epsilon(t)\|$ is descending and ultimately smaller than ϵ^*; that is,

$$\lim_{t\to\infty} \|\epsilon(t)\| \leq \epsilon^* \tag{4.36}$$

where $\epsilon^* = \frac{2\delta_{max} L_b^x e^{(L_b^x+1)\Delta(D+T_m)}\left(\sqrt{n}d_m + \frac{2L_b^w \phi_w}{L_b^x+1}\right) + \frac{2L_b^w \phi_w}{L_b^x+1}\delta_{max} + \frac{4\theta_m n_y \phi_v}{R_{min}}}{\min\{\theta_i q_{m_i}\}\alpha^2_{min} - 2\delta_{max} n_y L^*_{b_{max}}}$.

Proof: In this section, since the model of the nonlinear subsystems and communication of their corresponding filters are the same as the described model in Section 3.6, the filter dynamics and the Riccati equation will be the same as (3.44) and (3.45), respectively. Also, it can be verified that by considering the same model of the centralized system and predictor, Proposition 2 holds as well. By considering the worst case scenario in communication delays and data losses among the distributed filters, we may follow the same proof described in Section 3.6; however in this scenario the analysis is performed within $t \in [t_{k-D-T_m}, t_k)$. Consequently, by picking $V_i = \epsilon_i(t)S_i(t)\epsilon_i(t)$, and $V = \sum_{i=1}^{n_y} V_i$ as the Lyapunov functions for each subsystem and the entire system, respectively, the equations (3.57)-(3.59) hold and based on Proposition 2 for $t \in [t_{k-D-T_m}, t_k]$ we obtain,

$$\begin{aligned}\|b(x^p,0)-b(x,w)\| \leq & \left((\sqrt{n}\|e(t_{k-D-T_m})\|+2\tfrac{L_b^w\phi_w}{L_b^x+1})\exp[(L_b^x+1)(t-t_{k-D-T_m})]-\tfrac{L_b^w\phi_w}{L_b^x+1}\right) \\ & \times(L_b^x+1)+L_b^w\phi_w\end{aligned} \tag{4.37}$$

In the worst case scenario, since the last communication among filters has been made at time t_{k-D-T_m}, then $\epsilon(t_{k-D-T_m}) = e(t_{k-D-T_m})$. So, based on (4.37) and Proposition 2, we can obtain

$$\begin{aligned}\frac{dV(t)}{dt} \leq & -\min\{\theta_i q_{m_i}\delta_{min_i}\}V(t)+2\|\epsilon(t)\|^2\|S(t)\|n_y L_{b^*_{max}}+\frac{4\theta_m n_y\phi_v}{R_{min}}\|\epsilon(t)\| \\ & +2(L_b^w\phi_w-\frac{L_b^w L_b^x\phi_w}{L_b^x+1})\|\epsilon(t)\|\|S(t)\|+\frac{4L_b^x L_b^w\phi_w}{L_b^x+1}e^{(L_b^x+1)(t-t_{k-D-T_m})}\|S(t)\|\|\epsilon(t)\| \\ & +2\sqrt{n}\|\epsilon(t)\|\|\epsilon(t_{k-D-T_m})\|\|S(t)\|L_b^x e^{(L_b^x+1)(t-t_{k-D-T_m})}\end{aligned} \tag{4.38}$$

where $\delta_{max} = \max_i\{\delta_{max_i}\}$, and $\delta_{min} = \min_i\{\delta_{min_i}\}$, $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$, $R_{min} = \min\{R_i\}$. According to the definition of Lyapunov function V and the boundedness of S such that

$\delta_{min} \leq \|S\| \leq \delta_{max}$, we obtain $\delta_{min}\|\epsilon(t)\|^2 \leq V(t) \leq \delta_{max}\|\epsilon(t)\|^2$ and based on (4.38) we get

$$\begin{aligned}\frac{dV(t)}{dt} \leq & -\min\{\theta_i q_{m_i}\delta_{min_i}\}\delta_{min}\|\epsilon(t)\|^2 + 2\|\epsilon(t)\|^2\delta_{max}n_y L_{b^*_{max}} + \frac{4\theta_m n_y \phi_v}{R_{min}}\|\epsilon(t)\| \\ & + 2\sqrt{n}\|\epsilon(t)\|\|\epsilon(t_{k-D-T_m})\|\delta_{max}L_b^x e^{(L_b^x+1)(t-t_{k-D-T_m})} + 2(L_b^w\phi_w - \frac{L_b^w L_b^x \phi_w}{L_b^x+1})\|\epsilon(t)\|\delta_{max} \\ & + \frac{4L_b^x L_b^w \phi_w}{L_b^x+1} e^{(L_b^x+1)(t-t_{k-D-T_m})}\delta_{max}\|\epsilon(t)\|\end{aligned} \tag{4.39}$$

As stated in Theorem 5, from condition (4.35) and the definitions of d_m and ϵ^*, we can verify that $\|\epsilon(t_{k-D-T_m})\| \leq d_m$,$\|t - t_{k-D-T_m}\| < (T_m + D)\Delta$ for $t \in [t_{k-D-T_m}, t_k)$ and

$$\frac{dV}{dt} < 0 \tag{4.40}$$

for all $\|\epsilon(t)\| > \epsilon^*$ and $t \in [t_{k-D-T_m}, t_k)$, which means that the overall Lyapunov function is decreasing. If we consider the distributed filters communicating with each other every $(D + T_m)\Delta$ based on worst-case scenario, ϵ is reset to be the same as e every sampling time, and using the above result recursively, it is proved that $\frac{dV}{dt} < 0$ for all time as long as $\|\epsilon(t)\| > \epsilon^*$. This implies that within finite time duration, the estimation error $\|\epsilon(t)\|$ will be reduced to be smaller than ϵ^*. Once $\|\epsilon(t)\| < \epsilon^*$, the time derivative of V may be positive and $\|\epsilon(t)\|$ may increase; however $\|\epsilon(t)\|$ will be always smaller than ϵ^* because $\frac{dV}{dt} < 0$ when $\|\epsilon(t)\| > \epsilon^*$. This proves the theorem 5. □

4.7 Application to a simulated chemical process

In this section, we apply the proposed distributed AHG-EKF to a simulated chemical process allowed for data losses and delays. This process is a 2-CSTR system which was described in Section 2.6. Constant heat inputs to the two reactors are used: $Q_{h_1} = 1.4 \times 10^4\ kJ/h$ and $Q_{h_2} = 1.4 \times 10^4\ kJ/h$. These inputs ensure the stability of the process.

The two temperatures T_1 and T_2 are assumed to be the continuously measured outputs.

Based on these measurements, the estimation of the entire system state is desired. The whole process is divided into two subsystems with respect to CSTR1 and CSTR2, and an AHG-EKF is designed for each subsystem. In the design of the AHG-EKFs, $Q_1 = Q_2 = \text{diag}([50, 5])$ and $R_1 = R_2 = 1$, respectively. In the simulations, the communication interval is 36 seconds, the maximum time delay is picked as 108 seconds and the maximum allowable consecutive data loss samples is 72 seconds (i.e. $\Delta = 10$, $D = 3$ and $T_m = 2$). For all the subsystem estimators the parameter $\theta_m = 10$. Also, a unique value is used for the calculation of the innovation terms, $d = 0.01h$. The parameters in the adaptation functions are selected as follows: $\Delta T = 0.001h$, $\beta_1 = 5$, $m_1 = 1$, $\lambda_1 = 1000$, $\beta_2 = 5$, $m_2 = 1$, $\lambda_2 = 1000$. In the simulations, the initial state of the process is $x(0) = [360, 3, 320, 3]^T$ and the initial guesses in the two distributed filters are $z_1(0) = [396,\ 2.1]^T$ and $z_2(0) = [352,\ 9.7]^T$. Also the uncertainty in the temperature measurement is considered as $1°K$ in each subsystem, the process noise for temperature and concentration dynamics are considered as $3°K/sec$ and 60 $mol/m^3/sec$, respectively. In order to have excited concentration trajectories in the simulations, a periodic F_0 is used; that is, $F_0 = 10(1 + 0.9\sin(0.008t))m^3/h$.

First, we compare the proposed design with regular distributed EKF design (i.e., the proposed design with $\theta = 1$ all the time). In both of the designs, the two distributed filters communicate to transmit information as in the proposed design. The communication between the two filters is subject to delays and losses. The sequences of the data loss and communication delays among the filters in the simulations are shown in Figure 4.3. Note that the predictors are used in both designs to predict the missing information. Figure 4.4 shows the simulation results given by the proposed DAHG-EKF and the regular distributed EKF. It can be seen from the figure that the system states can be tracked in both of the designs. However, the proposed DAHG-EKF converges to the actual system states much faster than the regular distributed EKF. This result is also reflected by the innovations. The innovations of the two subsystems under the proposed DAHG-EKF decrease much faster to values close to zero compared with the innovations under the regular EKF. Moreover, it

can be seen that the gain of each subsystem filter in the proposed design increases to a high value quickly when the initial innovation is large and then decreases to 1 when the innovation becomes small. The adaptive gains in the proposed design ensure that the estimates are not sensitive to measurement noise which is demonstrated in the next set of simulations. The ability of the distributed filters to track more challenging dynamics is illustrated in Figure 4.5. In order to provide more visible variations in the states, the heat inputs to the tanks are assumed constant and the input flow-rate F_0 to the first tank varies with a sinusoidal trend. Figure 4.5 shows that although sinusoidal input on the first subsystem results in more variations on the actual states, the local filters can follow them effectively.

Next, we show the effectiveness of adaptive gains in reducing the sensitivity of the estimates to measurement noise. Specifically, in this set of simulations we use constant high gains in the two distributed EKF. Figure 4.6 shows the simulation results when the gains remain high all the time. From Figure 4.6, we see that the concentration estimates are very noisy with high fluctuations. This is expected when high gains are used in the filter designs. If comparing the results in Figure 4.6 with the results of the proposed distributed AHG-EKF shown in Figure 4.4, we can conclude that the proposed distributed AHG-EKF maintains the fast convergence of the estimates while significantly reducing the sensitivity of the estimates to measurement noise.

Then, we evaluate the importance of state predictors in the proposed design. Figure 4.7 shows the simulation results of the proposed design and a design without the predictors. Note that when there is no predictor, the latest available information is used to approximate the interactions between subsystems. Figure 4.7 shows that the estimates of the proposed design with predictors converge much faster to the actual system state compared with the case without predictors. From Figure 4.7, it can be inferred that the use of state predictors significantly improves the estimation performance.

Finally, the performance of the proposed approach is evaluated under the consideration of model-process mismatch in the design of predictors. In this set of simulations, the model

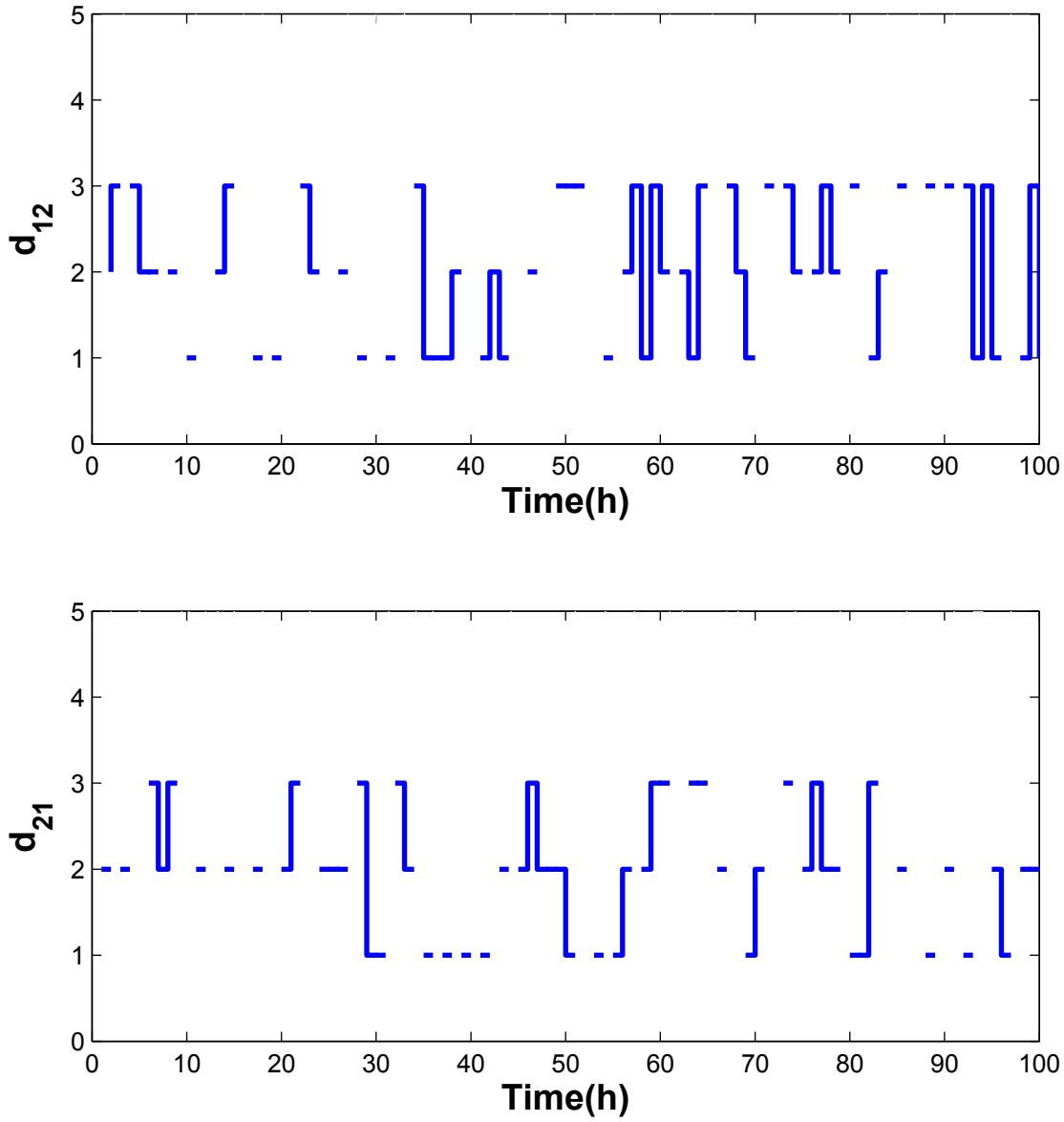

Figure 4.3: Communication delay and data loss sequnces between the subsystems with $D = 3$ and $T_m = 2$. The delay and loss of information of subsystem j to subsystem i is denoted by d_{ij}. The delay is indicated by solid lines and the discontinuity of the lines represents data losses.

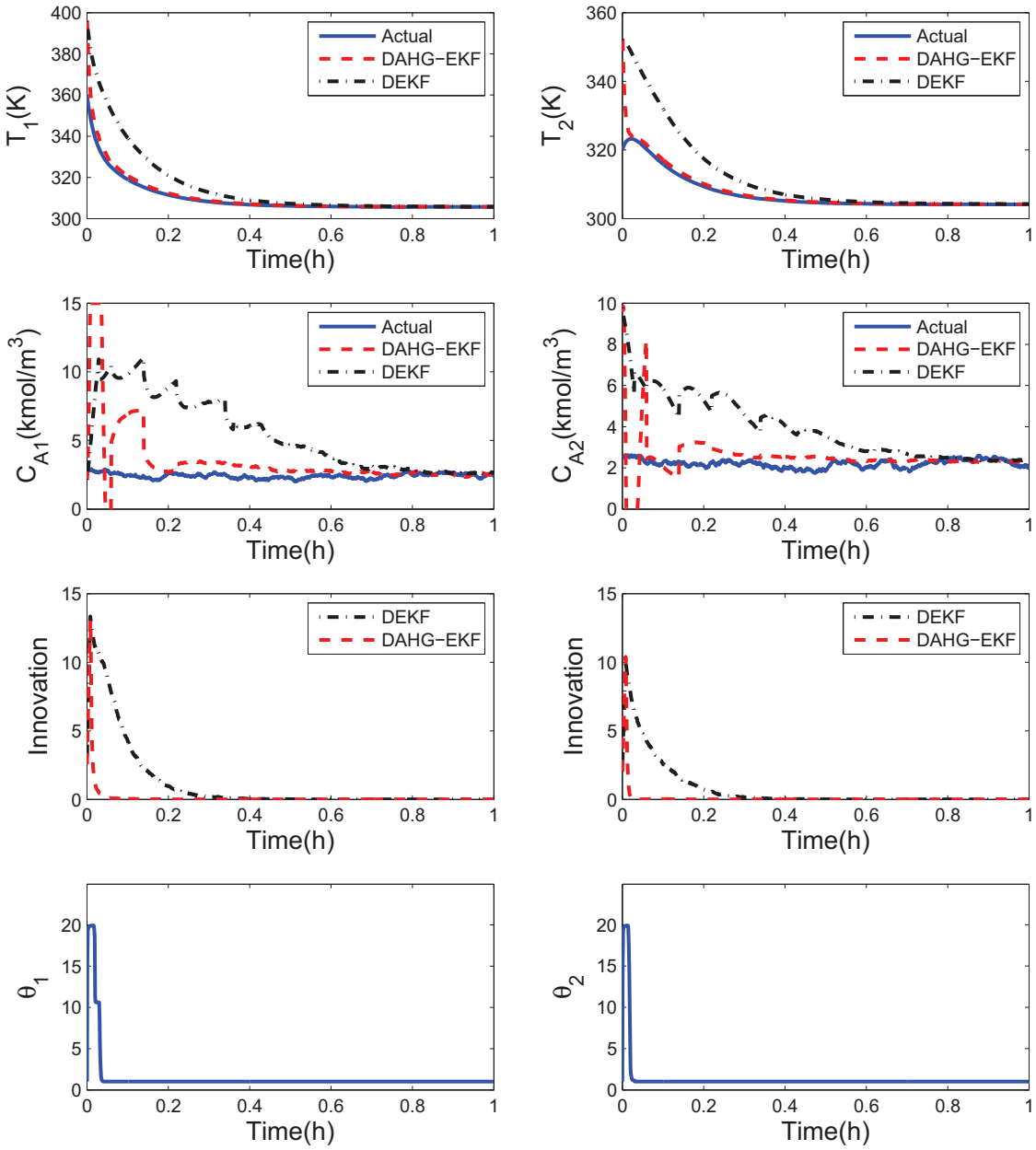

Figure 4.4: Trajectory of the states and their estimates, and trajectories of θ and the corresponding innovation under the proposed distributed AHG-EKF and the regular distributed EKF with communication delay and data loss.

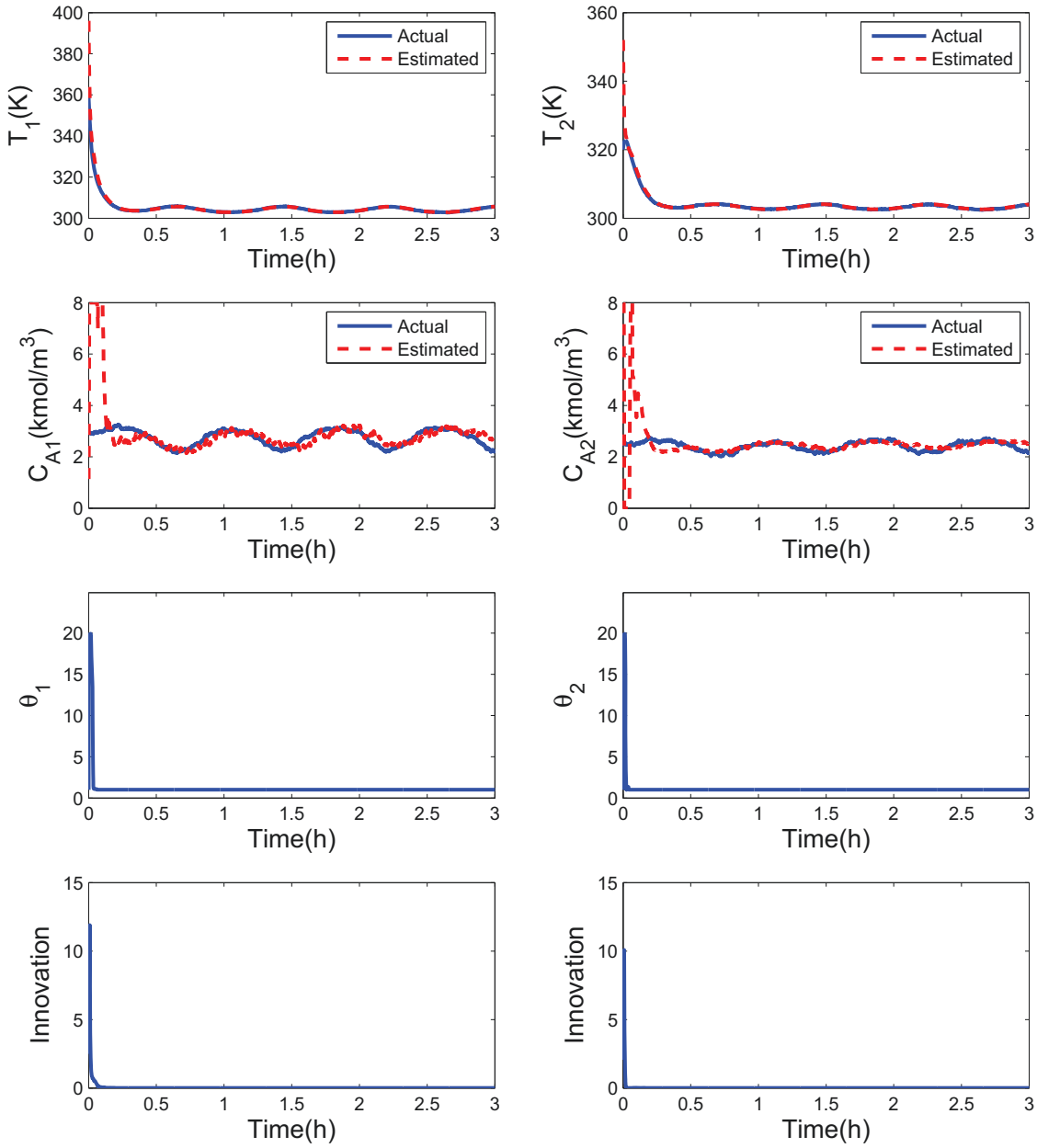

Figure 4.5: Trajectory of the states and their estimates, and trajectories of θ and the corresponding innovation under the proposed distributed AHG-EKF with sinusoidal flow input F_0, communication delay and data loss.

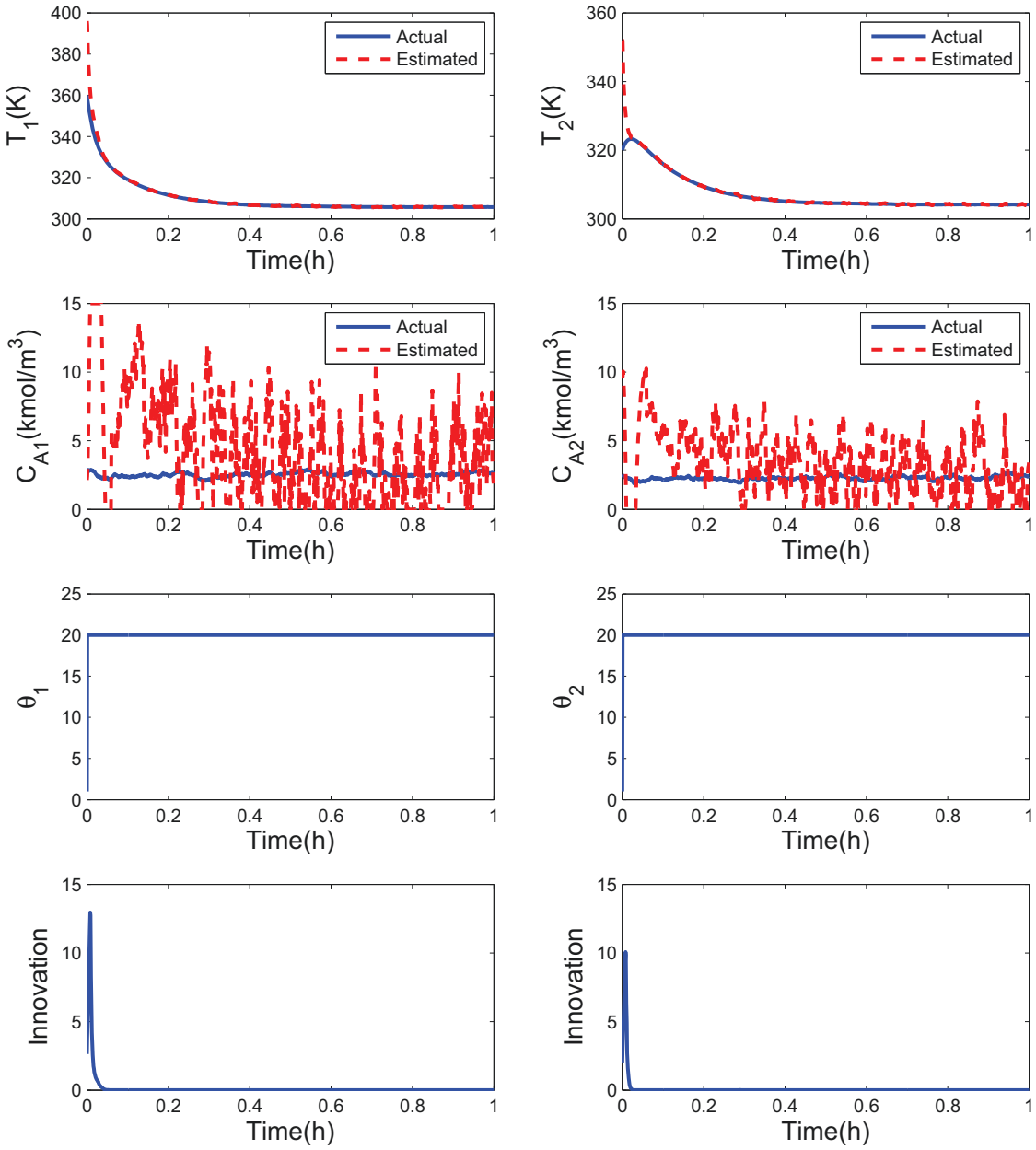

Figure 4.6: Trajectory of the states and their estimates, and trajectories of θ and the corresponding innovation under distributed high-gain EKF without gain adaptation.

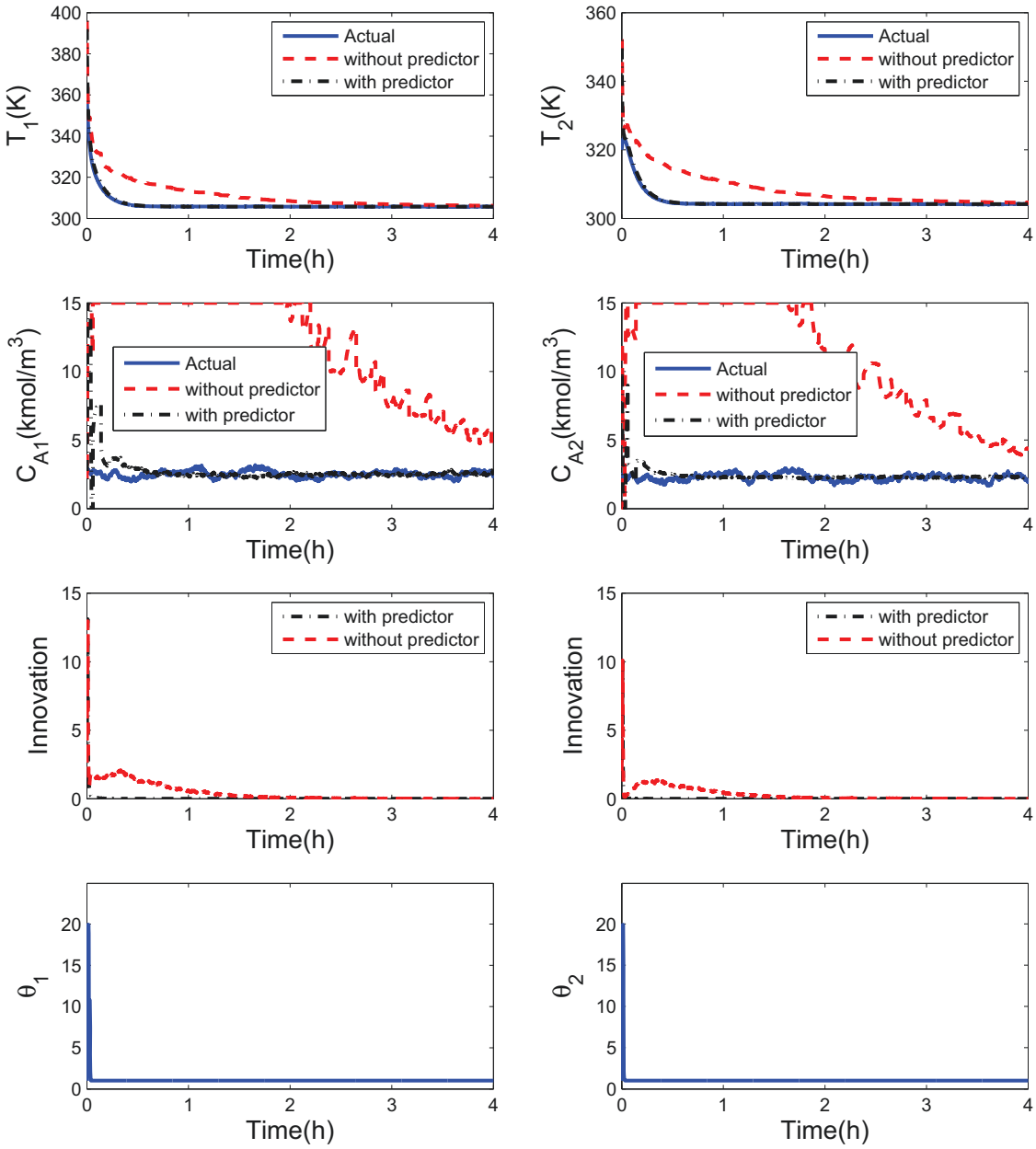

Figure 4.7: Trajectory of the states and their estimates, and trajectories of θ and the corresponding innovation under the proposed distributed AHG-EKF with and without state predictors.

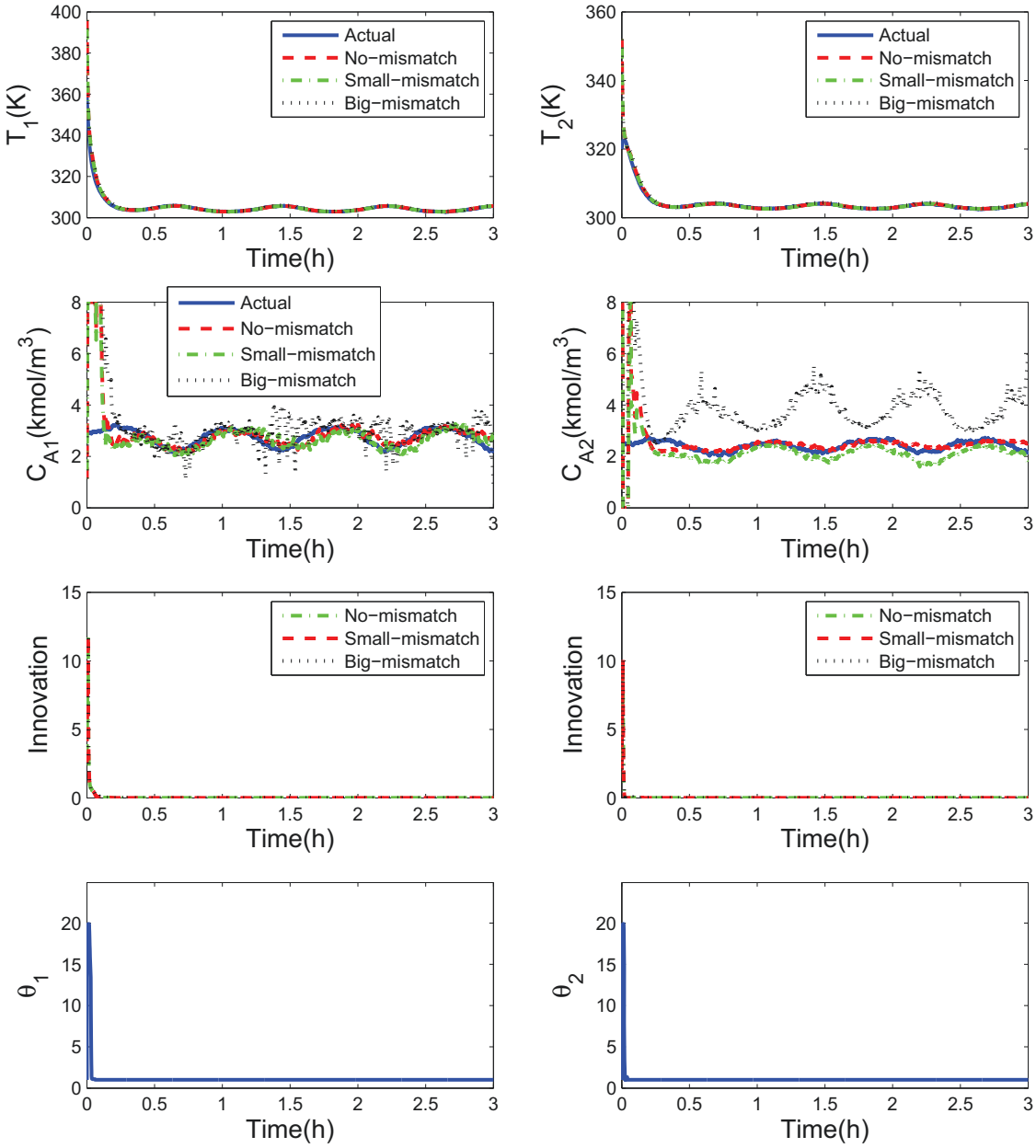

Figure 4.8: Trajectory of the states and their estimates, and trajectories of θ and the corresponding innovation under the proposed distributed AHG-EKF with model mismatch in predictors.

mismatch is considered as an uncertainty in the recycle flow-rate (F_r) to the first CSTR. Figure 4.8 shows the results of a small mismatch (i.e., the recycle flow rate in the predictor is $F_r + 1m^3/h$) and a large mismatch (i.e., the recycle flow rate in the predictor is $F_r + 5m^3/h$). It is seen that when the mismatch is small, the distributed filters still work well due to the inherit robustness of the design. However, when the mismatch is too large, the use of the predictor may not improve the performance of the distributed filters.

4.8 Conclusions

In this chapter, we demonstrated that the distributed adaptive high-gain extended Kalman filtering is a practical strategy to handle communication delays and data losses for a class of nonlinear systems made of several interacting subsystems. In this approach, a local adaptive high-gain extended Kalman filter is designed for each subsystem. The subsystem estimators are capable of exchanging information to effectively compensate for the interactions via their communications. In the proposed design, a prediction of the entire system state is generated for each subsystem estimator to provide a contribution for handling communication delays and data losses. Sufficient conditions were established by a rigorous analysis under which the proposed distributed estimation approach provides a bounded overall estimation error. Finally, a simulated chemical process is used to evaluate the practical performance of the proposed approach.

Chapter 5

Distributed adaptive high-gain extended Kalman filters with triggered communication

5.1 Introduction

In this chapter, a DAHGEKF scheme with triggered communication is designed for nonlinear systems. Particularly, a class of continuous-time nonlinear systems composing of several interconnecting subsystems is considered. In the proposed approach, an adaptive-gain EKF is first designed for each subsystem together with a trigger which determines when the information transmission is required. Each filter sends information through the network whenever its corresponding triggering condition is satisfied. Second, the implementation algorithm of the proposed design in distributed framework is discussed. This algorithm describes when the distributed filters communicate and what information should be transmitted. Conditions are specified under which the stability of the DAHGEKF with triggering communication is guaranteed. A simulated chemical process is used to demonstrate the applicability and effectiveness of the proposed design.

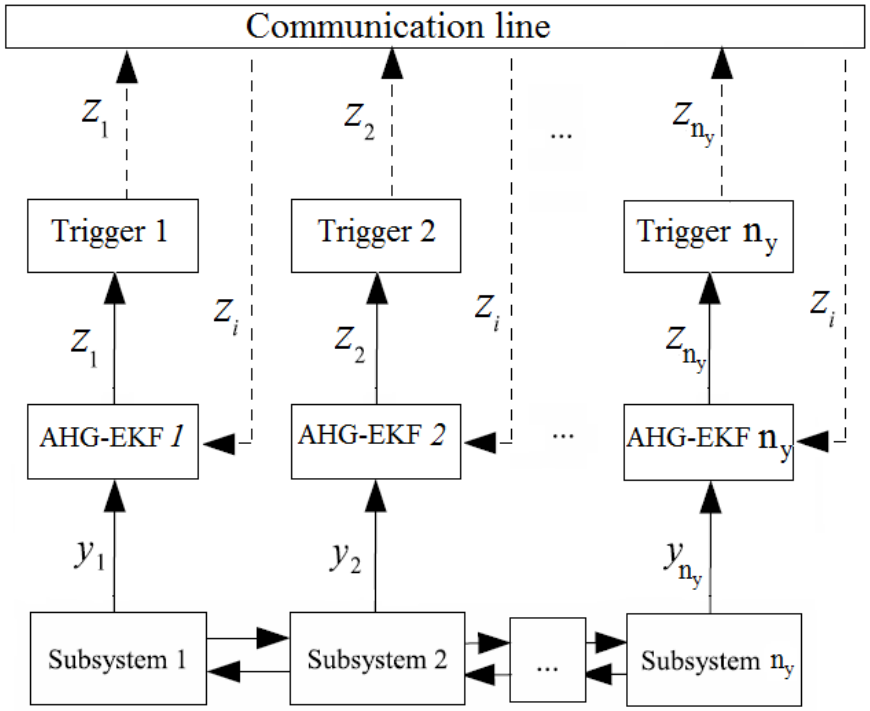

Figure 5.1: Proposed distributed state estimation design with triggered information transmission.

5.2 Distributed estimation scheme with triggered communication

In this section, we describe the proposed distributed adaptive high-gain extended Kalman filtering (DAHG-EKF) design with triggered communication. Figure 5.1 illustrates the structure of the proposed design. In the proposed design, each estimator has an AHG-EKF and a transmission trigger which determines if the information of the local estimator should be sent out to the other estimators at a sampling time. Based on the triggered strategy, the frequency of information transmission between the estimators will be reduced and the ultimate boundedness of the estimation error demands the local estimators in previous chapters to be redesigned to take the lack of state updates between estimators into consideration. The reduction of communication frequency is achieved by triggering the transmission of information among filters. In order for better explanation of the methodology, we may consider

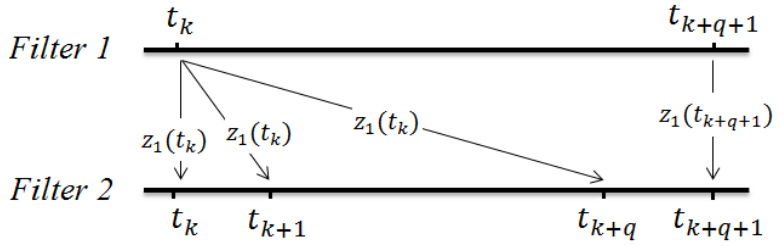

Figure 5.2: An illustration of the triggered communication.

two distributed filters communicating each other through their states as shown in Figure 5.2. A trigger is designed for each filter which sends out information whenever its triggering condition is satisfied. As shown in Figure 5.2, the triggering condition of trigger 1 is satisfied at discrete time instant t_k, and filter 1 transmits information to filter 2. Then, the triggering condition of filter 1 is reset. In Figure 5.2, the triggering condition of trigger 1 is not satisfied again for the next q sampling times and filter 2 uses the information received at time t_k (i.e. $z_1(t_k)$) within the time period from t_k to t_{k+q}. Filter 1 sends information to filter 2 again at t_{k+q+1} when the triggering condition of trigger 1 is satisfied.

This study is aimed to incorporate the DAHG-EKF scheme in [39, 62] to develop an approach accounting for the frequency reduction in information transmission between the estimators. Based on the difference between the current state estimate and the last sent state estimate, the triggering condition is designed for each trigger.

5.3 Implementation algorithm

In this work, we assume that the output measurements of the subsystems are available for the corresponding estimators continuously. The triggers of subsystems are checked at discrete time instants $\{t_k \geq 0\}$ where $t_k = t_0 + k\Delta$, t_0 is the initial time, Δ is the constant time interval (a positive constant) and k is a positive integer. In order for the initialization at the initial time instant (i.e., $t = 0$), initial subsystem state guesses of all the subsystems

(i.e., $z_j(0)$ with $j \in \mathbb{I}$), the actual subsystem output measurement (i.e., $y_i(0)$), and the initial value of its adaptive gain (i.e., $\theta_i(0) = 1$) are required to be available for estimator i. After this step, the following steps are carried out:

1. At time instant $t_k > 0$

   (a) AHGEKF i has direct and immediate access to the local output measurement $y_i(t_k)$.

   (b) Based on the local measurements $y_i(t)$ and the latest information received from other estimators $z_l(t_q^l)$ for $l \in \mathbb{I}_i$, with t_q^l as the last time instant that AHGEKF i receives information from AHGEKF l, AHGEKF i calculates the current state estimate $z_i(t)$.

   (c) Trigger i receives the corresponding state estimate $z_i(t_k)$ and checks the triggering condition. If the condition is satisfied, the trigger updates the transmission time $t_q^i = t_k$ and sends $z_i(t_k)$ to estimators that need the states of filter i; otherwise, no information will be sent from subsystem i, and other filters continue to use the last updates from estimator i, i.e. $z_i(t_q^i)$.

2. Filters continue evaluating state estimates based on the the latest available information and go to step 1 at the next sampling time t_{k+1}.

5.4 Design of communication triggers

The triggering condition in each trigger is determined based on the difference between the state estimates of its corresponding AHGEKF when it sent information to other estimators. Particularly at time instant t_k, the triggering condition for trigger i is designed as follows:

$$W_i(t_k) = \begin{cases} 1 & if \quad \|z_i(t_k) - z_i(t_q^i)\| \geq L_{z_i} \\ 0 & if \quad \|z_i(t_k) - z_i(t_q^i)\| < L_{z_i} \end{cases} \tag{5.1}$$

where $z_i(t_k)$ is the current state estimate of AHGEKF i, $z_i(t_q^i)$ is the last transmitted state estimate of AHGEKF i, t_q^i is the last time instant when AHGEKF i sent information to other AHGEKFs and L_{z_i} is a predetermined threshold for trigger i. The triggering condition is not satisfied when $W_i(t_k) = 0$ and no information will be sent out from AHGEKF i to other AHGEKFs and they continue to use $z_i(t_q^i)$. On the other hand, the triggering condition will be satisfied when $W_i(t_k) = 1$ and AHGEKF i updates $t_q^i = t_k$ and sends out $z_i(t_k)$. Note that the triggering condition of each estimator is only dependent on the states of the corresponding subsystem, and as a result, each estimator may send out information at different time instants.

5.5 Design of subsystem filters

For each subsystem, the local estimators are designed based on the centralized AHG-EKF presented in [52] and [44], and the appropriate modifications are applied accounting for the interactions between subsystems. The state estimator for subsystem i ($i \in \mathbb{I}$) is designed as follows:

$$\dot{z}_i = A_i z_i + b_i(z^i, 0) - S_i^{-1} C_i^T R_{\theta_i}^{-1}(C_i z_i - y_i) \tag{5.2}$$

where z_i is the state of the filter i, $z^i(t) = [z_1(t_q^1), \ldots, z_{i-1}(t_q^{i-1}), z_i(t), z_{i+1}(t_q^{i+1}), \ldots, z_{n_y}(t_q^p)]$ is composed of the estimated states of distributed filters at different time instants, $R_{\theta_i} = \dfrac{1}{\theta_i} R_i$ with R_i being a positive scalar and θ_i being the adaptive gain parameter of filter i. The vector $z^i(t)$ includes the current estimate of the corresponding filter, $z_i(t)$, since each filter is assumed to have immediate and direct access to its corresponding subsystem's estimates. It should be noted that t_q^i is the last time instant when AHGEKF i sent information to other AHGEKFs, and will be explained further in the following subsections. S_i in (5.2) is determined by solving the following matrix Riccati equation:

$$\dot{S}_i = -\left(A_i + b_i^*(z^i)\right)^T S_i - S_i \left(A_i + b_i^*(z^i)\right) + C_i^T R_{\theta_i}^{-1} C_i - S_i Q_{\theta_i} S_i \tag{5.3}$$

where $b_i^*(z^i)$ denotes the Jacobian of $b_i(z^i, 0)$ with respect to z_i (i.e. $b_i^*(z^i) = \frac{\partial b_i(z^i,0)}{\partial z_i}$) and $Q_{\theta_i} = \theta_i Q_i$ with Q_i being a n_{x_i} by n_{x_i} symmetric positive definite matrix. The tuning strategy for the high-gain parameter θ_i will follow the one described in Section 3.5. Note that the variable J_i in the adaptation function in this design is defined as follows:

$$\begin{aligned} J(t) &= \sum_{i=1}^{n_y} J_i(t) \\ J_i(t) &= \int_{t-d}^{t} \|y_i(t-d, x(t-d), \tau) - y_i(t-d, z_i(t-d), \tau)\|^2 d\tau \end{aligned} \tag{5.4}$$

where $y_i(t-d, z(t-d), \tau)$ for each filter is calculated based on the last estimates received from the other filters, since no predictor is designed for the distributed filters.

5.6 Stability analysis

In this section, the stability of the proposed distributed AHG-EKF design with triggered information transmission will be evaluated. The estimation error for each subsystem is defined as $\epsilon_i = z_i - x_i$, $i \in \mathbb{I}$. Based on (5.2)-(5.3), the error dynamics for each subsystem can be obtained as

$$\dot{\epsilon}_i = \dot{z}_i - \dot{x}_i = \left(A_i - S_i^{-1} C_i^T R_{\theta_i}^{-1} C_i\right) \epsilon_i + b_i(z^i, 0) - b_i(x, w) \tag{5.5}$$

Let us pick $\epsilon_i^T S_i \epsilon_i$ as the Lyapunov function for subsystem i, $i \in \mathbb{I}$. Based on (5.3) and (5.5) it can be obtained that:

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} = \theta_i \left[-\epsilon_i^T S_i Q_i S_i \epsilon_i - \epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i + \frac{2}{\theta_i} \Big(b_i^T(z^i, 0) S_i \epsilon_i - b_i^T(x, w) S_i \epsilon_i - \epsilon_i^T S_i b_i^*(z^i) \epsilon_i \Big) \right] \tag{5.6}$$

The significant role of (5.6) in establishing the convergence of the proposed distributed state estimation with triggered communication will be shown in the following theorem.

Theorem 6. *Consider system (3.3) with subsystems described by (3.1) with the local filters*

designed following (5.2)-(5.4). If the following conditions are satisfied:

1. *nonlinear function $b_i(.,.)$ is locally Lipschitz with respect to its arguments,*
2. *the Jacobian $b_i^*(z^i)$ is bounded,*
3. *the local Riccati matrices S_i with $i \in \mathbb{I}$ are bounded (Theorem 2.18 of [60] are satisfied),*
4. *$\|\dot{z}_i(t)\| \leq M_i$ for any $t \geq 0$,*
5. *for certain q_{m_i} the positive definite matrix Q_i is chosen such that $Q_i \geq q_{m_i} I$ and*

$$q_{m_{min}} > 2n_y^2 L_{b_{max}}^x \frac{\alpha_{max}}{\alpha_{min}^2} + 2n_y L_{b_{max}^*} \frac{\alpha_{max}}{\alpha_{min}^2} \tag{5.7}$$

where α_{min} and α_{max} are the minimum and maximum values of the lower and upper bounds of all S_i, and $q_{m_{min}} = \min_i\{q_{m_i}\}$, $L_{b_{max}}^x = \max_i\{L_{b_i}\}$, $L_{b_{max}^*} = \max_i\{L_{b_i^*}\}$ with L_{b_i} and $L_{b_i^*}$ being the Lipschitz constant for b_i and b_i^*, respectively,

then for any initial condition of subsystems and filters (i.e. $x_i(t_0) \in \mathbb{X}_i$ and $z_i(t_0) \in \mathbb{X}_i$) for all $i \in I$ in $\mathbb{X}_i$, the norm of the whole system's estimation error $\|\epsilon(t)\|$ is descending and ultimately smaller than ϵ_t; or

$$\lim_{t\to\infty} \|\epsilon(t)\| \leq \epsilon_t \tag{5.8}$$

where

$$\epsilon_t = \frac{2n_y(n_y-1)L_{b_{max}}^x(L_{z_{max}} + M_{max}\Delta)\alpha_{max} + 2\alpha_{max}n_y L_b^w \phi_w + 4\theta_m n_y \frac{\phi_v}{R_{min}} + \delta}{\alpha_{\min}^2 \min_i\{\theta_i q_{m_i}\} - 2n_y^2 L_{b_{max}}^x \alpha_{max} - 2n_y L_{b_{max}^*}\alpha_{max}} \tag{5.9}$$

with $R_{min} = \min_i\{R_i\}$ and δ a small positive constant.

Proof: For a positive definite R_i, it can be verified that $\epsilon_i^T C_i^T R_i^{-1} C_i \epsilon_i \geq 0$. Also, since Q_i is a positive definite matrix, there exists a q_{m_i} such that $Q_i \geq q_{m_i} I$. By using (5.6) we can obtain that

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \epsilon_i^T S_i^2 \epsilon_i + 2\epsilon_i^T S_i \left(b_i(z^i, 0) - b_i(x, w) - b_i^*(z^i)\epsilon_i\right) + 2\theta_i \epsilon_i^T C_i^T R_i v_i \tag{5.10}$$

According to Theorem 2.18 in [60], there exist scalars $\alpha_{min_i} > 0$ and $\alpha_{max_i} > 0$ for AHGEKF i such that the local Riccati matrix is bounded as $\alpha_{min_i} I \leq S_i(t) \leq \alpha_{max_i} I$. So, from (5.10) it can be obtained that

$$\frac{d\left(\epsilon_i^T S_i \epsilon_i\right)}{dt} \leq -\theta_i q_{m_i} \alpha_{min_i} \epsilon_i^T S_i \epsilon_i + 2\epsilon_i^T S_i \left(b_i(z^i, 0) - b_i(x, w) - b_i^*(z^i)\epsilon_i\right) + 2\theta_i \epsilon_i^T C_i^T R_i v_i \quad (5.11)$$

In order to analyze the stability of the overall system for $t \in [t_k, t_{k+1})$, we add (5.11) for all $i \in \mathbb{I}$ together to obtain,

$$\begin{aligned} \frac{d}{dt} \sum_{i=1}^{n_y} \epsilon_i^T S_i \epsilon_i(t) \leq & -\min_i\{\theta_i q_{m_i} \alpha_{\min_i}\} \epsilon^T S \epsilon(t) + \sum_{i=1}^{n_y} 2\theta_i \epsilon_i^T(t) C_i^T R_i v_i + 2\epsilon^T S(t) \\ \times & \left(\begin{bmatrix} b_1(z^1, 0) - b_1(x, w) \\ b_2(z^2, 0) - b_2(x, w) \\ \vdots \\ b_{n_y}(z^p, 0) - b_{n_y}(x, w) \end{bmatrix} - \begin{bmatrix} b_1^*(z^1) & 0 & \dots & 0 \\ 0 & b_2^*(z^2) & 0 & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & b_{n_y}^*(z^p) \end{bmatrix} \begin{bmatrix} \epsilon_1(t) \\ \epsilon_2(t) \\ \vdots \\ \epsilon_{n_y}(t) \end{bmatrix} \right) \end{aligned} \quad (5.12)$$

where z^i includes the estimation of all interacting states, either received from other filters at time t_q^l (i.e. $z_l(t_q^l)$) or computed at the local filter i at time t (i.e. $z_i(t)$). Note that t_q^l is the last sampling time when the information from filter l is sent to the other communicating filters.

According to the Lipschitz property for the nonlinear terms of subsystem i we have:

$$\|b_i(z^i, 0) - b_i(x, w)\| \leq L_{b_i}^x \|z^i - x\| + L_b^w \phi_w \quad (5.13)$$

Based on the triangular inequality and the design of the triggering conditions, it can be

obtained that

$$\begin{aligned}
\|z^i(t)-x(t)\| &\leq \sum_{l\in\mathbb{I}_i, l\neq i} \|x_l(t)-z_l(t_q^l)\| + \kappa_i\|x_i(t)-z_i(t)\| \\
&\leq \kappa_i\|x_i(t)-z_i(t)\| + \sum_{l\in\mathbb{I}_i, l\neq i} \big(\|z_l(t_k)-z_l(t_q^l)\| + \|x_l(t)-z_l(t_k)\|\big) \\
&\leq \kappa_i\|x_i(t)-z_i(t)\| + \sum_{l\in\mathbb{I}_i, l\neq i} \big(\|z_l(t_k)-z_l(t_q^l)\| + \|z_l(t_k)-z_l(t)\| + \|z_l(t)-x_l(t)\|\big)
\end{aligned} \tag{5.14}$$

where κ_i can be either 0 or 1 representing whether the local state estimates are present in the set of interacting state for subsystem i.

Whenever the estimates of the subsystem l do not change significantly in the consecutive sampling times from t_q^l to t_k, the triggering condition for the estimates of subsystem l is not satisfied and we would have $\|z_l(t_k)-z_l(t_q^l)\| < L_{z_l}$. Since $\|\dot{z}_i(t)\| \leq M_i$, for any $t \in [t_k, t_{k+1})$ we can obtain that

$$\|z_l(t_k)-z_l(t)\| \leq M_l\Delta \tag{5.15}$$

Hence, from (5.14) and (5.15) and the triggering condition ($\|z_l(t_k)-z_l(t_q^l)\| \leq L_{z_l}$) we obtain

$$\|z^i(t)-x(t)\| \leq \kappa_i\|\epsilon_i(t)\| + \sum_{l\in\mathbb{I}_i, l\neq i} (L_{z_l} + M_l\Delta + \|\epsilon_l(t)\|) \leq (n_y-1)(L_{z_{max}} + M_{max}\Delta) + n_y\|\epsilon(t)\| \tag{5.16}$$

On the other hand, using the Lipschitz property we have

$$\left\| \begin{bmatrix} b_1(z^1)-b_1(x,w) \\ b_2(z^2)-b_2(x,w) \\ \vdots \\ b_{n_y}(z^p)-b_{n_y}(x,w) \end{bmatrix} \right\| \leq n_y L_{b_{max}}^x \left((n_y-1)(L_{z_{max}} + M_{max}\Delta) + n_y\|\epsilon(t)\|\right) + n_y L_b^w \phi_w \tag{5.17}$$

Since $\alpha_{min_i} I \leq S_i \leq \alpha_{max_i} I$, from (5.12) and (5.17) it can be obtained that

$$\frac{d}{dt}\sum_{i=1}^{n_y}\epsilon_i^T S_i \epsilon_i(t) \leq -\min_i\{\theta_i q_{m_i}\alpha_{\min_i}\}\epsilon^T S\epsilon(t) + 2n_y^2 L_{b_{max}}^x \frac{\alpha_{max}}{\alpha_{min}}\epsilon^T S\epsilon(t) + 2n_y L_{b^*_{max}}\frac{\alpha_{max}}{\alpha_{min}}\epsilon^T S\epsilon(t)$$
$$+2n_y(n_y-1)L_{b_{max}}^x(L_{z_{max}} + M_{max}\Delta)\|\epsilon(t)\|\|S(t)\| + 2\|\epsilon(t)\|\|S(t)\|n_y L_b^w \phi_w + 4\theta_m \frac{p\phi_v}{R_{min}}\|\epsilon(t)\| \tag{5.18}$$

where $L_{b^*_{max}} = \max_i\{L_{b_i^*}\}$. If

$$\|\epsilon(t)\| > \epsilon_t \tag{5.19}$$

where ϵ_t is defined in (5.9), then by multiplying both sides of (5.19) by $\alpha_{min}\|\epsilon(t)\|$ we obtain that

$$\alpha_{min}\|\epsilon(t)\|^2 > \frac{2n_y(n_y-1)L_{b_{max}}^x(L_{z_{max}} + M_{max}\Delta)\alpha_{max} + 2\alpha_{max}n_y L_b^w\phi_w + 4\theta_m n_y\frac{\phi_v}{R_{min}} + \delta}{\alpha_{\min}\min_i\{\theta_i q_{m_i}\} - 2n_y^2 L_{b_{max}}^x\frac{\alpha_{max}}{\alpha_{min}} - 2n_y L_{b^*_{max}}\frac{\alpha_{max}}{\alpha_{min}}}\|\epsilon(t)\| \tag{5.20}$$

and since $V(t) = \epsilon^T S\epsilon(t) \geq \alpha_{min}\|\epsilon(t)\|^2$, based on (5.7), from (5.20) it can be obtained that

$$\left(-\min_i\{\theta_i q_{m_i}\alpha_{\min_i}\} + 2n_y^2 L_{b_{max}}^x\frac{\alpha_{max}}{\alpha_{min}} + 2n_y L_{b^*_{max}}\frac{\alpha_{max}}{\alpha_{min}}\right)V(t)$$
$$+2\|\epsilon(t)\|\left(n_y(n_y-1)L_{b_{max}}^x(L_{z_{max}} + M_{max}\Delta)\alpha_{max} + \alpha_{max}n_y L_b^w\phi_w + 2\theta_m n_y\frac{\phi_v}{R_{min}} + \frac{\delta}{2}\right) < 0 \tag{5.21}$$

Also, due to the boundedness of the Riccati matrix (i.e. $\|S(t)\| \leq \alpha_{max}$), from (5.21) we obtain that,

$$\left(-\min_i\{\theta_i q_{m_i}\alpha_{\min_i}\} + 2n_y^2 L_{b_{max}}^x\frac{\alpha_{max}}{\alpha_{min}} + 2n_y L_{b^*_{max}}\frac{\alpha_{max}}{\alpha_{min}}\right)V(t) + 2\|\epsilon(t)\|\|S(t)\|n_y L_b^w\phi_w$$
$$+2n_y(n_y-1)L_{b_{max}}^x(L_{z_{max}} + M_{max}\Delta)\|\epsilon(t)\|\|S(t)\| + 4\theta_m\frac{n_y\phi_v}{R_{min}}\|\epsilon(t)\| < -\delta\|\epsilon(t)\| < -\delta\epsilon_t \tag{5.22}$$

By comparing (5.22) with the right hand side of (5.18), it can be obtained that $\dot{V}(t) < -\delta\epsilon_t$ which implies that the estimation error decreases when its norm is greater than ϵ_t. It also implies that the error will become smaller than ϵ_t in finite time. After the error is smaller than ϵ_t, the error may not decrease but the error will be bounded in ϵ_t. This proves the theorem. □

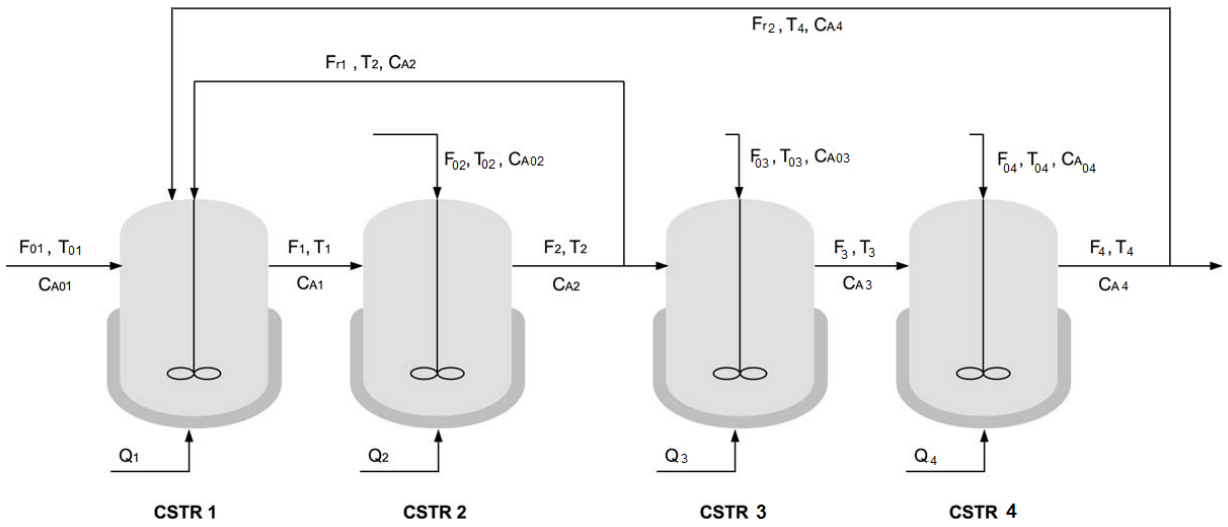

Figure 5.3: Two connected CSTRs with recycle stream.

Remark 8. *Note that due to process and measurement noises and the use of triggered communication, the estimation error will not decrease to zero. The value of ϵ_t which ultimately bounds the estimation error depends on the magnitudes of the process noise, measurement noise and the thresholds used in the communication triggers as can be seen from the definition of ϵ_t in (5.9). However it should be noted that the high gain parameter plays an important role in the value of ϵ_t. When the estimation error is big, the gain increases and causes ϵ_t to decrease which contributes the estimation error to be confined to a smaller region.*

5.7 Application to a chemical process

In this section, the performance of the proposed DAHGEKF and its triggering communication will be evaluated via the application to a chemical process composed of four connected continuous-stirred tank reactors (CSTRs). This process is described in Section 3.7 in details. In the simulations, the heat inputs to the four reactors are selected as constant values: $Q_{h_1} = 1.0 \times 10^4\ kJ/h$, $Q_{h_2} = 2 \times 10^4\ kJ/h$, $Q_{h_3} = 2.5 \times 10^4\ kJ/h$ and $Q_{h_4} = 1.0 \times 10^4\ kJ/h$. The stability of the process is guaranteed under the utilization of these inputs.

It is assumed that the four temperatures T_1, T_2, T_3 and T_4 are the continuously measured

outputs of the process, and it is desired to estimate the entire system state based on these measurements. In this process, each reactor plays the role of one subsystem in the process and an AHGEKF is designed for each CSTR. The covariance parameters of each AHG-EKF are tuned as, $Q_1 = Q_2 = Q_3 = Q_4 = \text{diag}\{[5, 5]\}$ and $R_1 = R_2 = R_3 = R_4 = 1$, respectively. In this set of simulations, unique values are selected for the bound of high-gain parameter, $\theta_m = 20$, and the forgetting horizon in the calculation of innovation terms, $d = 0.03h$. The other parameters included in the adaptation of high-gain parameter are selected as follows: $\Delta T = 0.001h,\ \beta_i = 150,\ m_i = 20,\ \lambda_i = 100$ for $i = 1, \ldots, 4$. The initial states of the subsystems and initial guesses of the corresponding filters are selected as $x_1(0) = [340, 2]^T$, $x_2(0) = [350, 3]^T$, $x_3(0) = [345, 2.5]^T$, $x_4(0) = [360, 4]^T$, $z_1(0) = [340, 2]^T$, $x_2(0) = [350, 3]^T$, $x_3(0) = [345, 2.5]^T$, $x_4(0) = [360, 4]^T$. In the simulations, the temperature measurements are corrupted with noise of zero mean and variance of $1^\circ K$. Also, the triggering condition for each subsystem is evaluated based on the normalized form of (5.1), i.e. $\left\| \left[\frac{\left(T_i(t_k) - T_i(t_q^i)\right)}{T_i(t_q^i)}, \frac{\left(C_{Ai}(t_k) - C_{Ai}(t_q^i)\right)}{C_{Ai}(t_q^i)} \right] H_i \right\|$, with the weighting matrix $H_i = \text{diag}\{[200, 1]\}$ and the triggering thresholds $L_{z_i} = 0.02$ for $i = 1, 2, 3, 4$.

In this set of simulations, the effects of the triggered communication on the proposed distributed state estimation design is studied. In this design, the triggers are checked every $18s$. When the triggering condition is satisfied, the filters send information through the communication network and when the triggering condition is not satisfied, there is no information transmission from the filter. Figure 5.4 shows the sequence of information transmission in the trigger of each subsystem. It can be seen that during the initial periods, each filter sends out information consecutively (i.e., every $18s$) because the subsystem estimates change significantly. When the system estimates change slowly after the initial periods, the filters may not send information at some time instants. The simulation results are illustrated in Figures 5.5 and 5.6. These figures show that the distributed filters track the actual process states, however a steady state error may remain between estimated and actual states due to the use of triggering conditions as discussed in the stability analysis in the previous section.

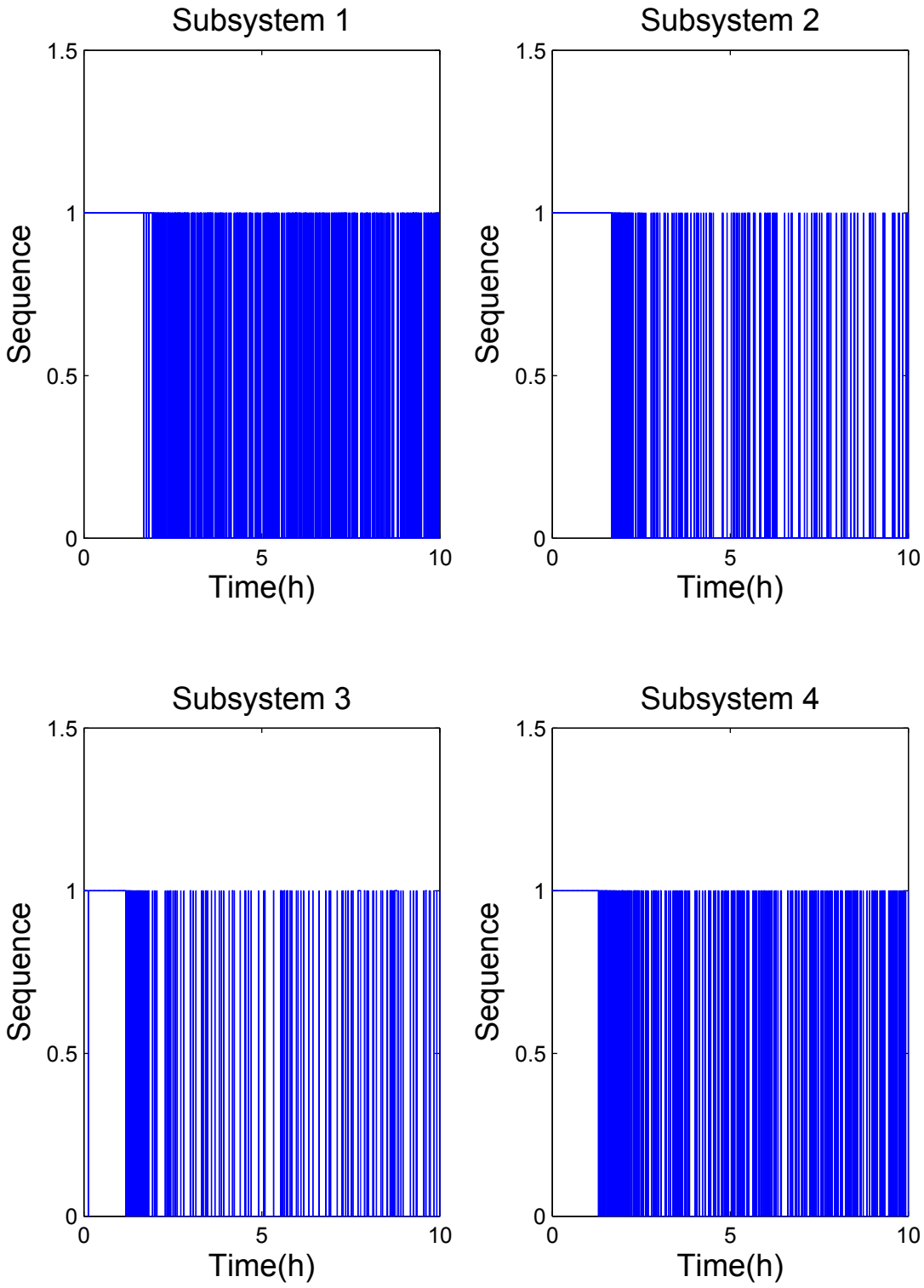

Figure 5.4: The triggering communication sequences

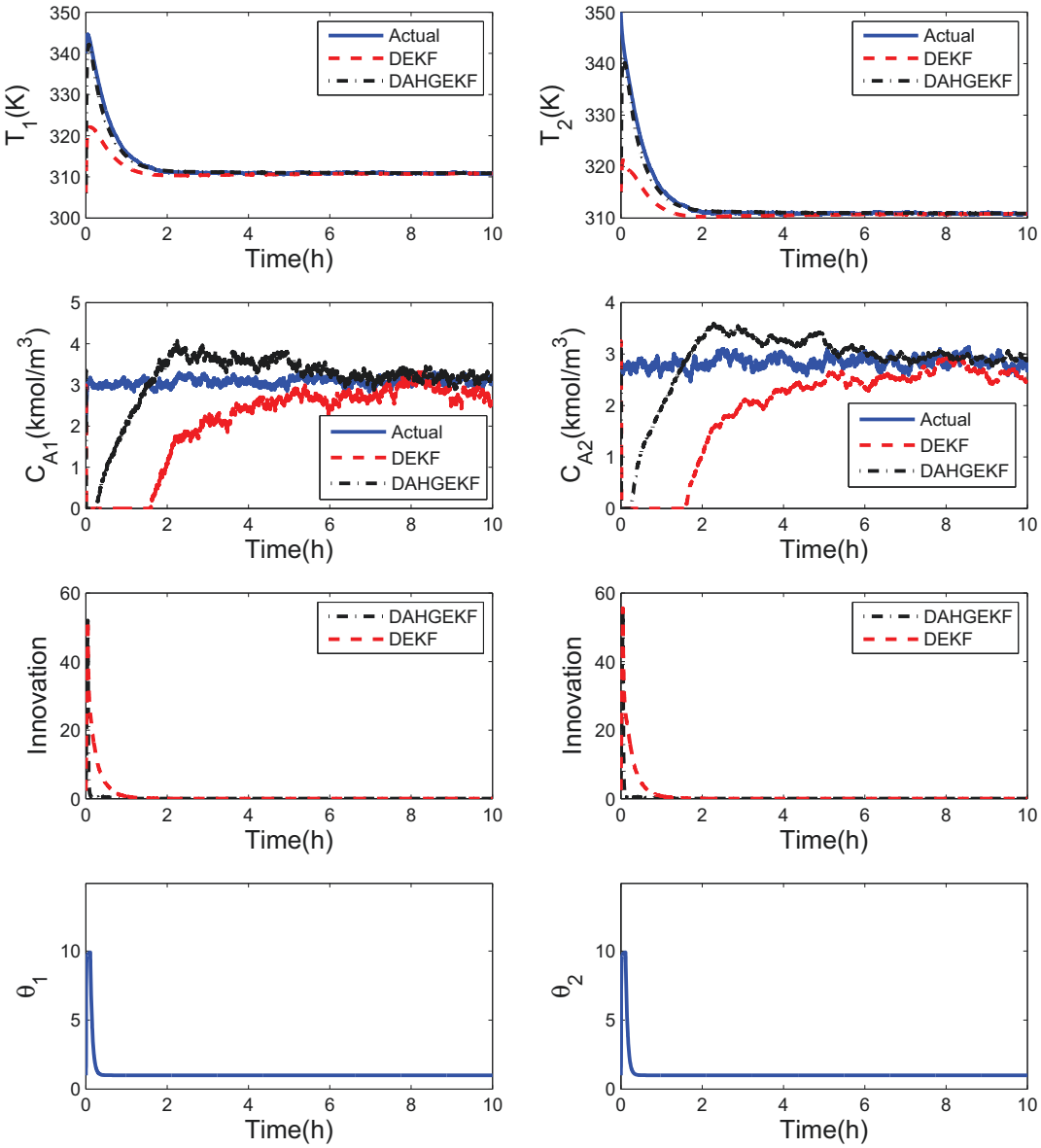

Figure 5.5: The trajectories of the states, high-gain parameters and innovations in the filters of CSTR 1 and CSTR 2 in the presence of measurement noise.

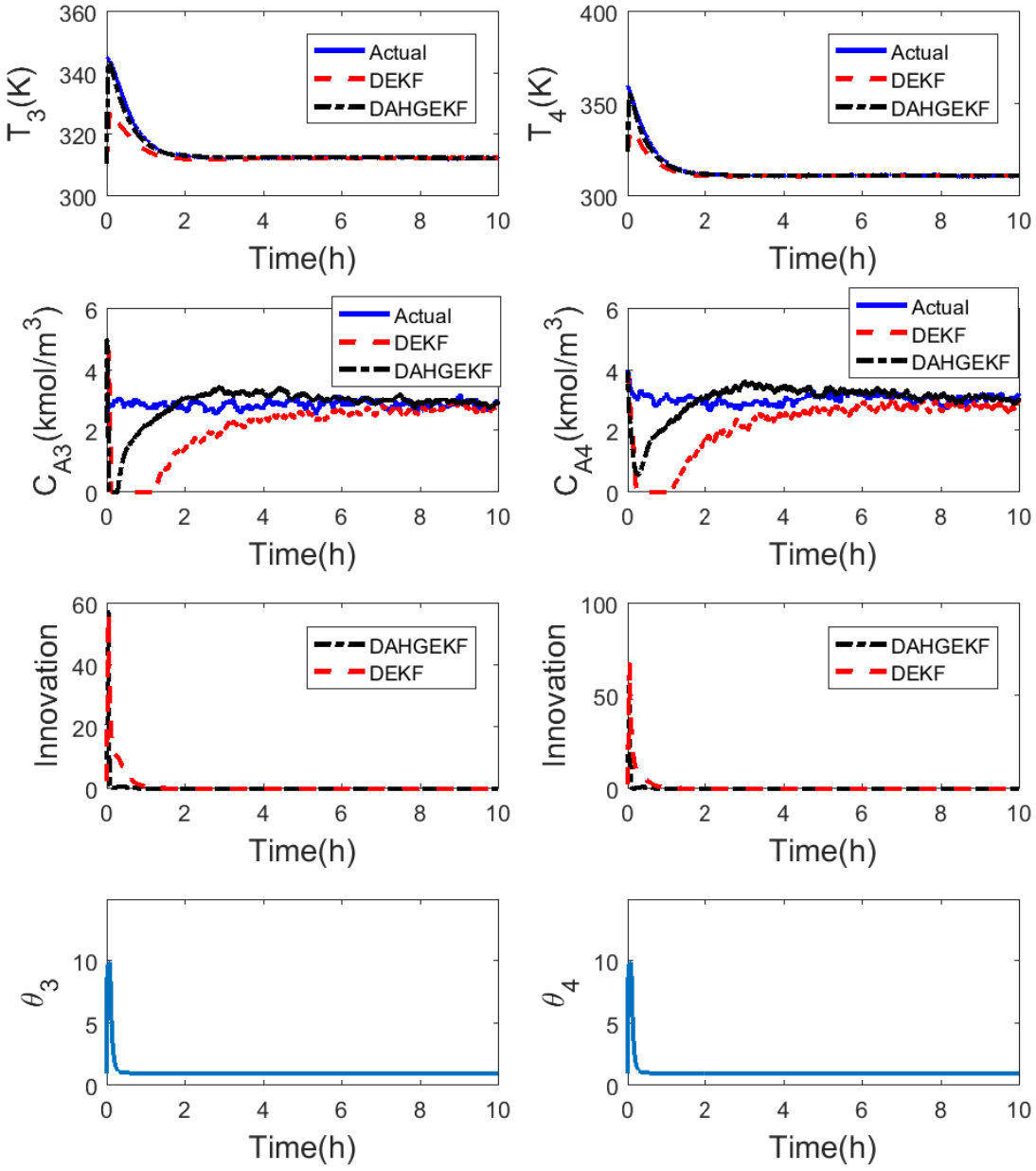

Figure 5.6: The trajectories of the states, high-gain parameters and innovations in the filters of CSTR 3 and CSTR 4 in the presence of measurement noise.

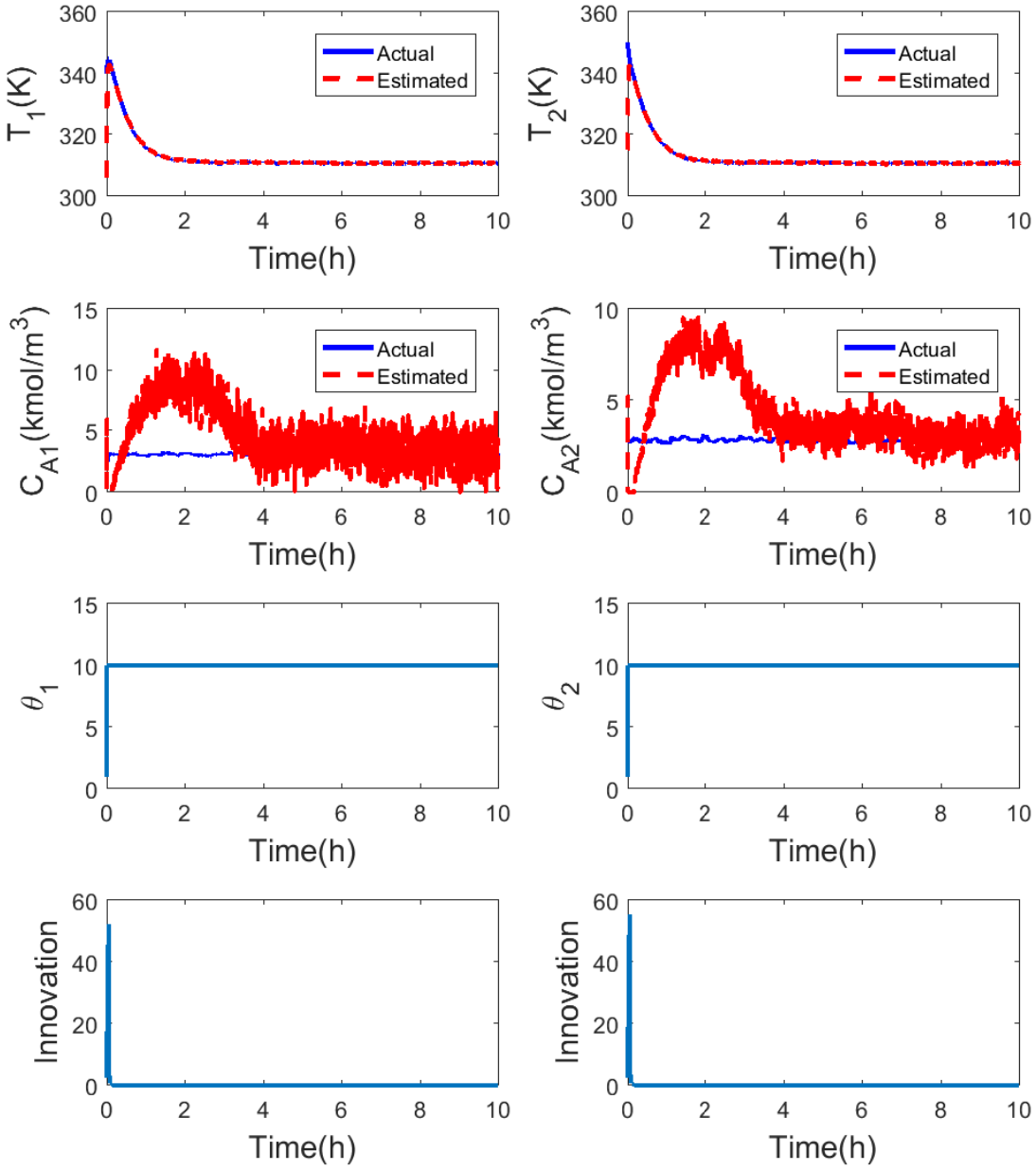

Figure 5.7: The effect of high-gain on the trajectories of the states and innovations in the CSTR 1 and CSTR 2 in the presence of process and measurement noises.

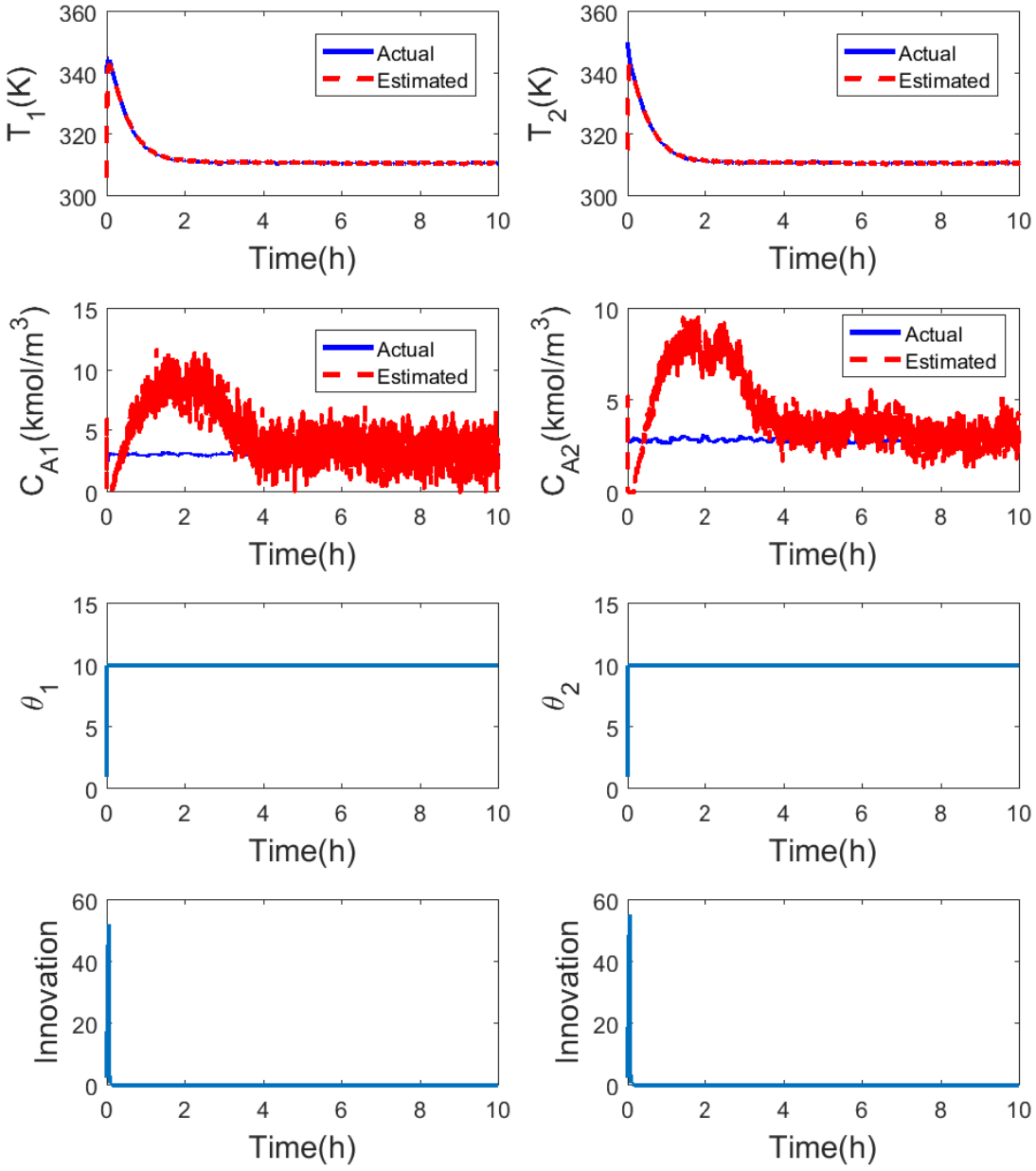

Figure 5.8: The effect of high-gain on the trajectories of the states and innovations in the CSTR 3 and CSTR 4 in the presence of process and measurement noises.

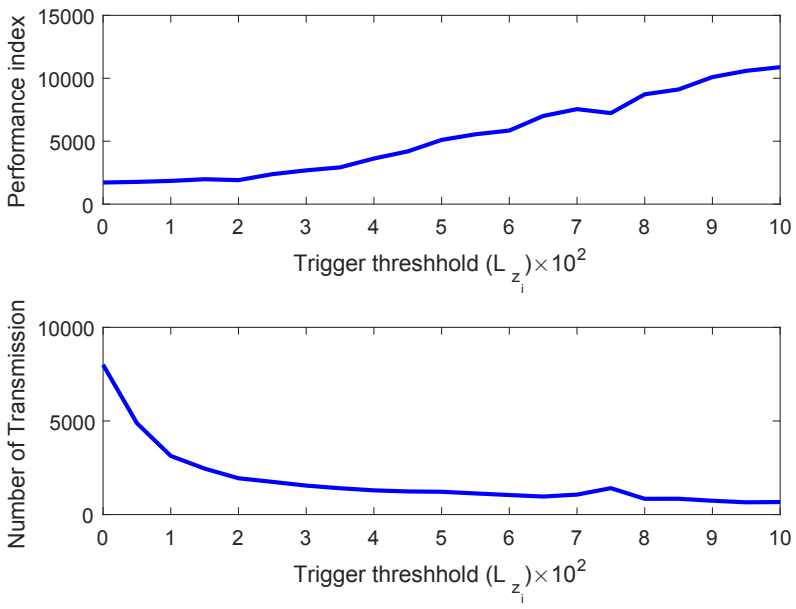

Figure 5.9: Performance index and the number of information transmission in the proposed DAHGEKF based on triggering condition (5.1) with L_{z_i}, $i = 1, 2, 3, 4$, varying from 0 to 0.1.

On the other hand, it can be seen from Figures 5.5 and 5.6 that the distributed adaptive-gain EKF can contribute to improve estimation performance compared with distributed regular EKF in terms of convergence speed and estimation error. Moreover, the estimated values of the concentrations converge the actual states slower compared with the temperatures, since the innovation is defined based on measured temperatures. Based on the definition of innovation in (5.4), the high-gain parameter decreases followed by an initial increment. Furthermore, the presence of process and measurement noises amplifies the uncertainty of the estimation, however the change of high-gain EKF to regular EKF smooths out the effect of noise. If the filters remain in the high-gain mode, they are not able to reduce the noise effect as illustrated in Figures 5.7 and 5.8.

Another set of simulations is also provided to evaluate the performance of the distributed estimation with triggered communication with varying triggering thresholds. Particularly, the simulations are conducted under random noise sequences and the performance index for AHGEKF i is designed as $U_i = \sum_{k=0}^{M} |z_i(t_k) - x_i(t_k)|^2_{G_i}$ where $i = 1, 2, 3, 4$, and the simulation

is run from $t_0 = 0$ to $t_M = 10h$. The parameters G_i, $i = 1, 2, 3, 4$ are used to compensate for the different scales of the states, and $G_i = \text{diag}\{[1, 200]\}$. The overall performance is calculated as $U = U_1 + U_2 + U_3 + U_4$. As shown in Figure 5.9, it can be inferred that the number of information transmission among filters decreases when triggering threshold increases. On the other hand, from the Figure 5.9 it can be implied that the overall trend of performance deteriorates by the increment of triggering threshold. Indeed, based on the bound of estimation error provided in Theorem 6, the increment of triggering threshold results in an increase of the value of ϵ_t. Based on both plots in Figure 5.9, it can be concluded that the optimal triggering threshold demands a balance between the performance index and the number of information transmission. In addition, Figure 5.9 shows that the trend of both performance index and the number of information transmission fluctuates due to the presence of process noise and measurement noise.

In the last set of simulations, the effect of various high-gain parameter bounds (θ_m) is illustrated in Figures 5.10 and 5.11. From these figures, we can see that by increasing the high-gain parameter of each filter (θ_i), the speed of convergence of measured states (T_i) increases and when their estimates get close enough to the actual states, the corresponding high-gain parameter drops quickly, and the estimator becomes regular EKF. On the other hand, Figures 5.10 and 5.11 show that increasing the gain makes an initial overshoot on the estimates of unmeasured states, and as θ_m is higher, the bigger overshoots occur. After the overshoots' occurrence, since the high-gain parameters are decreased to 1 and the standard EKFs take the action to reduce the estimation error for the rest of simulation, a bigger value for θ_m increases the convergence time of the estimates of the concentrations. That is improper tuning of parameter θ_m may affect the estimation performance.

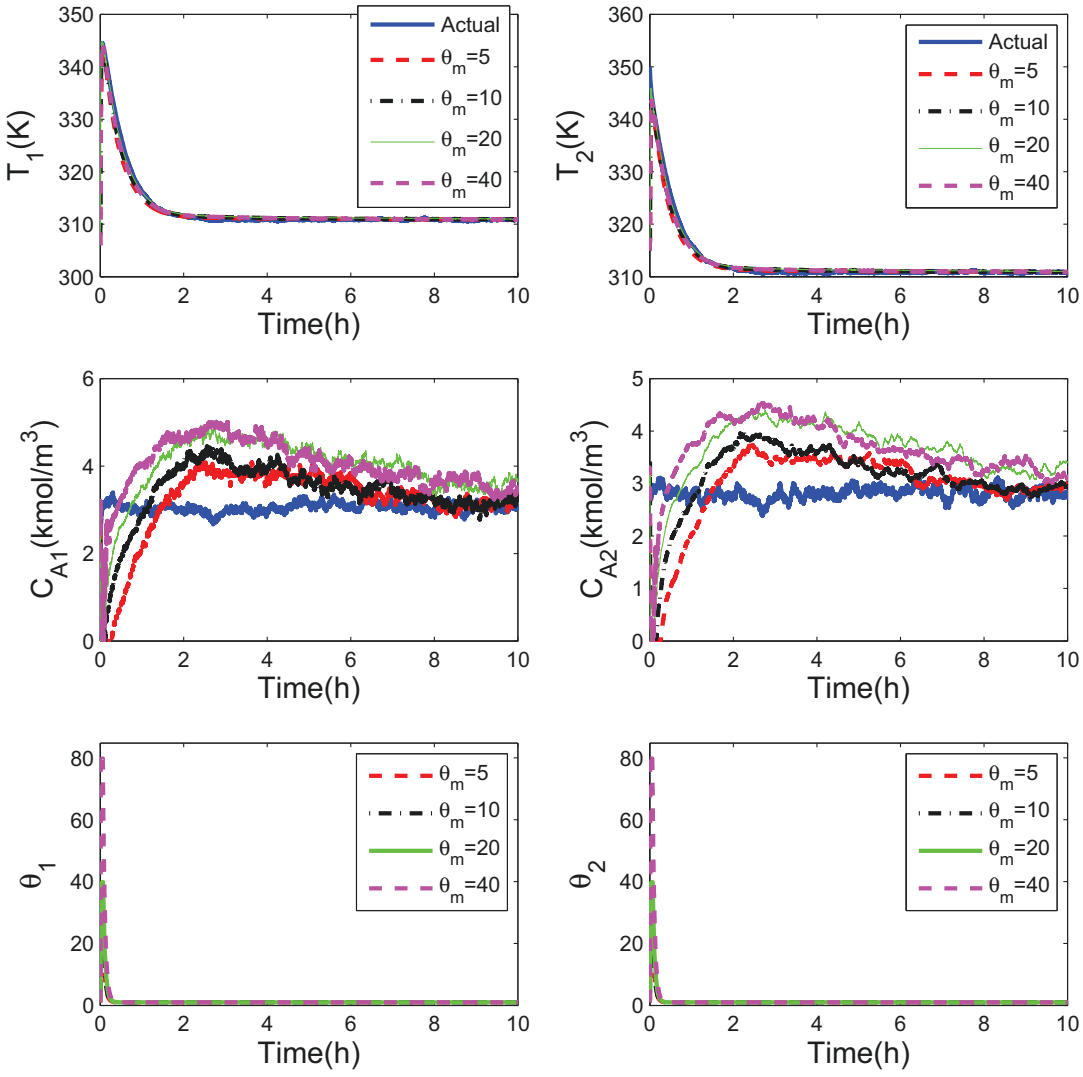

Figure 5.10: The effect of different bounds of high-gain parameters on the convergence rate of state estimates in CSTRs 1&2.

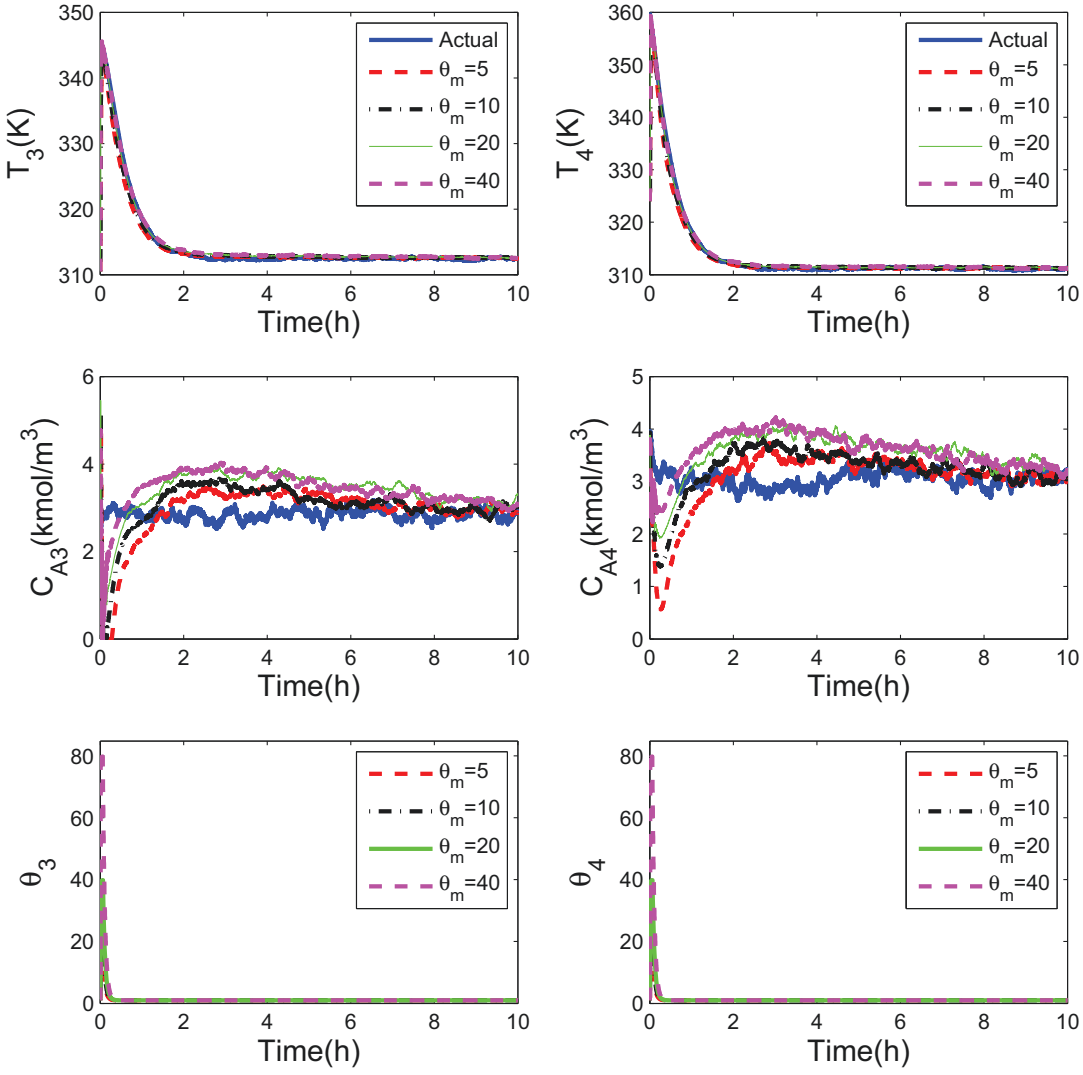

Figure 5.11: The effect of different bounds of high-gain parameters on the convergence rate of state estimates in CSTRs 3&4.

5.8 Conclusions

We developed a distributed adaptive high-gain extended Kalman filter with triggered communication for nonlinear systems. Specifically, we considered a class of nonlinear systems which can be decomposed into several interacting subsystems. Each subsystem takes the advantage of a local adaptive high-gain EKF equipped with a trigger to schedule information transmission leading to the reduction of communication frequency. Each of the designed triggers sends out information when a local triggering condition is satisfied. The triggering condition is determined based on the difference between the current state estimate and the last estimate sent out. The ultimate boundedness of the estimation error was established with sufficient conditions. Finally, a simulated chemical process example was utilized to illustrate the performance of the proposed approach.

Chapter 6

Application of distributed filtering approach to a Three-Tank system

6.1 Introduction

In this chapter we apply the proposed distributed adaptive high-gain extended Kalman filter to a Three-Tank experimental system. The system under consideration is located in process control laboratory at the University of Alberta. This system is composed of three tanks that are connected through pipes from the bottom, middle and top of the tanks as shown schematically in Figure 6.1. The water is pumped from the discharge tank to tanks 1 and 3 and the water may also flow to tank 2 when the related valves are open, and finally returns to the discharge tank through valves V_5, V_7 and V_9.

The valves V_1 to V_9 are solenoid valves which are in either ON or OFF positions. By opening or closing each of the valves in this system, a new dynamics for the whole system will be obtained. As shown in Figure 6.1, there is a level transmitter for each tank and a flow transmitter for each pump to measure the corresponding water levels and flow-rates, respectively. The information that is measured by the flow and level transmitters is sent and recorded by MATLAB on a computer connected to the system through the OPC connection.

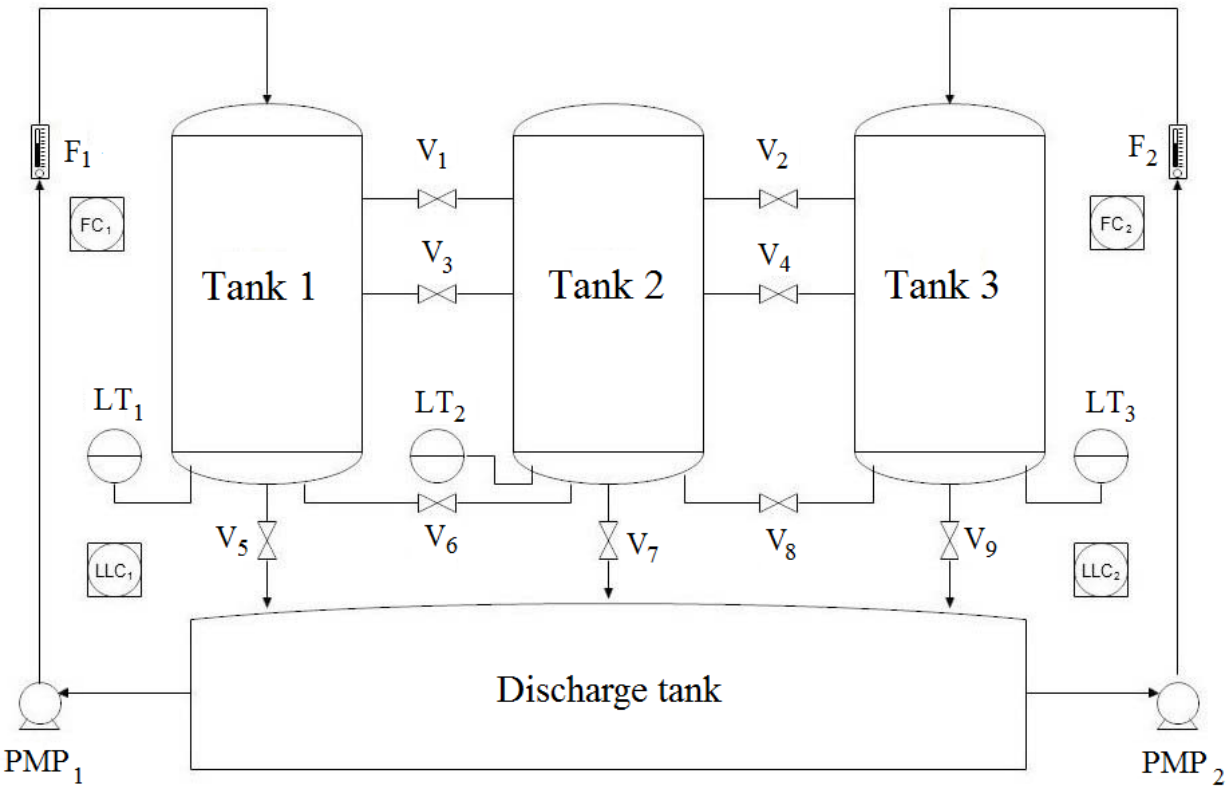

Figure 6.1: A schematic view of the experimental system

The control loops are designed and tuned in a Simulink file, and based on the received data the control signal is generated and sent to the level and flow controllers through the OPC connection.

In order to increase the robustness and reduce the disturbances to the flow rate, a cascade control circuit is designed for level control in tanks 1 and 3. This circuit is composed of two feedback control loops, primary and secondary loops, that are designed based on level and flow measurements. The primary control loop which controls the level of the tanks determines the set point of the flow control loop, and the secondary loop tunes the flow loop to reduce the flow disturbances. The cascade control loop for tank 1 is shown in Figure 6.2, where the controllers designed for level and flow are PID controllers.

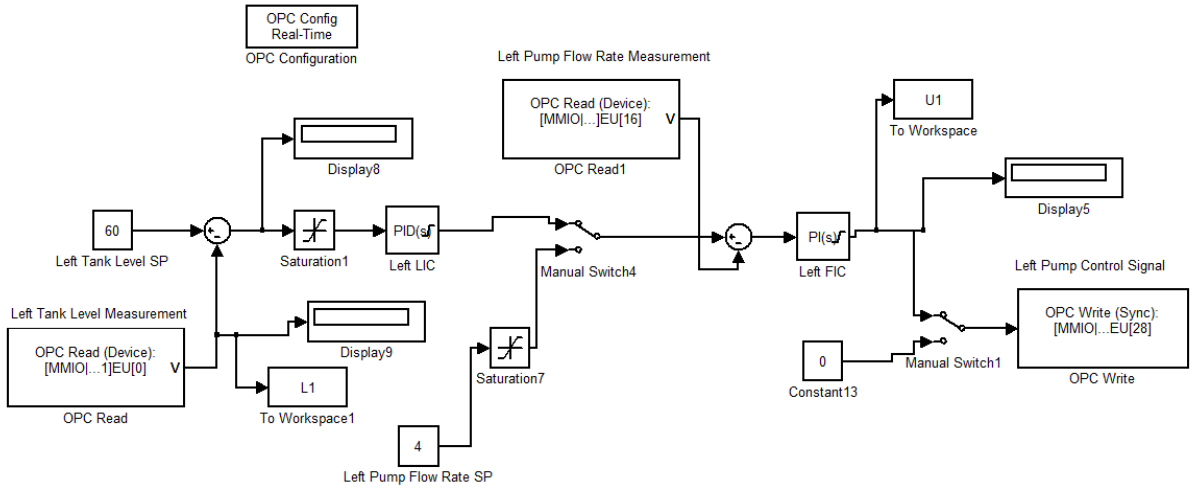

Figure 6.2: Cascade control loops for tank 1 in Three-Tank system

6.2 Model description

In spite of PID control strategies, a model of the process is required to design state estimators for any process. In order to derive a first principle model for the Three-Tank system, mass balance equations can be used. Based on the mass balance equations for each of the three tanks, i.e.

$$\{rate\ of\ accumulation\ of\ mass\} = \{rate\ of\ mass\ in\} - \{rate\ of\ mass\ out\} \tag{6.1}$$

and by considering the dynamics of interactions between the neighbor tanks, the following differential equations can be obtained:

$$\begin{aligned}
\dot{x}_1 &= (b_1 u_1 - S_{p1} a_{p1} sgn(x_2)\sqrt{2g|x_2|} - S_{d1} a_{d1}\sqrt{2g x_1})/S \\
\dot{x}_3 &= (S_{p1} a_{p1} sgn(x_2)\sqrt{2g|x_2|} + S_{p2} a_{p2} sgn(x_4)\sqrt{2g|x_4|} - S_{d2} a_{d2}\sqrt{2g x_3})/S \\
\dot{x}_5 &= (b_2 u_2 - S_{p2} a_{p2} sgn(x_4)\sqrt{2g|x_4|} - S_{d3} a_{d3}\sqrt{2g x_5})/S \\
\dot{x}_2 &= \dot{x}_1 - \dot{x}_3 \\
\dot{x}_4 &= \dot{x}_5 - \dot{x}_3
\end{aligned} \tag{6.2}$$

Table 6.1: Definition of parameters and variables for the Three-Tank system.

x_1	Level of tank 1
x_2	Interaction between tanks 1 and 2 $(x_1 - x_3)$
x_3	Level of tank 2
x_4	Interaction between tanks 3 and 2 $(x_5 - x_3)$
x_5	Level of tank 3
u_1	Input flow-rate for tank 1
u_2	Input flow-rate for tank 2
S	Cross section area of all three tanks
S_{p_i}	Pipe cross section area between tank i and $i+1$
S_{d_i}	Discharge pipe cross section area of tank i
a_{p_i}	Out-flow coefficient of the flow between tanks i and $i+1$
a_{d_i}	Out-flow coefficient from tank i to the discharge tank
g	Gravitational acceleration
b_i	Multiplier of input i which compensates for its unknown unit

where the variables and parameters are defined in Table 6.1. It should be noted that the calibration multipliers b_1 and b_2 are considered for the flow inputs. Model (6.2) is a general model for the Three-Tank system without the prior knowledge about the direction of the flow between the tanks and can be described in the following control-affine form:

$$\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned} \tag{6.3}$$

However, based on the structure of the system and the performed experiments, it rarely occurs that the level of tank 2 be higher than the levels of other tanks. As a result, model (6.2) can remain valid for the most of the operating regions of the systems even by excluding the sign function. It should be noted that, in the Three-Tank system model (6.2), the interactions are assumed to be unmeasured states and will be the objective of distributed estimation design in the following sections. Also, in this model the inputs are the flow rates to tanks 1 and 3 and the outputs are the tanks' levels that can be measured by the level

sensors, i.e.

$$y_1 = x_1,\ y_2 = x_3,\ y_3 = x_5 \tag{6.4}$$

6.2.1 Observability

Since the model in (6.2) is a multi-input multi-output nonlinear model, the observability can be verified by linearization of the model equations around the possible operating points. Indeed, based on this approach the observability can be verified by checking the rank of the observability matrix for linearized systems, i.e. $O = [C,\ CA,\ CA^2,\ \ldots]^T$ where C is the jacobian of output matrix $h(x)$ with respect to x and A is the jacobian of the state transition matrix f with respect to x. According to the performed verification, the overall system is found to be observable.

6.2.2 Parameter estimation

In model (6.2), there are some out-flow parameters which are unknown. Furthermore, the unit of the measured inputs in the actual system is not specified and two calibration parameters (b_1 and b_2) can be added as the input multipliers. In order to identify the model parameters, some persistently exciting inputs can be used and the resulting outputs together with the inputs can be used to identify the model parameters. Since there is feedback control loop for the flow-rates to the tanks, the inputs should be carefully selected. Figure 6.3 shows the elements that usually exist in closed loop identification. This figure shows that although the reference input to the Three-Tank system is injected based on the set point, the immediate input U and output Y should be collected for identification of the system. Here it should be noted that the system is nonlinear and the input signals should be able to excite the system sufficiently. For this purpose, a pseudo-random binary sequence signal can be designed where the inputs can move the system to operate within different operating points. Then by collecting the injected inputs and the resulted outputs, we collected around 12000 samples, from which 9000 samples were used for training the model and the rest was used to

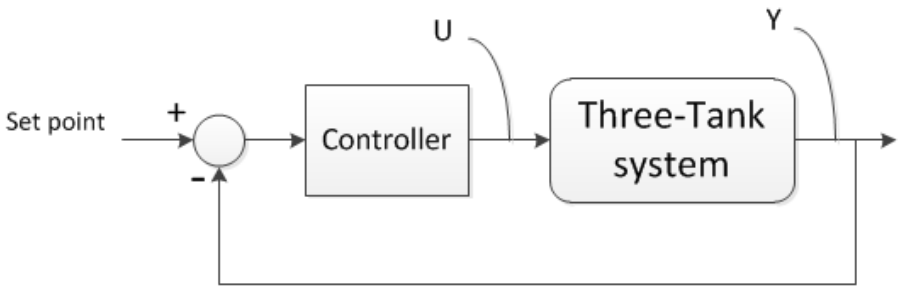

Figure 6.3: Inputs and outputs for closed loop identificaion

verify the identified model. Figure 6.4 verifies that the trajectories in identified model can track the actual states. From Figure 6.4, it is clear that the model is not accurate enough to regenerate the state of tank 2, however the mismatch between the model and actual process is inevitable from any modeling exercise. In this case, since we will focus on tanks 1 and 3, this model can be used for state estimation purposes. Finally, the identified parameters were found as follows:

$$\begin{aligned} &a_{p_1} = 3.43 && && a_{p_2} = 2.15 \\ &a_{d_1} = 18.15 && a_{d_2} = 7.78 && a_{d_3} = 18.98 \\ &b_1 = 4.56 \times 10^{-4} && && b_2 = 4.16 \times 10^{-4} \end{aligned} \tag{6.5}$$

The identified parameters in (6.5) to some extent prove the symmetry in the model, since the symmetric parameters are very close to each other. Moreover, by comparing the values of outflow parameters a_{p_1} and a_{p_2} with other outflow parameters, it can be concluded that the resistance over the flows between the tanks is higher than that between tanks and the discharge tank. In other words, the impact of the levels in tanks 1 and 3 on the level of tank 2 is not strong. Also, it should be noted that the sampling time in the experimental set-up is 1 second, and it is important to determine the sampling time in the identification. Indeed, for any specified sample time, the parameters should have been identified accordingly.

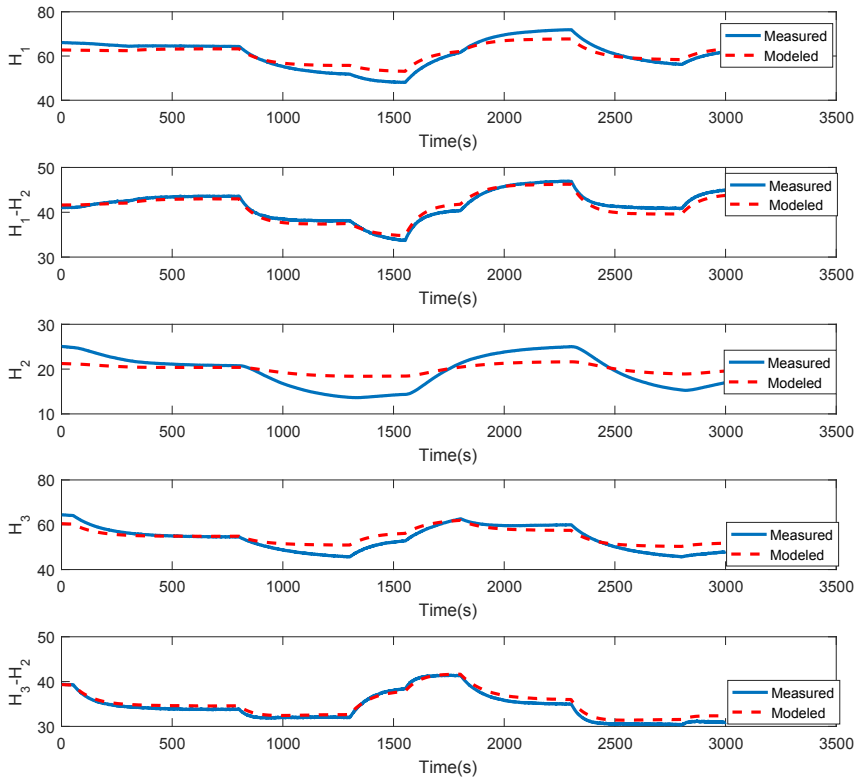

Figure 6.4: Measured and modeled trajectories to verify the identified model.

6.3 System decomposition

In order to implement the proposed distributed filtering algorithm, it is required to decompose the system into several interconnected subsystems. For this purpose, we can select the level of tank 1 and its interaction with tank 2 as the states of the first subsystem, and similarly select the level of tank 3 and its interaction with tank 2 as the states of the third subsystem and the remaining level of tank 2 as the single state of the second subsystem as shown in Figure 6.5. Based on this decomposition, the state space equations of the

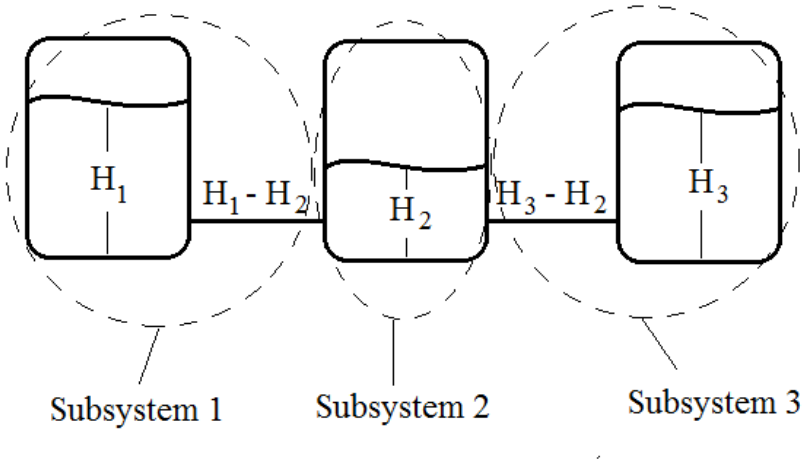

Figure 6.5: Decomposition of system to three subsystems

subsystems can be written as follows:

$$
\begin{aligned}
Subsystem\ 1 \begin{cases} \dot{r}_{11} = (b_1 u_1 - S_{p1} a_{p1} \sqrt{2 g r_{12}} - S_{d1} a_{d1} \sqrt{2 g r_{11}})/S \\ \dot{r}_{12} = \dot{r}_{11} - \dot{r}_2 \\ y_1 = r_{11} \end{cases} \\
Subsystem\ 2 \begin{cases} \dot{r}_2 = (S_{p1} a_{p1} \sqrt{2 g r_{12}} + S_{p2} a_{p2} \sqrt{2 g r_{32}} - S_{d2} a_{d2} \sqrt{2 g r_2})/S \\ y_2 = r_2 \end{cases} \\
Subsystem\ 3 \begin{cases} \dot{r}_{31} = (b_2 u_2 - S_{p2} a_{p2} \sqrt{2 g r_{32}} - S_{d3} a_{d3} \sqrt{2 g r_{31}})/S \\ \dot{r}_{32} = \dot{r}_{31} - \dot{r}_2 \\ y_3 = r_{31} \end{cases}
\end{aligned}
\tag{6.6}
$$

where $r_{11} = x_1$ and $r_{12} = x_2$ are the states of subsystem 1, $r_2 = x_2$ is the state of subsystem 2, and $r_{31} = x_5$ and $r_{32} = x_4$ are the states of subsystem 3. Furthermore, it can be seen that each subsystem has a single-input single-output model.

6.3.1 Transformation to normal form

Since distributed high-gain observer design requires the subsystems to be in normal form, we can transform all the subsystems' coordinates in (6.6) using the method described in Section

1.3. Consequently, the new coordinates for the subsystems would be obtained as follows:

$$\begin{aligned} z_1 &= r_{11} \\ z_2 &= L_f y_1 = (-S_{p1} a_{p1} \sqrt{2gr_{12}} - S_{d1} a_{d1} \sqrt{2gr_{11}})/S \\ z_3 &= r_2 \\ z_4 &= r_{31} \\ z_5 &= L_f y_3 = (-S_{p2} a_{p2} \sqrt{2gr_{32}} - S_{d3} a_{d3} \sqrt{2gr_{31}})/S \end{aligned} \tag{6.7}$$

and the dynamics of subsystems within the new coordinates can be obtained as:

$$\begin{aligned} \dot{z}_1 &= z_2 + \frac{b_1 u_1}{S} \\ \dot{z}_2 &= -\frac{1}{S}(S_{p_1} a_{p_1} \sqrt{2g} \frac{\dot{r}_{12}}{2\sqrt{r_{12}}} + S_{d_1} \sqrt{2g} \frac{\dot{z}_1}{2\sqrt{z_1}}) \\ \dot{z}_3 &= \dot{r}_2 \\ \dot{z}_4 &= z_5 + \frac{b_2 u_2}{S} \\ \dot{z}_5 &= -\frac{1}{S}(S_{p_2} a_{p_2} \sqrt{2g} \frac{\dot{r}_{32}}{2\sqrt{r_{32}}} + S_{d_3} \sqrt{2g} \frac{\dot{z}_4}{2\sqrt{z_4}}) \end{aligned} \tag{6.8}$$

where the transformations of r_{12}, $\dot{r}_{12}$, $\dot{r}_2$, r_{32} and $\dot{r}_{32}$ can be achieved from (6.7) and (6.6).

6.4 Distributed filtering design

In this section, we design the distributed filters based on the method described in Chapter 2. In order to show the effectiveness of the proposed method, we used a time-varying input signal for the Three-Tank system. These inputs cause time-varying trajectories of the states, however the distributed estimators should be tuned such that the estimates follow the actual trajectories well.

For each of the tanks an AHG-EKF is designed, however only EKF can be designed for the tank 2 since it has only one state. In the design of AHG-EKFs, $Q_1 = Q_3 = \text{diag}([0.5, 20])$, $Q_2 = 0.5$ and $R_1 = R_2 = R_3 = 0.5$. The bound of high-gain parameter is selected as $2\theta_m = 6$ and is uniquely defined for all the estimators. The forgetting horizon in the calculation of

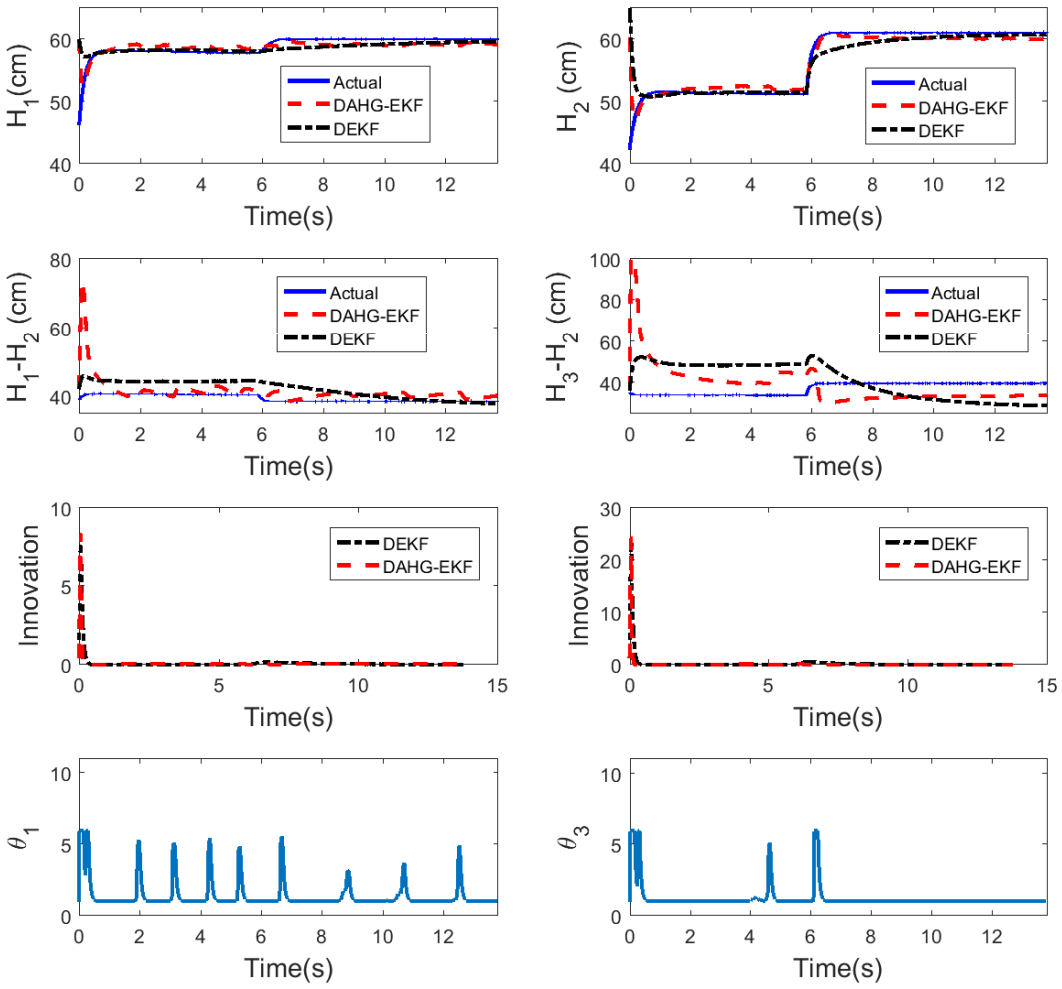

Figure 6.6: Trajectories of the levels, interactions and innovations in DEKF and DAHG-EKF frameworks with continuous communication for the first and third subsystems.

innovation terms is uniquely selected as $d = 0.03h$, and the parameters in the adaptation functions are as follows: $\Delta T = 0.001h$, $\beta_i = 450$, $m_i = 0.1$, $\lambda_i = 20$, for $i = 1, 2, 3$. Also, the initial state of the process based on (6.2) is $x(0) = [51.08, 35.02, 16.05, 40.88, 56.93]$ and the initial guesses in the three filters are $\hat{x}_1(0) = [60, 50]$, $\hat{x}_2(0) = 15$, $\hat{x}_1(0) = [50, 67]$.

Figure 6.6 shows the distributed estimation with continuous communication results for the Three-Tank system when the system takes almost constant inputs, and the system experiences a step change in inputs in the middle of operation. It should be noted that the levels of the tanks may change between $0cm$ and $100cm$, and the estimates cannot go beyond this range. As shown in Figure 6.6 the state estimates in both DEKF and DAHG-EKF

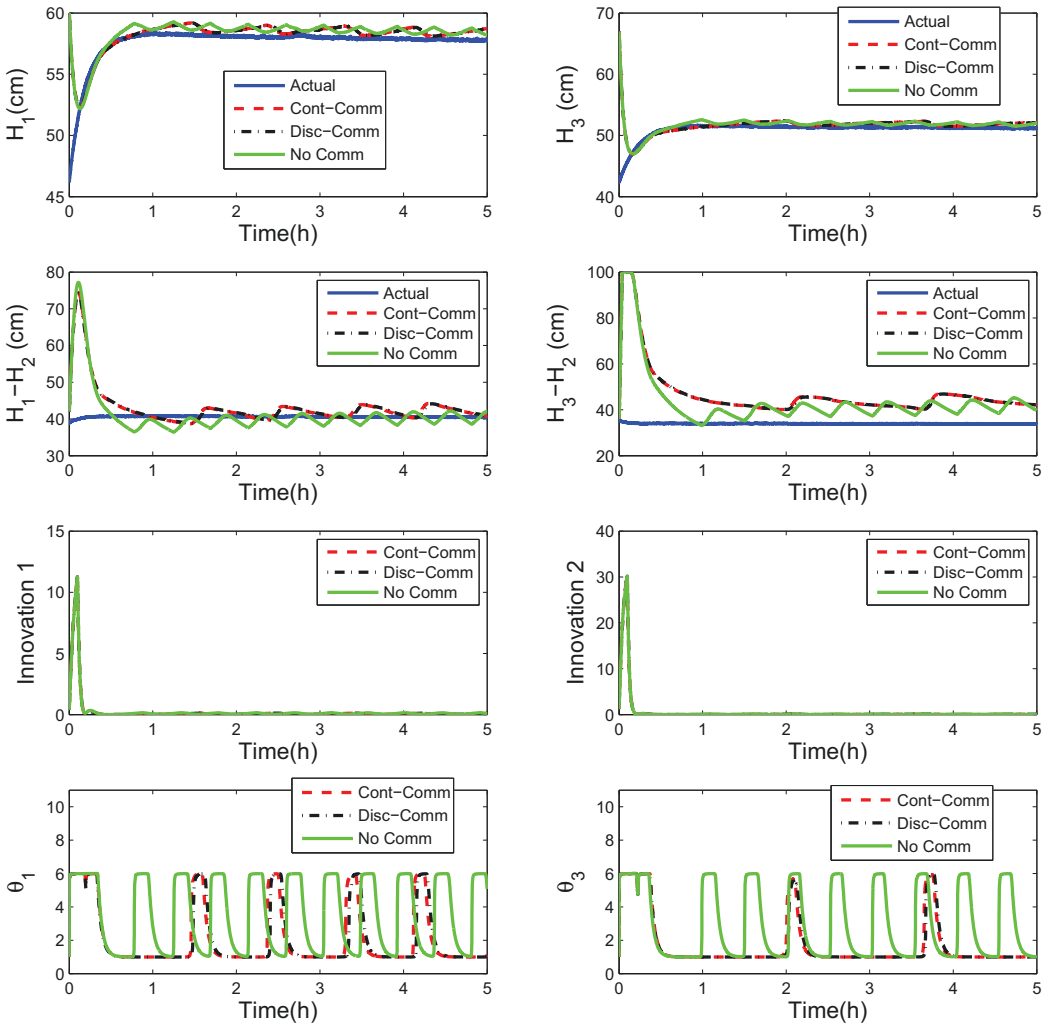

Figure 6.7: Trajectories of the levels, interactions and innovations in DAHGEKF with continuous, discrete and absence of communication for the first and third subsystems.

travel towards the actual states, however due to the process-model mismatch the estimates of the unmeasured states may have some bias. Figure 6.6 also compares the performance of distributed framework with adaptive high-gain EKF and the one with standard EKF formulation. As shown in the figure, the high-gain parameter increases whenever the innovation becomes higher than the threshold and this contributes to the result that the state estimates get closer to the actual states.

In another case, we evaluate the effects of communication on the distributed state estimation performance. Figure 6.7 shows the results of distributed estimation for three cases:

continuous communication, discrete communication with 90 seconds communication interval and decentralized estimation where no communication exists among the distributed filters. As shown in Figure 6.7, the trajectories of estimates of DAHGEKF with continuous and discrete communications are very similar while the absence of communication reduces the performance resulting in the local filters to be in high-gain mode more frequently.

6.5 Conclusion

In this chapter, we verified the proposed distributed adaptive high-gain extended Kalman filter on a Three-Tank system. This system includes three interconnected tanks and the dynamical equations of the system were derived in which the interactions between the tanks are considered as virtual states. The unknown parameters of the system were identified by injecting persistently exciting inputs to the system. Then, the identified model was decomposed into three subsystems and a DAHG-EKF was designed for each subsystem. The performance of the DAHG-EKF was compared with that of distributed regular EKF (DEKF) based on continuous communication assumption among the filters. The results show that the DAHGEKF gives improved estimates than the DEKF. Moreover, by comparing the trajectories of DAHGEKF with continuous, discrete and no communication cases, the effects of communication on the performance of DAHGEKF were evaluated.

Chapter 7

Conclusions and future work

7.1 Conclusions

This thesis considers the development and implementation of distributed adaptive high-gain extended Kalman filtering approaches on nonlinear systems. Specifically a class of nonlinear systems are considered which can be decomposed into several interacting subsystems.

To date, many algorithms have been developed for centralized estimation in which a single filter/estimator performs the task of estimation of the entire system state and decentralized estimation in which each local filter estimates the states of the corresponding subsystem only based on the local measurements. However, the former design is not favourable in terms of fault tolerance and computational loads and the latter one degrades the performance of estimation due to the absence of communication between filters. To acquire the desired performance, this thesis proposes a distributed estimation framework which removes the above difficulties by providing a communication channel between interconnected subsystems. This communication network is used to exchange state estimates between the distributed filters to compensate for the subsystems interactions.

Chapter 2 dealt with the distributed adaptive high-gain EKF (DAHGEKF) in which the distributed filters communicated continuously. The system under consideration was a type

of deterministic nonlinear system composed of interconnected subsystems. Each subsystem was assigned a local filter which receives measurement from the corresponding subsystem and state estimates from interacting filters. Indeed, the presented local filter design was an extension of the research in [10] where the multi-input multi-output system is broken into several multi-input single-output subsystems. The methodology for the communication of the distributed filters was described and the exponential convergence of the proposed approach was ensured under certain conditions. In order to demonstrate the performance of the proposed DAHGEKF, a simulated chemical process was used.

Since the continuous transmission of digital data is difficult to be established practically, we took a more realistic step in Chapter 3 and considered the filters to communicate at discrete time instants. In this framework, a state predictor was designed for each local filter to compensate for the missing information within communication intervals. The design of predictors has been previously studied in the prediction of missing measurements in distributed model predictive control [59, 17]. The asymptotic stability of the proposed approach is ensured under sufficient conditions within both deterministic and stochastic schemes of the system. To illustrate the applicability of the proposed algorithm, a 4-CSTR process is used, and the effectiveness of the predictor design is illustrated through the simulations.

In Chapter 4, one more step ahead was taken to consider data loss and delay within the discrete communication among the distributed filters. This chapter describes an extension of possible scenarios in communication losses and delays in distributed model predictive control [58] and distributed moving horizon estimation [46]. To compensate for the missing communications caused by the delays and data dropouts, a state predictor was designed for each local filter. By considering maximum allowable time delay and maximum consecutive samples of data losses within both deterministic and stochastic system structures, the worst case of information transmission was introduced under which the adequate conditions were derived to provide the stability of the distributed estimation approach. In the stochastic structure, the algorithm ensures that the overall estimation error remains bounded, however

it asymptotically converges to zero within the deterministic scheme.

The results in Chapter 3 were obtained with periodic discrete communication, however the network's capacity may not support the volume of information transmission. In Chapter 5, a DAHGEKF scheme was proposed to reduce the frequency of information transmission through the communication network. For this purpose, a triggering strategy was introduced to determine when each filter should sent out information to other filters. Indeed, this strategy contributes to reducing network traffic in the communication network which may lead to delay or loss of information. Consequently, a communication trigger was designed for each filter which sends out local information whenever the corresponding triggering condition is satisfied. The appropriate conditions are derived under which the convergence and ultimate boundedness of the estimation error is ensured.

In Chapter 6, a Three-Tank system was used to evaluate the applicability of the proposed distributed filtering algorithm. First, a dynamical model was developed for the Three-Tank system and the model parameters were identified appropriately. Then the system was decomposed into three subsystems and a local AHGEKF was designed for each subsystem. The obtained distributed estimation results demonstrated the advantage of adaptive high-gain EKF over standard EKF in distributed framework. Furthermore, the distributed estimation results illustrated that the performance is improved when there is communication between the subsystems compared with the decentralized case.

7.2 Future work directions

Within the course of the work in this thesis, some potential areas of future work can be considered. A few of these areas are listed as follows:

7.2.1 Extension to continuous-discrete case

This thesis provided an insight to the distributed adaptive high-gain extended Kalman filtering design with applied communication types. Within these designs, we always considered the subsystems which send the measurements to the corresponding filters continuously; however this may not hold in practice. Consequently, the combination of discrete communication and discrete measurements in distributed estimation may pave ways to more applicable results. Moreover, another scenario can be considered in which the sensors send the measurements to the corresponding filters with random time delays in addition to the communication issues in the distributed filtering framework.

7.2.2 Extension to coordinated DAHGEKF

Although distribute framework surpasses the centralized scheme in terms of fault tolerance and computational complexities, it may not be able to achieve the centralized performance. In order to obtain the ideal performance, a coordinator is required to make a two-level filtering, in which the upper level (coordinator) coordinates the decision making process of local filters at lower level.

7.2.3 Integration of estimation and control

State estimation task can be used both for process monitoring and control. In order for controller design in large scale systems, distributed estimation framework can be utilized to estimate the unmeasured states required for state feedback controls. So, the integration of estimation and control tasks can contribute to the applicability of the proposed approach even for unstable large scale systems.

Bibliography

[1] E. Busevelle and J-P. Gauthier. High-gain and non high-gain observers for nonlinear systems. *World Scientific*, 2002.

[2] B. Boulkroune, M. Darouach, and M. Zasadzinski. Moving horizon state estimation for linear discrete-time singular systems. *IET Control Theory and Applications*, 4:339–350, 2010.

[3] A. G. O. Mutambara. *Decentralized estimation and control for multisensor systems*. Boca Raton, FL: CRC, 1998.

[4] H. R. Hashemipour, S. Roy, and A. J. Laub. Decentralized structures for parallel Kalman filtering. *IEEE Transactions on Automatic Control*, 33:88–94, 1988.

[5] T. Keviczky, F. Borrelli, and G. J. Balas. Decentralized receding horizon control for large scale dynamically decoupled systems. *Automatica*, 42:2105–2115, 2006.

[6] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In proceedings of the *46th IEEE Conference on Decision and Control*, pages 5492–5498, New Orleans, LA, USA, 2007.

[7] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. Distributed Kalman filtering using consensus strategies. In proceedings of the *46th IEEE Conference on Decision and Control*, pages 5486–5491, New Orleans, LA, USA, 2007.

[8] R. Olfati-Saber. Distributed Kalman filter with embedded consensus filters. In proceedings of the *44th IEEE Conference on Decision and Control*, pages 8179–8184, Seville, Spain, 2005.

[9] H. Bai, R. A. Freeman, and K. M. Lynch. Distributed Kalman filtering using the internal model average consensus estimator. In proceedings of the *American Control Conference*, pages 1500–1505, San Francisco, CA, USA, 2011.

[10] Nicolas Boizot. *Adaptive high-gain extended Kalman filter and applications*. PhD thesis, University of Luxembourg, 2010.

[11] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Distributed moving horizon estimation for linear constrained systems. *IEEE Transactions on Automatic Control*, 55:2462–2475, 2010.

[12] 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.

[13] J. Zhang and J. Liu. Distributed moving horizon estimation for nonlinear systems with bounded uncertainties. *Journal of Process Control*, 23:1281–1295, 2013.

[14] J. Zhang and J. Liu. Distributed moving horizon state estimation with triggered communication. In *Proceedings of the American Control Conference*, pages 5700–5705, Portland, Oregon, USA, 2014.

[15] U. A. Khan and J. M. F. Moura. Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, 56:4919 – 4935, 2008.

[16] 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.

[17] Y. Sun and N. H. El-Farra. Quasi-decentralized model-based networked control of process systems. *Computers and Chemical Engineering*, 32:2016–2029, 2008.

[18] M. Mercangöz and F. J. Doyle. Distributed model predictive control of an experimental four-tank system. *Journal of Process Control*, 17:297–308, 2007.

[19] 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.

[20] M. Farina and R. Scattolini. Distributed model predictive control: A non-cooperative algorithm with neighbor-to-neighbor communication for linear systems. *Automatica*, 48:1088–1096, 2012.

[21] Y. Zheng, S. Li, , and N. Li. Distributed model predictive control over network information exchange for large-scale systems. *Control Engineering Practice*, 19:757–769, 2011.

[22] T. Keviczky, F. Borelli, and G.J. Balas. Stability analysis of decentralized RHC for decoupled systems. In *Proceedings of the Joint 44th IEEE Conference on Decision and Control and European Control Conference*, pages 1689–1694, Seville, Spain, 2005.

[23] B. Ding, L. Xie, and W. Cai. Distributed model predictive control for constrained linear systems. *International Journal of Robust and Nonlinear Control*, 20:1285–1298, 2010.

[24] H. Li and Y. Shi. Distributed model predictive control of constrained nonlinear systems with communication delays. *Systems & Control Letters*, 62:819–826, 2013.

[25] M. Heidarinejad, J. Liu, and P. D. Christofides. Distributed model predictive control of switched nonlinear systems. In *Proceedings of the American Control Conference*, pages 3198–3203, Montreal, Canada, 2012.

[26] R. Vadigepalli and F. J. Doyle III. A distributed state estimation and control algorithm for plantwide processes. *IEEE Transactions on Control Systems Technology*, 11:119–127, 2003.

[27] 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*, 21:767–774, 2011.

[28] S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Consensus based overlapping decentralized estimator. *IEEE Transactions on Automatic Control*, 54:410–415, 2009.

[29] M. S. Mahmoud and H. M. Khalid. Distributed Kalman filtering: a bibliographic review. *IET Control Theory and Applications*, 7:483–501, 2013.

[30] Y. Chetouani, N. Mouhab, J. M. Cosmao, and L. Estel. Application of extended Kalman filtering to chemical reactor fault detection. *Chemical Engineering Communications*, 189:1222–1241, 2002.

[31] M. A. Myers and R. H. Luecke. Process control applications of an extended Kalman filter algorithm. *Computers & Chemical Engineering*, 15:853–857, 1991.

[32] G. A. Madrid and G. J. Bierman. Application of kalman filtering to spacecraft range residual prediction. *IEEE Transactions on Automatic Control*, pages 430 – 433, 1978.

[33] A. G. O. Mutambara and H. E. Durrant-Whyte. Estimation and control for a modular wheeled mobile robot. *IEEE Transactions on Control Systems Technology*, 8:35–46, 2000.

[34] S. S. Stanković, M. S. Stanković, and D. M. Stipanović. Consensus based overlapping decentralized estimation with missing observations and communication faults. *Automatica*, 45:1397–1406, 2009.

[35] Y. Zhu, Z. You, J. Zhao, K. Zhang, and X. R. Li. The optimality for the distributed Kalman filtering fusion with feedback. *Automatica*, 37:1489–1493, 2001.

[36] M. Reinhardt, B. Noack, S. Kulkarni, and U. D. Hanebeck. Distributed Kalman filtering in the presence of packet delays and losses. In *proceedings of the 17th International conference on information fusion*, pages 1–7, Salamanca, Spain, 2014.

[37] R. Vadigepalli and Francis J. Doyle III. Structural analysis of large-scale systems for distributed state estimation and control applications. *Control Engineering Practice*, 11:895–905, 2003.

[38] A. G. O. Mutambara. *Decentralized estimation and control for multisensor systems*. CRC Press, 1998.

[39] M. Rashedi, J. Liu, and B. Huang. Distributed adaptive high-gain extended Kalman filtering for nonlinear systems. In *proceedings of the 9th International Symposium on Advanced Control of Chemical Processes*, volume 48, pages 158–163, Whistler, BC, Canada, 2015.

[40] C. V. Rao, J. B. Rawlings, and J. H. Lee. Constrained linear state estimation - A moving horizon approach. *Automatica*, 37:1619–1628, 2001.

[41] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Moving-horizon partition-based state estimation of large-scale systems. *Automatica*, 46:910 – 918, 2010.

[42] 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 (ECC) 2013*, pages 2621–2626, Zurich, Switzerland, 2013.

[43] J. Zhang and J. Liu. Observer-enhanced distributed moving horizon state estimation subject to communication delays. *Journal of Process Control*, 24:672–686, 2014.

[44] N. Boizot. *Adaptive High-Gain Extended Kalman Filter and Applications*. PhD thesis, Universite de Bourgogne, 2011.

[45] N. Boizot, E. Busevelle, and J. P. Gauthier. Adaptive-gain extended Kalman filter: extension to the continuous-discrete case. In *10th European Control Conference (ECC'09)*, pages 4570–4575, Budapest, Hungary, 2009.

[46] J. Zeng and J. Liu. Distributed moving horizon state estimation: Simultaneously handling communication data losses. *Systems & Control Letters*, 75:56–68, 2015.

[47] A. Anta and P. Tabuada. To sample or not to sample: Self-triggered control for nonlinear systems. *IEEE Transactions on Automatic Control*, 55:2030–2042, 2010.

[48] P. Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *IEEE Transactions on Automatic Control*, 52:1680–1685, 2007.

[49] L. Li, M. Lemmon, and X. Wang. Event-triggered state estimation in vector linear processes. In *Proceedings of the American Control Conference*, pages 2138–2143, Baltimore, Maryland, USA, 2010.

[50] J. Weimer, J. Araujo, and K. H. Johansson. Distributed event-triggered estimation in networked systems. In *Proceedings of the 4th IFAC Conferences on Analysis and Design of Hybrid Systems*, pages 178–185, Eindhoven, Netherlands, 2012.

[51] Y. Hu and N. H. El-Farra. Quasi-decentralized output feedback model predictive control of networked process systems with forecast-triggered communication. In *Proceedings of the American Control Conference*, pages 2612–2617, Washington, DC, USA, 2013.

[52] N. Boizot, E. Busvelle, and J.-P. Gauthier. An adaptive high-gain observer for nonlinear systems. *Automatica*, 46:1483–1488, 2010.

[53] J.R.C. Alarcon, H.R. Cortes, and E.V. Vivas. Extended Kalman tuning in attitude estimation from inertial and magnetic field measurements. In *proceedings of the 6th*

International Conference on Electrical Engineering, Computing Science and Automatic Control, pages 1–6, Toluca, Mexico, 2009.

[54] G. Lin, Z. Jing, and Z. Liu. Tuning of extended Kalman filter using improved particle swarm optimization for sensorless control of induction motor. *Journal of Computational Information Systems*, 10:2455–2462, 2014.

[55] B.N. Alsuwaidan, J.L. Crassidis, and Y. Cheng. Convergence properties of autocorrelation-based generalized multiple-model adaptive estimation. In *AIAA Guidance, Navigation and Control Conference*, pages 1–20, Honolulu, Hawaii, USA, 2008.

[56] K. Robenack. Computation of multiple lie derivatives by algorithmic diffrentiation. *Journal of Computational and Applied Mathematics,*, 213:454–464, 2008.

[57] R. Marino and P. Tomei. *Nonlinear control design*. Prentice Hall Europe, 1995.

[58] 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*, 46:52–61, 2010.

[59] B. T. Stewart, A. N. Venkat, J. B. Rawlings, S. J. Wright, and G. Pannocchia. Cooperative distributed model predictive control. *Systems and Control Letters*, 59:460–469, 2010.

[60] J-P Gauthier and I. Kupka. *Deterministic observation theory and applications*. Cambridge University Press, Cambridge, UK, 2001.

[61] O. Hlinka, O. Sluciak, F. Hlawatsch, P. M. Djuric, and M. Rupp. Likelihood consensus and its application to distributed particle filtering. *IEEE Transactions on Signal Processing*, 60:4334–4349, 2012.

[62] M. Rashedi, J. Liu, and B. Huang. Communication delays and data losses in distributed adaptive high-gain EKF. *AICHE*, 62:4321–4333, 2016.

Appendix A

Derivation of the Riccati equation

A.1 Preliminaries

For an unforced, time-varying and linear dynamic system

$$\dot{x}(t) = A(t)x(t) \tag{A.1}$$

with a known initial condition $x(t_0)$, the general solution is

$$x(t) = \phi(t, t_0)x(t_0) \tag{A.2}$$

From (A.2), the followings properties can be inferred for $\phi(t, t_0)$:

$$\begin{aligned} &\phi(t_0, t_0) = I \\ &\phi(t_0, t) = \phi^{-1}(t, t_0) \\ &\phi(t_2, t_0) = \phi(t_2, t_1)\phi(t_1, t_0) \end{aligned} \tag{A.3}$$

By replacing (A.2) into (A.1) we also obtain,

$$\dot{\phi}(t, t_0)x(t_0) = A(t)\phi(t, t_0)x(t_0) \;\Rightarrow\; \dot{\phi}(t, t_0) = A(t)\phi(t, t_0) \tag{A.4}$$

In another case, let us consider a forced, time-varying linear system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{A.5}$$

For the case of $u(t) \neq 0$ we seek to replace $x(t_0)$ by a function $g(t)$ which satisfies

$$x(t) = \phi(t, t_0)g(t) \tag{A.6}$$

where $g(t)$ is a vector of unknown functions and it is clear that $g(t_0) = x(t_0)$. By differentiating (A.6)

$$\dot{x}(t) = \phi(t, t_0)\dot{g}(t) + \dot{\phi}(t, t_0)g(t) = \phi(t, t_0)\dot{g}(t) + A(t)\phi(t, t_0)g(t) \tag{A.7}$$

Substituting (A.6) in (A.5) and comparing the result with (A.7) we obtain,

$$\phi(t, t_0)\dot{g}(t) + A(t)\phi(t, t_0)g(t) = A(t)\phi(t, t_0)g(t) + B(t)u(t) \tag{A.8}$$

Therefore,

$$\dot{g}(t) = \phi^{-1}(t, t_0)B(t)u(t) \tag{A.9}$$

whose integration result in,

$$g(t) = x(t_0) + \int_{t_0}^{t} \phi^{-1}(\tau, t_0)B(\tau)u(\tau)d\tau \tag{A.10}$$

Substituting (A.10) into (A.6), the general solution would be obtained

$$x(t) = \phi(t, t_0)x(t_0) + \phi(t, t_0) \int_{t_0}^{t} \phi^{-1}(\tau, t_0)B(\tau)u(\tau)d\tau \tag{A.11}$$

According to (A.3) we know that

$$\phi(t_0,\tau) = \phi(t_0,t)\phi(t,\tau) \;\Rightarrow\; \phi^{-1}(\tau,t_0) = \phi^{-1}(t,t_0)\phi(t,\tau) \tag{A.12}$$

Finally, by plugging the result of (A.12) into (A.11) we obtain

$$x(t) = \phi(t,t_0)x(t_0) + \int_{t_0}^{t} \phi(t,\tau)B(\tau)u(\tau)d\tau \tag{A.13}$$

A.2 Description

We consider an observable nonlinear system whose subsystems' dynamics are described in the following canonical form:

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i(t) + b_i(x(t),u(t)) + w_i(t), \quad w_i(t) \sim N(0,Q_i(t)) \\ y_i(t) &= C_i x_i(t) + v_i(t), \quad v_i(t) \sim N(0,R_i(t)) \end{aligned} \tag{A.14}$$

Then the dynamics of filter i can be as follows:

$$\dot{z}_i(t) = A_i z_i(t) + b_i(z(t),u(t)) + k_i(t)(y_i(t) - C_i z_i(t)) \tag{A.15}$$

Defining the subsystem's estimation error as $\epsilon(t) = z(t) - x(t)$ we obtain

$$\dot{\epsilon}_i(t) = F_i(t)\epsilon_i(t) + b_i(z(t),u(t)) - b_i(x(t),u(t)) + \lambda_i(t) \tag{A.16}$$

where $F_i(t) = A_i - k_i(t)C_i$, $\lambda_i(t) = k_i(t)v_i(t) - w_i(t)$. From the definition of λ_i, it can be found that,

$$E\{\lambda_i(t)\lambda_i(t)^T\} = Q_i(t) + k_i(t)R_i(t)k_i(t)^T \tag{A.17}$$

Also, by linearizing $b_i(x)$ in the neighborhood of z we obtain that

$$b_i(x,u) \approx b_i(z,u) + \frac{\partial b_i(x,u)}{\partial x}|_z(x-z) = b_i(z,u) - b_i^*(z)\epsilon_i + \sum_{j\in\mathbb{I}_i\setminus i} \frac{\partial b_i(z,u)}{\partial z_j}(x_j - z_j) \quad \text{(A.18)}$$

where $b_i^*(z) = \dfrac{\partial b_i(z)}{\partial z_i}$. If we neglect the last term on the right hand side of (A.18), and use the remaining in (A.16) we obtain

$$\dot{\epsilon}_i(t) = (F_i(t) + b_i^*(z(t),u(t)))\epsilon_i(t) + \lambda_i(t) \quad \text{(A.19)}$$

Using the matrix exponential solution, the solution to the system (A.19) can be found as

$$\epsilon_i(t) = \phi_i(t,t_0)\epsilon_i(t_0) + \int_{t_0}^{t} \phi_i(t,\tau)\lambda_i(\tau)d\tau \quad \text{(A.20)}$$

where $\dot{\phi}_i(t,t_0) = [F_i(t) + b_i^*(z(t),u(t))]\phi_i(t,t_0)$. The state error covariance is defined by

$$P_i(t) = E\{\epsilon_i(t)\epsilon_i(t)^T\} \quad \text{(A.21)}$$

Substituting (A.20) into (A.21) and assuming that $\lambda_i(t)$ and $\epsilon_i(t_0)$ are uncorrelated,

$$P_i(t) = \phi_i(t,t_0)P_i(t_0)\phi_i^T(t,t_0) + \int_{t_0}^{t} \phi_i(t,\tau)[Q_i(t) + k_i(t)R_i(t)k_i(t)^T]\phi_i^T(t,\tau)d\tau \quad \text{(A.22)}$$

Taking the time derivative of (A.22) gives

$$\begin{aligned}
\dot{P}_i(t) &= \frac{\partial \phi_i(t,t_0)}{\partial t}P(t_0)\phi_i^T(t,t_0) + \phi_i(t,t_0)P_i(t_0)\frac{\partial \phi_i^T(t,t_0)}{\partial t} \\
&+ \textstyle\int_{t_0}^{t} \dfrac{\partial \phi_i(t,\tau)}{\partial t}[Q_i(\tau) + k_i(\tau)R_i(\tau)k_i(\tau)^T]\phi_i^T(t,\tau)d\tau \\
&+ \textstyle\int_{t_0}^{t} \phi_i(t,\tau)[Q_i(\tau) + k_i(\tau)R_i(\tau)k_i(\tau)^T]\dfrac{\partial \phi_i^T(t,\tau)}{\partial t}d\tau \\
&+ \phi_i(t,t)[Q_i(t) + k_i(t)R_i(t)k_i(t)^T]\phi_i^T(t,t)
\end{aligned} \quad \text{(A.23)}$$

Using the properties of the matrix exponential we have

$$\frac{\partial \phi_i(t,t_0)}{\partial t} = (F_i(t) + b_i^*(z(t),u(t)))\phi_i(t,t_0) \tag{A.24}$$

and (A.23) would be changed to

$$\begin{aligned}
&\dot{P}_i(t) = (F_i(t) + b_i^*(z(t),u(t)))\phi_i(t,t_0)P_i(t_0)\phi_i^T(t,t_0) + \phi_i(t,t_0)P_i(t_0)\phi_i^T(t,t_0)(F_i(t) + b_i^*(z(t)))^T \\
&+(F_i(t) + b_i^*(z(t),u(t))) \int_{t_0}^{t} \phi_i(t,\tau)[Q_i(\tau) + k_i(\tau)R_i(\tau)k_i(\tau)^T]\phi_i^T(t,\tau)d\tau \\
&+ \int_{t_0}^{t} \phi_i(t,\tau)[Q_i(\tau) + k_i(\tau)R_i(\tau)k_i(\tau)^T]\phi_i^T(t,\tau)d\tau(F_i(t) + b_i^*(z(t),u(t)))^T \\
&+\phi_i(t,t)[Q_i(t) + k_i(t)R_i(t)k_i(t)^T]\phi_i^T(t,t)
\end{aligned} \tag{A.25}$$

Based on (A.22), (A.25) changes to

$$\begin{aligned}
&\dot{P}_i(t) = [A_i - k_i(t)C_i + b_i^*(z(t),u(t))]P_i(t) + P_i(t)[A_i - k_i(t)C_i + b_i^*(z(t),u(t))]^T + Q_i(t) \\
&+k_i(t)R_i(t)k_i(t)^T = [A_i + b_i^*(z(t),u(t))]P_i(t) + P_i(t)[A_i + b_i^*(z(t),u(t))]^T + Q_i(t) \\
&+k_i(t)R_i(t)k_i(t)^T - k_i(t)C_iP_i(t) - P_i(t)C_i^Tk_i^T(t) + P_i(t)C_i^TR_i^{-1}(t)C_iP_i(t) \\
&-P_i(t)C_i^TR_i^{-1}(t)C_iP_i(t) = [A_i + b_i^*(z(t),u(t))]P_i(t) + P_i(t)[A_i + b_i^*(z(t),u(t))]^T \\
&+Q_i(t) - P_i(t)C_i^TR_i^{-1}(t)C_iP_i(t) + \left[k_i(t)R_i(t) - P_i(t)C_i^T\right]R_i^{-1}\left[k_i(t)R_i(t) - P_i(t)C_i^T\right]^T
\end{aligned} \tag{A.26}$$

We need to have $k_i(t)$ such that $P_i(t)$ is as small as possible. So, if $k_i(t)R_i(t) = P_i(t)C_i^T$ then $\dot{P}_i(t)$ would be minimized and the optimal gain would be,

$$k_i(t) = P_i(t)C_i^TR_i^{-1}(t) \tag{A.27}$$

Replacing (A.27) into (A.26) we obtain

$$\dot{P}_i(t) = [A_i + b_i^*(z(t), u(t))]P_i(t) + P_i(t)[A_i + b_i^*(z(t), u(t))]^T + Q_i(t) - P_i(t)C_i^T R_i^{-1}(t)C_i P_i(t) \tag{A.28}$$

and defining $S_i(t) = P_i^{-1}(t)$, from (A.28) it can be obtained that,

$$\dot{S}_i(t) = -[A_i + b_i^*(z(t), u(t))]^T S_i(t) - S_i(t)[A_i + b_i^*(z(t), u(t))] - S_i^T(t)Q_i(t)S_i(t) + C_i^T R_i^{-1}(t)C_i \tag{A.29}$$

Appendix B

Lipschitz constant of normal form nonlinearities

In Chapter 2, it was discussed that the Lipschitz constants of the vector filed $b(x,u)$ and $\tilde{b}(\tilde{x},u)$ as well as the bound of $\tilde{b}^*(\tilde{x},u)$ and $b(x,u)$ are equivalent. Although this result is proved on page 215 of [60] for a centralized filter, here we prove it for distributed framework. First we recall from Chapter 2 that the high-gain parameter satisfies $\theta(t) \geq 1$. Then, from the observable structure of the nonlinear function $\tilde{b}_i(\cdot,u)$ for subsystem i ($i \in \mathbb{I}$), i.e.

$$b_i(x,u) = \begin{bmatrix} b_{i,1}(x_{i,1},u) \\ b_{i,2}(x_{i,1},x_{i,2},u) \\ \vdots \\ b_{i,n_{x_i}-1}(x_{i,1},x_{i,2},\ldots,x_{i,n_{x,i}-1},u) \\ b_{i,n_{x_i}}(x,u) \end{bmatrix} \tag{B.1}$$

we consider a component $\tilde{b}_i^k(\cdot,u)$, $k \in \{1,\ldots,n_{x_i}-1\}$ and from the definition of $\tilde{b}$ (i.e. $\tilde{b}_i(\cdot,u) = \Theta_i b_i(\Theta_c^{-1}\cdot,u)$), we can obtain that

$$\tilde{b}_{i,k}(x,u) = \frac{1}{\theta^{n^*-n_{x_i}+k-1}} b(\theta^{n^*-n_{x_i}}x_{i,1},\ldots,\theta^{n^*-n_{x_i}+k-1}x_{i,k},u) \tag{B.2}$$

and by considering L_b as the Lipschitz constant of $b(x,u)$ with respect to variable x, it can be obtained that,

$$\begin{aligned}
&\|\tilde{b}_{i,k}(x,u)-\tilde{b}_{i,k}(z,u)\|= \\
&\tfrac{1}{\theta^{n^*-n_{x_i}+k-1}}\|b(\theta^{n^*-n_{x_i}}x_{i,1},\ldots,\theta^{n^*-n_{x_i}+k-1}x_{i,k},u)-b(\theta^{n^*-n_{x_i}}z_{i,1},\ldots,\theta^{n^*-n_{x_i}+k-1}z_{i,k},u)\| \\
&\leq \tfrac{L_b}{\theta^{n^*-n_{x_i}+k-1}}\|(\theta^{n^*-n_{x_i}}x_{i,1},\ldots,\theta^{n^*-n_{x_i}+k-1}x_{i,k},u)-(\theta^{n^*-n_{x_i}}z_{i,1},\ldots,\theta^{n^*-n_{x_i}+k-1}z_{i,k},u)\| \\
&\leq \tfrac{L_b}{\theta^{n^*-n_{x_i}+k-1}}\theta^{n^*-n_{x_i}+k-1}\|(x_{i,1},\ldots,x_{i,k},u)-(z_{i,1},\ldots,z_{i,k},u)\| \\
&\leq L_b\|(x_{i,1},\ldots,x_{i,k},u)-(z_{i,1},\ldots,z_{i,k},u)\|
\end{aligned} \tag{B.3}$$

which means the Lipschitz constants of both b and $\tilde{b}$ are equivalent. It should be noted that in the derivation of (B.3) we considered all the elements of $\tilde{b}_i$ except the last one which may be a function of system's overall states x. Now we consider the element $\tilde{b}_{i,n_{x_i}}$ and prove that its Lipschitz constant remains consistent.

For the states of the overall system we know that

$$\|\Theta_c^{-1}x-\Theta_c^{-1}z\|\leq \theta^{n^*-1}\|x-z\| \tag{B.4}$$

and also based on the definition of the matrix Θ_c for all n_{x_i}, $i\in\mathbb{I}$, we have

$$\tilde{b}_{i,n_{x_i}}(x,u)=\frac{1}{\theta^{n^*-1}}b_{i,n_{x_i}}(x,u) \tag{B.5}$$

Hence, according to (B.4) and (B.5) we obtain,

$$\|\tilde{b}_{i,n_{x_i}}(\tilde{x},u)-\tilde{b}_{i,n_{x_i}}(\tilde{z},u)\|\leq L_b\|\tilde{x}-\tilde{z}\| \tag{B.6}$$

which implies that the change of coordinates does not change the Lipschitz constant.

The above consistency holds for the Jacobian matrix $\tilde{b}^*(\cdot,u)$ as well. In order to prove this, we consider an element of the matrix $\tilde{b}$ denoted as $\tilde{b}^*_{i,j}$ and based on the definition of

the matrix $\tilde{b}$ we obtain that

$$\tilde{b}^*_{i,j}(\tilde{z}, u) = \frac{1}{\theta^p} b^*_{i,j}(\tilde{z}, u)\theta^q \tag{B.7}$$

where p and q are natural numbers and $p \geq q$. Not that due to the structure of b_i, as described in (B.3), the Jacobian matrix is lower triangular. Consequently,

$$\|\tilde{b}^*_{i,j}(\tilde{z}, u)\| \leq \theta^{q-p} \|b^*_{i,j}(\tilde{z}, u)\| \leq \|b^*_{i,j}(\tilde{z}, u)\| \leq L_{b^*} \tag{B.8}$$