Distributed Moving Horizon State Estimation of Nonlinear Systems

by

Jing Zhang

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

Master of Science

in

Chemical Engineering

Department of Chemical and Materials Engineering
University of Alberta

Abstract

Large-scale complex chemical processes increasingly appear in the modern process industry due to their economic efficiency. Such a large-scale complex chemical process usually consists of several unit operations (subsystems), which are connected together through material and energy flows. Because of the increased process scale and the significant interactions between different subsystems, it poses great challenges in the design of automatic control systems for such large-scale complex chemical processes which are desired to fulfill the fundamental safety, environmental sustainability and profitability requirements. In recent years, predictive process control has emerged as an attractive control approach to handle the scale and interactions of large-scale complex chemical processes. It has been demonstrated that predictive process control can achieve improved closed-loop performance compared with decentralized control while preserving the flexibility of the decentralized framework. However, almost all of the existing distributed predictive process control designs are developed under the assumption that the state measurements of subsystems or the entire system are available. This assumption does not hold in many applications.

This thesis presents a robust distributed moving horizon state estimation (DMHE) scheme that is appropriate for output feedback distributed predictive control of nonlinear systems as well as approaches for reducing the communication demand of the proposed DMHE scheme and a strategy for handling delays in the communication between subsystem estimators. First, the proposed robust DMHE scheme is presented for a class of nonlinear systems that are composed of several subsystems. It is assumed that the subsystems interact with each other via their states only. Subsequently, two triggered communication algorithms are introduced for the proposed DMHE scheme to reduce the number of information transmissions between subsystems. Following this, an approach is proposed to handle the potential time-varying delays in the communication between the subsystem estimators. The applicability and effectiveness of the proposed approaches are illustrated via their applications to different chemical process examples.

Acknowledgements

It gives me great pleasure in acknowledging the patience, support and help of Professor Jinfeng Liu, as well as his friendship with me during my graduate studies. This thesis would have remained a dream had it not been for the guidance of Prof. Liu and his mentorship was significant in assisting to build up my long-time career goals and experiences. As an undergraduate student graduated from the Department of Polymer Science and Engineering, I know little about mathematics and programming. I was not sure if I could make it to transfer from experimental science to computational science and I was ready to get my master's degree in three years. To help me better involved in this specialty, Prof. Liu even gave me lectures on the basics of math, control theories and programming. I would like to give my appreciation to Prof. Liu again for his trust, cultivation and comprehensiveness.

I would like to thanks Professor Biao Huang for his help in system identification in my project and the access to Prof. Huang's general group meeting. I am indebted to my many colleagues in CPC groups who supported me: Tianbo Liu, Shuning Li, Kangkang Zhang, Ruben Gonzalez, Yaojie Lu, Ming Ma, Liu Liu, Zhankun Xi, Mulang Chen, Chen Li, Su Liu, Yujia Zhao, Ruomu Tan and Ouyang Wu. I also wish to thank my friends who makes my graduate life so wonderful: Ran Li, Qian Fu, Qi Zhang, Xi Huang and Steven Myroniuk.

I owe my deepest gratitude to my parents and siblings. Without their support and understanding of my graduate studies, I would never accomplished such a huge improvement in my career. I hope They'd like to be proud of me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Due to the increasing global competition, large-scale complex chemical processes is common appearances in the modern process industry due to their economic efficiency. In recent years, distributed predictive control has emerged as an attractive control approach to handle the scale and interactions of large-scale complex chemical processes. It has been demonstrated that distributed predictive control can achieve improved (sometimes the centralized) closed-loop performance while preserving the flexibility of the decentralized framework. However, almost all of the existing distributed predictive control designs were developed under the assumption that the state measurements of subsystems or the entire system are available. It is in general difficult to measure all the state variables in a process system. In order to maintain the structural flexibility of distributed predictive control, distributed or decentralized state estimation systems should be used instead of centralized observers.

There are many existing results on decentralized deterministic observer designs for different classes of systems (e.g., [1, 2, 3, 4]) and distributed Kalman filtering based on consensus algorithms with applications to sensor networks (e.g, [5, 6, 7, 8]). These results are primarily developed in the context of linear systems. Recently, in [9], a distributed state estimation approach for linear systems was developed in the framework of moving horizon estimation (MHE) which was extended to nonlinear systems in [10]. However, these designs are not appropriate for output-feedback control. Motivated by the above considerations, in this thesis, we present a robust distributed MHE (DMHE) design for a class of nonlinear systems with bounded output measurement noise and process disturbances. The proposed DMHE design has the potential to be used in output-feedback control.

1.2 Background

Large-scale complex chemical processes increasingly appear in the modern process industry due to their economic efficiency. Such a large-scale complex chemical process usually consists of several unit operations (subsystems), which are connected together through material and energy flows. Because of the increased process scale and the significant interactions between different subsystems, it poses great challenges in the design of automatic control systems for such large-scale complex chemical processes which are desired to fulfill the fundamental safety, environmental sustainability and profitability requirements. Traditionally, control and state estimation of large-scale systems has been studied primarily within the centralized and the decentralized frameworks. While the centralized framework is shown to provide the best performance, it is not favorable from the computational and fault tolerance view points. In a decentralized framework, the interactions between subsystems in general are either not taken into account or accounted for in conservative fashions such as worst case scenarios (e.g., [11, 12] and references therein). Decentralized framework in general has a reduce complexity in the controller and observer design and implementation. However, it may lead to deteriorated performance or even lost of closed-loop stability.

In recent years, distributed model predictive control (DMPC) has emerged as an attractive control approach to handle the scale and interactions of large-scale complex chemical processes; please see [13, 14, 15] for reviews of results on DMPC. The existing DMPC algorithms can be broadly classified into non-cooperative and cooperative DMPC algorithms based on the cost function used in the local controller optimization problem [13]. In a non-cooperative DMPC algorithm, each local controller optimizes a local cost function while in a cooperative DMPC algorithm, a local controller optimizes a global cost function. Non-cooperative DMPC algorithms include [16, 17, 18, 19, 20, 21]. Cooperative DMPC was first proposed in [22] and was developed in [13, 23, 24]. Lyapunov-based cooperative DMPC algorithms for nonlinear systems were also developed in [25, 26] in recent years. It has been demonstrated that DMPC has the potential to achieve the performance of the centralized control while preserving the flexibility of decentralized frameworks [23, 15]. In addition to DMPC, other important work within process control includes the development of a quasi-decentralized control framework for multi-unit plants that achieves the desired closed-loop objectives with minimal cross communication between the plant units under state feedback control [27]. However, almost all of the above results are derived under the assumptions that the system states are available all the times or that a centralized state observer is available. These assumptions, however, either fail in many applications or are inconsistent with the distributed framework which is not favorable from a fault tolerance point of view. Therefore, it is desirable to develop state estimation schemes in the distributed framework.

In the literature, a majority of the existing results on state observer designs are derived in the centralized framework. For linear systems, Kalman filters and Luenberger observers are standard solutions. In the context of nonlinear systems, observer designs including high-gain observers for different specific classes of nonlinear systems are available (e.g., [28, 29, 30, 31, 32, 33, 34, 35, 36, 37]). In a recent work [38], observers for systems with delayed measurements were also developed. It is worth noting that the capability of high-gain observers to be used in output feedback control designs has made high-gain observers very popular in output feedback control of nonlinear systems (e.g., [39, 40, 41, 42, 43, 44]).

In another line of work, MHE has become popular because of its ability to handle explicitly nonlinear systems and constraints on decision variables (e.g., [45, 46, 47, 48]). In MHE, the state estimate is determined by solving online an optimization problem that minimizes the sum of squared errors. In order to have a finite dimensional optimization problem, the horizon (estimation window size into the past) of MHE is in general chosen to be finite. At a sampling time, when a new measurement is available, the oldest measurement in the estimation window is discarded, and the finite horizon optimization problem is solved again to get the new estimate of the state [49, 45]. In a recent work [50], a robust MHE scheme was developed which effectively integrates deterministic (high-gain) observers into the MHE framework. The resulting robust MHE scheme gives bounded estimation error and has a tunable convergence rate.

In order to maintain the structural flexibility of DMPC, distributed or decentralized state estimation systems should be used instead of centralized observers. There are many existing results on decentralized deterministic observer designs for different classes of systems (e.g., [1, 2, 3, 4]) and distributed Kalman filtering based on consensus algorithms with applications to sensor networks (e.g, [5, 6, 7, 8]). These results are primarily developed in the context of linear systems. Recently, in [9], a distributed state estimation approach for linear systems was developed in the framework of moving horizon estimation (MHE) which was extended to nonlinear systems in [10]. The distributed MHE (DMHE) schemes developed in [9, 10] were also based on consensus algorithms which require the use of the entire system model in each individual MHEs. Along this line of work, in [51, 52], DMHE schemes based on subsystem models were developed for both linear and nonlinear systems. Since the above DMHE schemes were developed based on the classical centralized MHE

[45, 46, 47], they maintain the advantages of MHE including the ability to handle nonlinearities, constraints and optimality considerations explicitly. However, as in the centralized MHE, the convergence of the above DMHE schemes to the actual system state requires a reliable approximation of the arrival cost. Even though different approaches including the extended Kalman filtering [53], the unscented Kalman filtering [54] and particle filters [55, 56] have been developed to approximate the arrival cost, it is in general a difficult task to determine the arrival cost for constrained nonlinear systems. Also, when there are bounded measurement noise and process disturbances, it is in general difficult to ensure the boundedness of the estimation error [47]. Moreover, the convergence rates of the estimates given by the above DMHE schemes to the actual system states are not tunable and is not favorable from an output feedback control point of view.

1.3 Thesis outline and contributions

This thesis is organized as follows:

In Chapter 2, a robust DMHE design for a class of nonlinear systems with bounded output measurement noise and process disturbances is presented. In this DMHE, each subsystem MHE communicates with subsystems that it interacts with every sampling time. In the design of each subsystem MHE, an auxiliary deterministic nonlinear observer is taken advantage of to calculate a confidence region that contains the actual system state every sampling time. The subsystem MHE is only allowed to optimize its state estimate within the confidence region. This strategy was demonstrated to guarantee the convergence and ultimate boundedness properties of the estimation error.

In Chapter 3, two algorithms are proposed to reduce the number of communications between subsystem based on the DMHE framework developed in Chapter 2. In particular, event-triggered approach is adopted to reduce the number of communication between subsystems. Specifically, in the first proposed algorithm, a subsystem sends out its current information when a triggering condition based on the difference between the current state estimate and a previously transmitted state estimate is satisfied; in the second proposed algorithm, the transmission of information from a subsystem to other subsystems is triggered by the difference between the current measurement of the output and its derivatives and a previously transmitted measurement of the output and its derivatives. Because of the triggered communication, a subsystem may not have the latest information of the other subsystems. The application to a chemical process illustrates the effectiveness of the proposed approaches in reducing the number of communications between the subsystems while maintaining the estimation performance.

In Chapter 4, the DMHE developed in Chapter 2 is extended to handle time-varying communication delays. In particular, an open-loop state predictor is designed for each subsystem to provide predictions of unavailable subsystem states. In the design of each predictor, the centralized system model is used. Based on the state predictions, an auxiliary nonlinear observer is used to generate a reference subsystem state estimate for each subsystem every sampling time. Based on the reference subsystem state estimate as well as the local output measurement, a confidence region is constructed for the actual state of a subsystem. A subsystem MHE is only allowed to optimize its state estimate within the corresponding confidence region at a sampling time. The proposed DMHE is proved to give decreasing and ultimately bounded estimation errors under the assumption that there is an upper bound on the time-varying delay. The theoretical results are illustrated via the application to a reactor-separator chemical process.

Chapter 5 summarizes the main results of this thesis and discusses future research directions.

Chapter 2

Distributed Moving Horizon State Estimation for Nonlinear Systems with Bounded Uncertainties*

2.1 Introduction

In this chapter, we propose a DMHE scheme for a class of nonlinear systems with bounded output measurement noise and process disturbances. Specifically, we consider a class of nonlinear systems that are composed of several subsystems and the subsystems interact with each other via their subsystem states. First, a distributed estimation algorithm is designed which specifies the information exchange protocol between the subsystems and the implementation strategy of the DMHE. Subsequently, a local MHE scheme is design for the each subsystem. In the design of each subsystem MHE, an auxiliary nonlinear deterministic observer that can asymptotically track the corresponding nominal subsystem state when the subsystem interactions are absent is taken advantage of. For each subsystem, the nonlinear deterministic observer together with an error correction term is used to calculate confidence regions for the subsystem states every sampling time. Within the confidence regions, the subsystem MHE is allowed to optimize its estimates. The proposed DMHE scheme is proved to give bounded estimation errors. It is also possible to tune the convergence rate of the state estimate given by the DMHE to the actual system state. The applicability and effectiveness of the proposed DMHE are illustrated via the application to a reactor-separator process example.

*This chapter is a revised version of "J. Zhang and J. Liu, Distributed moving horizon state estimation for nonlinear systems with bounded uncertainties. Journal of Process Control, 23:1281-1295, 2013."

2.2 Notation

Throughout this thesis, we operator $|\cdot|$ denotes Euclidean norm of a scalar or a vector while $|\cdot|_Q^2$ indicates the square of the weighted Euclidean norm of a vector, defined as $|x|_Q^2 = x^T Q x$ where Q is a positive definite square matrix. A function $f(x)$ is said to be locally Lipschitz with respect to its argument x if there exists a positive constant L_f^x such that $|f(x') - f(x'')| \leq L_f^x |x' - x''|$ for all x' and x'' in a given region of x and L_f^x is the associated Lipschitz constant. A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and satisfies $\alpha(0) = 0$. A function $\beta(r, s)$ is said to be a class $\mathcal{KL}$ function if for each fixed s, $\beta(r, s)$ belongs to class $\mathcal{K}$ with respect to r, and for each fixed r, it is deceasing with respect to s, and $\beta(r, s) \to 0$ as $s \to \infty$. The symbol diag(v) denotes a diagonal matrix whose diagonal elements are the elements of vector v. The symbol '\' denotes set subtraction such that $\mathbb{A} \setminus \mathbb{B} := \{x \in \mathbb{A}, x \notin \mathbb{B}\}$. The superscript $^{(s)}$ denotes the s-th order time derivative of a function. The matrix (or vector) A^+ denotes the pseudoinverse of a matrix (or vector) A. The set $\mathbb{I} = \{1, \ldots, m\}$.

2.3 System description

Throughout this thesis, we consider a class of nonlinear systems composed of m interconnected subsystems where the i-th subsystem can be described by the following state-space model:

$$\begin{aligned} \dot{x}_i(t) &= f_i(x_i(t), w_i(t)) + \tilde{f}_i(X_i(t)) \\ y_i(t) &= h_i(x_i) + v_i(t) \end{aligned} \tag{2.1}$$

where $i \in \mathbb{I}$, $x_i(t) \in \mathbb{R}^{n_{x_i}}$ denotes the vector of state variables of subsystem i, $w_i(t) \in \mathbb{R}^{n_{w_i}}$ denotes disturbances associated with subsystem i, and the vector function f_i characterizes the dependence of the dynamics of x_i on itself and the associated disturbances. The vector function $\tilde{f}_i$ characterizes the interactions between subsystem i and other subsystems. The state vector $X_i(t) \in \mathbb{R}^{n_{X_i}}$ denotes the vector of states that involved in characterizing the interactions. The vector $y_i \in \mathbb{R}^{n_{y_i}}$ is the measured output of subsystem i and $v_i \in \mathbb{R}^{n_{v_i}}$ is the measurement noise vector. This system will also be used in Chapter 3 and Chapter 4.

In the following discussion, we use $\mathbb{I}_i \subset \mathbb{I}$, $i \in \mathbb{I}$, to denote the set of subsystem indices whose corresponding subsystem states are involved in X_i. For example, if X_1 contains states of subsystem 1, subsystem 3 and subsystem 4, then $\mathbb{I}_1 = \{1,\ 3,\ 4\}$. It is assumed that the sets $\mathbb{I}_i$, $i \in \mathbb{I}$, are known.

It is assumed that the subsystem states x_i, $i \in \mathbb{I}$, satisfy the constraint:

$$x_i \in \mathbb{X}_i \tag{2.2}$$

where $\mathbb{X}_i$, $i \in \mathbb{I}$, are convex compact sets and the system disturbances and measurement noise are bounded such as $w_i \in \mathbb{W}_i$ and $v_i \in \mathbb{V}_i$, $i \in \mathbb{I}$, where

$$\begin{array}{c} \mathbb{W}_i := \{w_i \in \mathbb{R}^{n_{w_i}} : |w_i| \leq \theta_{w_i}\} \\ \mathbb{V}_i := \{v_i \in \mathbb{R}^{n_{v_i}} : |v_i| \leq \theta_{v_i}\} \end{array} \tag{2.3}$$

with θ_{w_i} and θ_{v_i}, $i \in \mathbb{I}$, known positive real numbers. The entire nonlinear system state vector and measured output vector are denoted as x and y which are compositions of the states and outputs of the m subsystems, respectively. That is, $x = [x_1^T \cdots x_i^T \cdots x_m^T]^T \in \mathbb{R}^{n_x}$ and $y = [y_1^T \cdots y_i^T \cdots y_m^T]^T \in \mathbb{R}^{n_y}$. The entire system can be described as follows:

$$\begin{array}{rcl} \dot{x}(t) & = & f(x(t), w(t)) + \tilde{f}(x(t)) \\ y(t) & = & h(x(t)) + v(t) \end{array} \tag{2.4}$$

where f, $\tilde{f}$, w, h, and v are appropriate compositions of f_i, $\tilde{f}_i$, w_i, h_i, and v_i, $i \in \mathbb{I}$, respectively.

The outputs of the m subsystems, y_i, $i \in \mathbb{I}$, are assumed to be sampled synchronously and periodically at time instants $\{t_{k \geq 0}\}$ such that $t_k = t_0 + k\Delta$ with $t_0 = 0$ the initial time, Δ a fixed sampling time interval and k positive integers. For each subsystem, a state estimator (observer) will be designed in the framework of MHE to estimate its state. It is assumed that the estimator associated with subsystem i has direct access to the measurements of y_i and can communicate with the other subsystems when necessary to exchange their subsystem output measurements and state estimates.

Remark 1 *In order to illustrate the system model representation, consider the following system with three one-dimensional subsystems:*

$$\begin{array}{rclcl} \dot{x}_1 & = & -x_1 + x_1 x_3 & = & f_1(x_1) + \tilde{f}_1(X_1) \\ \dot{x}_2 & = & -0.5x_2 + x_2 x_1 + x_3^2 & = & f_2(x_2) + \tilde{f}_2(X_2) \\ \dot{x}_3 & = & -x_3 + x_1 x_2 + 0.1x_2^3 & = & f_3(x_3) + \tilde{f}_3(X_3) \end{array}$$

In this example, $X_1 = [x_1,\ x_3]^T$ with $\mathbb{I}_1 = \{1,\ 3\}$, $X_2 = [x_1,\ x_2,\ x_3]^T$ with $\mathbb{I}_2 = \{1,\ 2,\ 3\}$ and $X_3 = [x_1,\ x_2]^T$ with $\mathbb{I}_3 = \{1,\ 2\}$. Note that in order to simplify the discussion but without loss of generality, inputs of the system are not considered in (2.4).

2.4 Nonlinear observers

Throughout this thesis, an auxiliary local nonlinear deterministic observer for each subsystem will be taken advantage of to calculate a confidence region for the actual system state every sampling time. In the context of nonlinear systems, there are extensive studies on nonlinear deterministic observers focusing on the design of centralized observers [57, 58, 31, 30, 29, 35, 59, 60, 37] with many successful applications to different areas including the control and monitoring of nonlinear chemical processes [42, 43, 32, 33, 34, 61, 44]. One important class of nonlinear observers is the so-called 'high-gain' observers [29, 30, 31, 39] which allow for effective separation principles in output feedback control designs. However, little attention has been paid to the design of nonlinear decentralized or distributed deterministic observers. Taking this fact into account, we assume that there exists a nonlinear deterministic observer for subsystem i, $i \in \mathbb{I}$, of the following form:

$$\dot{z}_i(t) = F_i(z_i(t), h_i(x_i(t))) \tag{2.5}$$

such that if $\tilde{f}_i(X_i(t)) \equiv 0$, $w_i(t) \equiv 0$ for all t, then z_i asymptotically approaches x_i for all the states $x_i \in \mathbb{X}_i$. This assumption implies that if $\tilde{f}_i(X_i(t)) \equiv 0$, $w_i(t) \equiv 0$ for all t, there exists a $\mathcal{KL}$ function β_i such that:

$$|z_i(t) - x_i(t)| \leq \beta_i(|z_i(0) - x_i(0)|, t) \tag{2.6}$$

where $z_i(0)$ and $x_i(0)$ are the initial states of the observer and the subsystem. The above observability assumption also implies that [62]:

$$\text{rank}(O_i(x_i)) = n_{x_i} \tag{2.7}$$

with $O_i(x_i) = \dfrac{d\Phi_i(x_i)}{dx_i}$ for all $x_i \in X_i$. Note that the convergence property of the nonlinear observer (2.5) is obtained based on continuous noise-free output measurements. We also assume that F_i, $i \in \mathbb{I}$, are locally Lipschitz functions. It is further assumed that the entire system of Eq. (2.4) is locally observable which essentially implies that the interactions between the subsystems do not damage the collective observability of the subsystems.

Note that in the above assumption of the nonlinear observers, the interactions between the subsystems are assumed to be absent (i.e., $\tilde{f}_i(X_i(t)) \equiv 0$ for all t). Note also that the convergence properties of the nonlinear observers are obtained based on continuous noise-free output feedback. In the proposed DMHE discussed in the next section, we will discuss in details on how to take advantage of observer (2.5) and to compensate for the interactions between the subsystems using information exchanged between subsystems.

2.5 The DMHE scheme

A schematic of the proposed DMHE design which includes m local MHEs for the nonlinear system of Eq. (2.4) is shown in Fig. 2.1. In the proposed DMHE design, each subsystem is

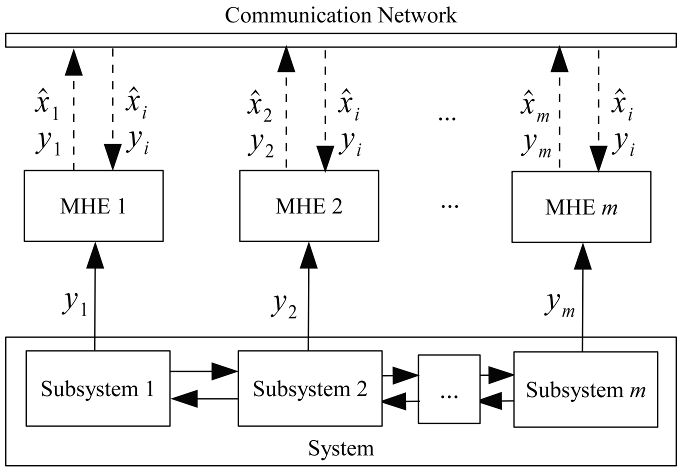


Figure 2.1: The proposed DMHE design.

associated with an MHE which is evaluated every sampling time. We also assume that a local MHE has immediate access to the output measurements of its associated subsystem and can communicate with the other subsystems to exchange their subsystem output measurements and state estimates which are used to compensate for the interactions between subsystems to improve their state estimates. The MHE associated with subsystem i ($i \in \mathbb{I}$) will be referred to as MHE i. In the remainder of this chapter, we will first introduce the proposed distributed estimation algorithm; subsequently, we present the design of the local MHEs and finally derive sufficient conditions under which the proposed DMHE gives bounded estimation error.

2.5.1 Distributed estimation algorithm

The proposed DMHE uses the following distributed state estimation algorithm to get an estimate of the entire system.

Algorithm 2 *Distributed state estimation algorithm*

1. *At* $t_0 = 0$*, all the MHEs are initialized with the initial subsystem state guesses* $\hat{x}_i(0)$*,* $i \in \mathbb{I}$*, and the actual subsystem output measurements* $y_i(0)$*,* $i \in \mathbb{I}$*.*
2. *At* $t_k > 0$*, carry out the following steps:*
   - 2.1. *Each MHE receives the output measurement of the subsystem that it is associated with; that is, MHE* i *receives* $y_i(t_k)$*.*
   - 2.2. *Each MHE requests and receives the output measurements and subsystem state estimates of the previous time instant from subsystems that directly affect its*

dynamics; that is, MHE i requests and receives $y_l(t_{k-1})$ and $\hat{x}_l(t_{k-1})$ (which denotes the state estimate of subsystem l at t_{k-1}) for all $l \in \mathbb{I}_i$.

2.3. Based on both the local measurement and information from other subsystems, each MHE calculates the estimate of its subsystem's state; that is, MHE i calculates $\hat{x}_i(t_k)$. The estimate of the entire system state is $\hat{x}(t_k) = [\hat{x}_1(t_k)^T \dots \hat{x}_m(t_k)^T]^T$.

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

From Algorithm 2, it can be seen that it is a non-iterative algorithm. At a sampling time, the MHEs are evaluated only once in parallel. This implementation compensates for the interactions between subsystems based on the state estimates and output measurements at the previous time instant. In addition, Algorithm 2 does not require an all-to-all communication between the MHEs. From Step 2.2, it can be seen that an MHE only communicates with subsystems that it interacts directly. For example, if the subsystems are connected in series, an MHE only has to receive information from its directly connected upstream subsystem and to send information to its directly connected downstream subsystem.

Remark 3 *Note that an iterative implementation algorithm may be used for the proposed DMHE design based on the current output measurements and the state estimates obtained at the previous iteration of the current sampling time. In this case, it is possible to achieve the convergence property of the state estimates by redesigning the local MHEs accordingly. An iterative implementation algorithm may lead to improved distributed estimation performance if the iterations converge to the global optimum which, however, is not ensured for general nonlinear systems due to the non-convexity of the optimization problems. Moreover, when an iterative implementation algorithm is used, it may significantly increase the computational complexity of the proposed DMHE design.*

2.5.2 Local MHE design

In the design of a local MHE, the subsystem model of Eq. (2.1), the corresponding nonlinear deterministic observer of Eq. (2.5) together with the information received from other subsystems are used. A confidence region that contains the actual subsystem state will be calculated every sampling time taking into account the boundedness of the measurement noise and process disturbances. The local MHE is only allowed to optimize the subsystem state estimate within the confidence region.

Specifically, the proposed design of MHE i at time instant t_k is formulated as follows:

$$\min_{\tilde{x}_i(t_{k-N}),\ldots,\tilde{x}_i(t_k)} \sum_{q=k-N}^{k-1} |w_i(t_q)|^2_{Q_i^{-1}} + \sum_{q=k-N}^{k} |v_i(t_q)|^2_{R_i^{-1}} + V_i(t_{k-N}) \tag{2.8a}$$

$$\text{s.t. } \dot{\tilde{x}}_i(t) = f_i(\tilde{x}_i(t), w_i(t_q)) + \tilde{f}_i(\hat{X}_i(t_q)),\ t \in [t_q, t_{q+1}],\ q = k-N, \ldots, k-1 \tag{2.8b}$$

$$v_i(t_q) = y_i(t_q) - h_i(\tilde{x}_i(t_q)),\ q = k-N, \ldots, k \tag{2.8c}$$

$$w_i(t_q) \in \mathbb{W}_i,\ v_i(t_q) \in \mathbb{V}_i,\ \tilde{x}_i(t_q) \in \mathbb{X}_i,\ q = k-N, \ldots, k \tag{2.8d}$$

$$\dot{z}_i(t) = F_i(z_i(t), y_i(t_{k-1})) + \tilde{f}_i(\hat{X}_i(t_{k-1})) + \sum_{l \in \mathbb{I}_i} K_{i,l}(\hat{x}_l)(y_l(t_{k-1}) - h_l(\hat{x}_l(t_{k-1}))) \tag{2.8e}$$

$$z_i(t_{k-1}) = \hat{x}_i(t_{k-1}) \tag{2.8f}$$

$$|\tilde{x}_i(t_k) - z_i(t_k)| \leq \kappa_i |y_i(t_k) - h_i(z_i(t_k))| \tag{2.8g}$$

where N is the estimation horizon, Q_i and R_i are the covariance matrices of w_i and v_i respectively, $V_i(t_{k-N})$ denotes the arrival cost which summarizes past information up to t_{k-N}, $\tilde{x}_i$ is the predicted x_i in the above optimization problem, $\hat{x}_i$ is the optimal estimate of x_i at previous time instants, $K_{i,l}$ for $l \in \mathbb{I}_i$ are gain matrices which are functions of $\hat{x}_l$, and κ_i is a positive constant. The roles of $K_{i,l}$ and κ_i will be made clear in the following discussion.

Once problem (2.8) is solved, an optimal trajectory of the system states, $\tilde{x}_i^*(t_{k-N}), \ldots, \tilde{x}_i^*(t_k)$, is obtained. The last element $\tilde{x}_i^*(t_k)$ is used as the optimal estimate of the state of subsystem i at t_k and is denoted as $\hat{x}_i(t_k)$. That is,

$$\hat{x}_i(t_k) = \tilde{x}_i^*(t_k). \tag{2.9}$$

Note that in the optimization problem (2.8), w_i and v_i are assumed to be piece-wise constant variables with sampling time Δ to ensure that (2.8) is a finite dimensional optimization problem.

In optimization problem (2.8), constraint (2.8a) is the cost function that needs to be minimized. The arrival cost $V_i(t_{k-N})$ summarizes the past information that is not covered in the estimation horizon. Constraint (2.8b) is the model of subsystem i. Because only state estimates at the sampling times are available, $\tilde{f}_i(\hat{X}_i(t_q))$ is used to approximate $\tilde{f}_i(X_i(t))$ from t_q to t_{q+1}. This also implies that each MHE should be able to store the previously received information from other MHEs. Constraint (2.8d) are bounds on the disturbances, noise and subsystem state.

Constraints (2.8e)-(2.8g) are used to calculate a confidence region that contains the actual subsystem state based on the deterministic nonlinear observer, a correction term and the information received from other subsystems. The estimate of the current subsystem state is only allowed to be optimized within the confidence region. Specifically, constraint (2.8e) is an augmented nonlinear observer taking into account explicitly the interactions between subsystem i and the other subsystems. The first term on the right-hand-side of (2.8e) comes from observer (2.5); the second term on the right-hand-side of (2.8e) is based on the interaction model and the previous state $\hat{X}_i(t_{k-1})$ which is the latest available information available to subsystem i; and the last term is used to correct the errors in the interaction model. The gain $K_{i,l}$ depends on $\hat{x}_l$ and is calculated at each sampling time as follows:

$$K_{i,l} = \frac{\partial \tilde{f}_i}{\partial x_l} \left(\frac{\partial h_l}{\partial x_l} \right)^+ \bigg|_{\hat{x}_l(t_{k-1})} \tag{2.10}$$

for $l \in \mathbb{I}_i$ and $i \in 1, \ldots, m$. The above calculation of $K_{i,l}$ implies that the estimation error caused by the interaction is compensated by its linearized approximation. This idea will be made explicit in Section 2.5.3.

Constraint (2.8g) explicitly defines the confidence region based on the parameter κ_i and the estimate given by the nonlinear observer (2.8e) as well as the actual output measurement. The parameter κ_i depends on the properties of the system. Conditions that the value of κ_i ($i \in \mathbb{I}$) needs to satisfy will be derived in Section 2.5.3. It will also be proven in Section 2.5.3 that the proposed approach leads to bounded estimation errors.

Remark 4 *In the proposed DMHE design, the interactions between subsystems (i.e., $\tilde{f}_i(X_i)$, $i \in \mathbb{I}$) are compensated for based on the interaction models as well as subsystem state estimates (i.e., $\hat{X}_i$). The error caused by the difference between $\hat{X}_i$ and X_i is further compensated for using the correction term $\sum_{l \in \mathbb{I}_i} K_{i,l}(y_l(t_{k-1} - h_l(\hat{x}_l(t_{k-1})))$ with $K_{i,l}$, $l \in \mathbb{I}_i$, determined following (2.10). The correction term with $K_{i,l}$ essentially compensates for the linear part of the error dynamics caused by the difference between $\hat{X}_i$ and X_i and neglects the higher order dynamics which will be made clear in the proof of Proposition 5 in Section 2.5.3. In many applications, a first order correction term like the one used in the present work is sufficient to achieve desired estimation performance. Please see Section 2.6 for the application of the proposed approach to a reactor-separator chemical process. If in an application it is necessary to compensate for the higher order error dynamics in system interactions, the proposed approach can be extended in a straightforward manner.*

2.5.3 Stability analysis

In this section, we study the robustness and stability properties of the proposed DMHE. Specifically, we first investigate the boundedness of the estimation error given by the nonlinear observer (2.8e) with $K_{i,l}$ determined following (2.10) taking into account measurement noise and process disturbances. Following this, we state the stability and ultimate boundedness of the estimation error of the proposed DMHE. Sufficient conditions will be provided.

Proposition 5 *Consider the nonlinear observer of Eq. (2.8e) for subsystem i, $i \in \mathbb{I}$, with initial condition $z_i(t_k) = \hat{x}_i(t_k)$ and output measurement $y_i(t_k)$. If $K_{i,l}$ for $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$ are determined as in (2.10) and $K_{i,l}$ are bounded, then the deviation of the observer state z_i in one sampling time Δ (i.e., at t_{k+1}) from the actual subsystem state x_i is bounded for all $x_i \in \mathbb{X}_i$, $i \in \mathbb{I}$, as follows:*

$$|e_{z,i}(t_{k+1})| \leq \beta_i(|e_{z,i}(t_k)|, \Delta) + \gamma_i(\Delta) + \sum_{l\in\mathbb{I}_i} L_{i,l}\Delta|e_{z,l}(t_k)|^2 \tag{2.11}$$

where $e_{z,i} = z_i - x_i$, $i \in \mathbb{I}$, and $\gamma_i(\tau) = L_{F_i}^{y_i} L_{h_i} M_i \tau^2/2 + L_{F_i}^{y_i}\theta_{v_i}\tau + L_{f_i}^{w_i}\theta_{w_i}\tau + \sum_{l\in\mathbb{I}_i} M_{K_{i,l}}\theta_{v_l}\tau + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \tau^2/2$ and $L_{i,l} = H_i^{\tilde{f}_i} + M_{K_{i,l}} H_l^{h_l}$ with $L_{F_i}^{y_i}$, L_{h_i}, $L_{f_i}^{w_i}$, and $L_{\tilde{f}_i}^{x_l}$ being the Lipschitz constants of F_i with respect to y_i, h_i with respect to x_i, f_i with respect to w_i, and $\tilde{f}_i$ with respect to x_l, respectively, and M_i, $M_{K_{i,l}}$, $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$, being constants that bound $\dot{x}_i$ in $\mathbb{X}_i$ and $K_{i,l}$ in $\mathbb{X}_l$, respectively, and $H_i^{\tilde{f}_i}$, $H_l^{h_l}$ being positive constants that associated with the Taylor expansions of $\tilde{f}_i$ and h_l.

Proof: We consider the local MHE i of Eq. (2.8) ($i \in \mathbb{I}$) and define $e_{z,i} = z_i - x_i$ where z_i denotes the trajectory of observer (2.8e) and x_i is the state trajectory of the actual subsystem of Eq. (2.1). In this proof, we consider the time interval from $t = t_k$ to $t = t_{k+1}$ and the initial condition $z_i(t_k) = \hat{x}_i(t_k)$. The derivative of $e_{z,i}$ is evaluated as follows:

$$\begin{aligned} \dot{e}_{z,i}(t) \;=\;& F_i(z_i(t), y_i(t_k)) - f_i(x_i(t), w_i(t)) \\ & + \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) + \textstyle\sum_{l\in\mathbb{I}_i} K_{i,l}(\hat{x}_l)(y_l(t_k) - h_l(\hat{x}_l(t_k))). \end{aligned} \tag{2.12}$$

From the Lipschitz properties of F_i, f_i and h_i, the fact that $y_i(t_k) = h_i(x_i(t_k)) + v_i(t_k)$, and $|v_i(t_k)| \leq \theta_{v_i}$, $|w_i| \leq \theta_{w_i}$, the following inequality can be obtained from (2.12):

$$\begin{aligned} |\dot{e}_{z,i}(t)| \;\leq\;& |F_i(z_i(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i}|x_i(t) - x_i(t_k)| + L_{F_i}^{y_i}\theta_{v_i} + L_{f_i}^{w_i}\theta_{w_i} \\ & + \left|\tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) + \textstyle\sum_{l\in\mathbb{I}_i} K_{i,l}(\hat{x}_l)(y_l(t_k) - h_l(\hat{x}_l(t_k)))\right| \end{aligned} \tag{2.13}$$

where $L_{F_i}^{y_i}$, L_{h_i} and $L_{f_i}^{w_i}$ are the Lipschitz constants associated with F_i, h_i and f_i, respectively.

Using Taylor series expansion, the following inequalities can be obtained:

$$\begin{aligned} \tilde{f}_i(X_i(t_k)) &= \tilde{f}_i(\hat{X}_i(t_k)) + \sum_{l\in\mathbb{I}_i} \frac{\partial \tilde{f}_i}{\partial x_l}(x_l(t_k) - \hat{x}_l(t_k)) + H.O.T_i^{\tilde{f}_i}, \\ h_l(x_l(t_k)) &= h_l(\hat{x}_l(t_k)) + \frac{\partial h_l}{\partial x_l}(x_l(t_k) - \hat{x}_l(t_k)) + H.O.T_l^{h_l} \end{aligned} \tag{2.14}$$

where $H.O.T_i^{\tilde{f}_i}$ and $H.O.T_l^{h_l}$ are high order terms associated with the expansions of $\tilde{f}_i$ and h_l. These high order terms satisfy the following constraints:

$$H.O.T_i^{\tilde{f}_i} \le H_i^{\tilde{f}_i}|X_i(t_k) - \hat{X}_i(t_k)|^2, \quad H.O.T_l^{h_l} \le H_l^{h_l}|x_l(t_k) - \hat{x}_l(t_k)|^2 \tag{2.15}$$

for all $x_i \in \mathbb{X}_i$ $(i \in \mathbb{I})$ with $H_i^{\tilde{f}_i}$, $i \in \mathbb{I}$, and $H_l^{h_l}$, $l = 1, \ldots, m$, are positive constants.

Let us define $A_i(t_k) = \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t_k)) + \sum_{l\in\mathbb{I}_i} K_{i,l}(\hat{x}_l)(h_l(x_l(t_k)) - h_l(\hat{x}_l(t_k)))$. From (2.14), the following equation can be written:

$$A_i(t_k) = \sum_{l\in\mathbb{I}_i} \left(-\frac{\partial \tilde{f}_i}{\partial \hat{x}_l}(x_l(t_k) - \hat{x}_l(t_k)) + K_{i,l}\frac{\partial h_l}{\partial \hat{x}_l}(x_l(t_k) - \hat{x}_l(t_k)) \right) + \sum_{l\in\mathbb{I}_i} K_{i,l} H.O.T_l^{h_l} - H.O.T_i^{\tilde{f}_i}. \tag{2.16}$$

If $K_{i,l}$ is determined following (2.10), from (2.15) and (2.16), it can be obtained that:

$$|A_i(t_k)| \le H_i^{\tilde{f}_i}|X_i(t_k) - \hat{X}_i(t_k)|^2 + \sum_{l\in\mathbb{I}_i} K_{i,l} H_l^{h_l}|x_l(t_k) - \hat{x}_l(t_k)|^2. \tag{2.17}$$

From (2.13), using the triangle inequality and taking into account (2.17), $|X_i(t_k) - \hat{X}_i(t_k)|^2 = \sum_{l\in\mathbb{I}_i} |x_l(t_k) - \hat{x}_l(t_k)|^2$ as well as $y_l(t_k) = h_l(x_l(t_k)) + v_l(t_k)$, and from the Lipschitz property of $\tilde{f}_i$ with respect to x_l $(l \in \mathbb{I}_i)$, the following inequality can be obtained:

$$\begin{aligned} |\dot{e}_{z,i}(t)| \le\ & |F_i(z_i(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i}|x_i(t) - x_i(t_k)| + L_{F_i}^{y_i}\theta_{v_i} + L_{f_i}^{w_i}\theta_{w_i} \\ & + \sum_{l\in\mathbb{I}_i}\left(H_i^{\tilde{f}_i} + K_{i,l}H_l^{h_l}\right)|e_{z,l}(t_k)|^2 + \sum_{l\in\mathbb{I}_i} K_{i,l}\theta_{v_l} + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l}|x_l(t) - x_l(t_k)| \end{aligned} \tag{2.18}$$

where $e_{z,l}(t_k) = \hat{x}_l(t_k) - x_l(t_k)$. Taking into account that $z_i(t_k) = \hat{x}_i(t_k)$ and condition (2.6) and integrating (2.18) from $t = t_k$ to $t = t_{k+1}$, the following inequality can be obtained:

$$\begin{aligned} |e_{z,i}(t_{k+1})| \le\ & \beta_i(|e_{z,i}(t_k)|, \Delta) + L_{F_i}^{y_i} L_{h_i} M_i \Delta^2/2 + L_{F_i}^{y_i}\theta_{v_i}\Delta + L_{f_i}^{w_i}\theta_{w_i}\Delta \\ & + \sum_{l\in\mathbb{I}_i}\left(H_i^{\tilde{f}_i} + M_{K_{i,l}}H_l^{h_l}\right)|e_{z,l}(t_k)|^2\Delta + \sum_{l\in\mathbb{I}_i} M_{K_{i,l}}\theta_{v_l}\Delta + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \Delta^2/2 \end{aligned} \tag{2.19}$$

where M_i, $i \in \mathbb{I}$, are constants that bounds $\dot{x}_i$ in $\mathbb{X}_i$ (i.e., $|\dot{x}_i| \le M_i$) and $M_{K_{i,l}}$, $l \in \mathbb{I}_i$, are constants that bounds $K_{i,l}$ in $\mathbb{X}_l$ (i.e., $|K_{i,l}| \le M_{K_{i,l}}$). If $\gamma_i(\tau)$ and $L_{i,l}$ are defined as in Proposition 5, (2.19) can be written in the form of Eq. (2.11). This proves Proposition 5. □

From the formulation of the local MHE of Eq. (2.8), it can be seen that observer (2.8e) is operated in open-loop since only information at t_{k-1} is used. The result of Proposition 5 shows that within one sampling time, the estimation error given by the open-loop observer (2.8e) is bounded and the upper bound depends on the Lipschitz properties of the system, the property of observer (2.5), the sampling time, the uncertainties involved in the system, and the interactions between the subsystems. Theorem 6 below takes advantage of the boundeness property of observer (2.8e) and provides sufficient conditions on the convergence and ultimate boundedness of the estimation error of the proposed DMHE.

Theorem 6 *Consider system (2.4) with the outputs of its subsystems y_i sampled at time instants $\{t_{k\geq 0}\}$. If the proposed DMHE implemented following Algorithm 2 with subsystem MHE designed as in (2.8) based on deterministic nonlinear observers satisfying (2.6) and $K_{i,l}$ determined following (2.10) that are bounded, and if there exist concave functions $g_i(\cdot)$, $i \in \mathbb{I}$, such that:*

$$g_i(|e_i|) \geq \beta_i(|e_i|, \Delta) \tag{2.20}$$

for all $|e_i| \leq d_i$ and if there exist vector constants $d_{s,i}$, d_i such that $0 \leq d_{s,i} \leq d_i$, and positive constants $a_i \geq 1$, $b_i > 0$, and $\epsilon_i > 0$, such that:

$$d_{s,i} - a_i \left(g_i(d_{s,i}) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta d_l^2 \right) - b_i \theta_{v_i} \geq \epsilon_i \tag{2.21}$$

for all $i \in \mathbb{I}$, and if κ_i for all $i \in \mathbb{I}$, are picked as follows:

$$0 \leq \kappa_i \leq \min\{(a_i - 1)/L_{h_i}, b_i\}, \tag{2.22}$$

then the estimation error $|e_i| = |\hat{x}_i - x_i|$ ($i \in \mathbb{I}$) is a decreasing sequence if $|e_i(0)| \leq d_i$ for all $i \in \mathbb{I}$ and is ultimately bounded as follows:

$$\lim_{t \to \infty} \sup |e_i(t)| \leq d_{i,\min} \tag{2.23}$$

for $i \in \mathbb{I}$ with $d_{i,\min} = \max\{|e_i(t + \Delta)| : |e_i(t)| \leq d_{s,i}\}$ for all $e_i(0) \leq d_i$ and $x_i \in \mathbb{X}_i$. This also implies that the entire system state estimation error is ultimately bounded.

Proof: We prove that the evolution of the estimation error of each subsystem state $|e_i| = |\hat{x}_i - x_i|$, $i \in \mathbb{I}$, under the proposed DMPC with the local MHE of Eq. (2.8) is a decreasing sequence and is ultimately bounded in a small region. The decrease and ultimate boundedness of subsystem estimation errors imply the decrease and ultimate boundedness of the entire system state estimation error. Specifically, we first focus on MHE i, $i \in \mathbb{I}$, for the time interval from t_k to t_{k+1} and then extend to the general case.

From constraint (2.8g) for MHE i, it can be written that:

$$|\hat{x}_i(t_{k+1}) - z_i(t_{k+1})| \leq \kappa_i |y_i(t_{k+1}) - h_i(z_i(t_{k+1}))|. \tag{2.24}$$

From the Lipschitz property of h_i, the fact that $y_i = h_i(x_i) + v_i$ and $|v_i| \leq \theta_{v_i}$, it is obtained that:

$$|\hat{x}_i(t_{k+1}) - z_i(t_{k+1})| \leq \kappa_i L_{h_i} |x_i(t_{k+1}) - z_i(t_{k+1})| + \kappa_i \theta_{v_i} \tag{2.25}$$

where L_{h_i} is the Lipschitz constant of h_i as defined in Proposition 5. Using the triangle inequality $|\hat{x}_i - x_i| \leq |\hat{x}_i - z_i| + |z_i - x_i|$, it is obtained from (2.25) that:

$$|\hat{x}_i(t_{k+1}) - x_i(t_{k+1})| \leq (1 + \kappa_i L_{h_i}) |x_i(t_{k+1}) - z_i(t_{k+1})| + \kappa_i \theta_{v_i}. \tag{2.26}$$

From Proposition 5 and (2.26), and noticing that $e_i(t_k) = e_{z,i}(t_k)$, the following inequality can be obtained:

$$|e_i(t_{k+1})| \leq (1 + \kappa_i L_{h_i}) \left(\beta_i(|e_i(t_k)|, \Delta) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta |e_l(t_k)|^2 \right) + \kappa_i \theta_{v_i}. \tag{2.27}$$

If condition (2.20) is satisfied, from (2.27), it can be obtained that:

$$|e_i(t_{k+1})| \leq (1 + \kappa_i L_{h_i}) \left(g_i(|e_i(t_k)|) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta |e_l(t_k)|^2 \right) + \kappa_i \theta_{v_i}. \tag{2.28}$$

If there exists $d_{s,i}$ satisfy (2.21) and κ_i is picked following (2.22), then (2.21) holds for all $d_{s,i} \leq |e| \leq d_i$ taking into account that $g_i(\cdot)$ is a concave function; that is:

$$|e_i| - (1 + \kappa_i L_{h_i}) \left(g_i(|e_i|) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta |e_l|^2 \right) - \kappa_i \theta_{v_i} \geq \epsilon_i \tag{2.29}$$

for all $d_{s,i} \leq |e_i| \leq d_i$ and $|e_l| \leq d_l$ ($l \in \mathbb{I}_i$). From (2.28) and (2.29), it can be obtained that:

$$|e_i(t_{k+1})| \leq |e_i(t_k)| - \epsilon_i \tag{2.30}$$

for all $d_{s,i} \leq |e_i| \leq d_i$. If $|e_i| \geq d_{s,i}$ for all the time from 0 to t_k, using (2.30) recursively, it can be obtained that:

$$|e_i(t_k)| \leq |e_i(0)| - k\epsilon_i \tag{2.31}$$

for all $d_{s,i} \leq |e_i(t_k)| \leq d_i$. This implies that $|e_i|$ decreases every sampling time and will become smaller than $d_{s,i}$ in finite steps. Once $|e_{s,i}| < d_{s,i}$, it will remain to satisfy $|e_i(t)| \leq$

$d_{i,\min}$ which is ensured by the definition of $d_{i,\min}$; that is, $\lim_{t\to\infty} \sup |e_i(t)| \leq d_{i,\min}$. Note that the above proof holds for all $i \in \mathbb{I}$.

The ultimate boundedness of subsystem state estimation error implies the ultimate boundedness of the entire system state estimation error. This can be seen from the inequality $|e| \leq \sum_{i=1}^{m} |e_i|$ which implies that:

$$\lim_{t\to\infty} \sup |e| \leq \sum_{i=1}^{m} d_{i,\min}. \tag{2.32}$$

This proves Theorem 6. □

Remark 7 *The purpose of the introduction of* g_i *as in condition (2.20) is to ensure that when condition (2.21) is satisfied for* $|e_i| = d_{s,i}$*, it is also satisfied for* $d_{s,i} \leq |e_i| \leq d$ *which implies that if* $|e_i| \geq d_{s,i}$*,* $|e_i|$ *will be decreasing. For many of the existing nonlinear observers (e.g., [31, 30, 33, 37]) that provide exponentially convergence rates such that* $|e_i(t)| \leq \lambda_i |e_i(0)| \exp(-\alpha_i t)$ *with* λ_i *and* α_i *positive numbers, condition (2.20) can be easily satisfied.*

Remark 8 *Referring to condition (2.21) in Theorem 6 (or (2.29) in the proof), the term* $g_i(|e_i(t_k)|)$ *denotes an upper bound on the error value after one sampling time (i.e.,* $|e_i(t_{k+1})|$*) if the initial error value is* $|e_i(t_k)|$ *for the nominal subsystem* i *without interactions under continuous output* y_i *feedback; the term* $\gamma_i(\Delta)$ *represents the effects of sampled-and-hold implementation of nonlinear observer (2.5), measurement noise and process disturbances; the term* $\sum_{l\in I_i} L_{i,l}\Delta d_l^2$ *bounds the effect of subsystem interactions; and the term* $\kappa_i\theta_{v_i}$ *represents the uncertainty introduced into condition (2.8g) due to measurement noise. Condition (2.21) essentially requires that the assumed nonlinear observer (2.5) for the nominal system without interactions converges to the actual nominal subsystem state fast enough such that its contribution to the decease of the estimation error dominates the effects caused by other factors that contribute to the increase of the estimation error.*

Remark 9 *Note that Theorem 6 provides a set of sufficient conditions that essentially decouple the error dynamics of each subsystem. Condition (2.21) involves the initial estimation error* d_l*,* $l \in \mathbb{I}_i$*, of the interacting subsystems of subsystem* i*. This set of sufficient conditions requires that the initial estimation errors of the subsystems should be sufficiently small. In other words, the convergence rates of the nonlinear observers of Eq. (2.5) should be high enough to reject the effects of the initial estimation errors.*

Remark 10 *From Theorem 6 as well as Remark 8, it can be seen that the convergence rates of the nonlinear observers of Eq. (2.5) for the nominal subsystems without interactions play an important role in the convergence rate of the estimates of the proposed DMHE to the actual system states. This implies that it is possible to tune the convergence rate of the proposed DMHE by tuning the convergence rates of the nonlinear observers of Eq. (2.5). By examining condition (2.21), it can be found that when we tune the nonlinear observers to increase their convergence rates (i.e.,* $g_i(e_i)$ *decreases), the negative effects caused by the sampled-and-hold implementation (i.e.,* $\gamma_i(\Delta)$*) increase at the same time which implies the increase of the value of* $d_{i,\min}$*. This is because with the increase of the convergence rates of the nonlinear observers, their Lipschitz constants (i.e.,* $L_{F_i}^{y_i}$*) increase.*

In order to overcome the above issue, two approaches may be used. First, a nonlinear observer of Eq. (2.5) with switched convergence rates is used in the design of the MHE of Eq. (2.8) for each subsystem. Specifically, a high convergence rate is adopted when the estimation error is large and a low convergence rate is used when the estimation error is small. By applying this approach, a high convergence rate of each local MHE can be achieved while keeping the value of $d_{i,\min}$ *small [63]. A second approach that may be used to further improve the performance of the first approach is to use/require continuous output measurements for the time period when the high convergence rate is used in the evaluation of the nonlinear observer of Eq. (2.5). This approach can significantly reduce the negative effects caused by the sampled-and-hold implementation of the nonlinear observer in the case of high convergence rate.*

Remark 11 *The proposed DMHE scheme integrates deterministic nonlinear observer design techniques and MHE. It can increase the robustness and reliability of the observer over either deterministic observers or classical MHE as will be demonstrated in Section 2.6 (see also [50, 64]). This is due to constraint (2.8g) in each subsystem MHE design which ensures that the MHE inherits the robustness of the deterministic nonlinear observer. The proposed approach, however, requires more efforts in the initial design stage and in the tuning of the parameters.*

Remark 12 *Note that in this work, we consider a type of bounded model mismatch (i.e., process disturbances). Other types of model mismatches, such as uncertainties in model parameters or model structure, can be considered in a similar fashion as long as the model mismatches are bounded and the auxiliary nonlinear observers are designed following the assumptions in Section 2.4. Note also that even we do not explicitly consider model mis-*

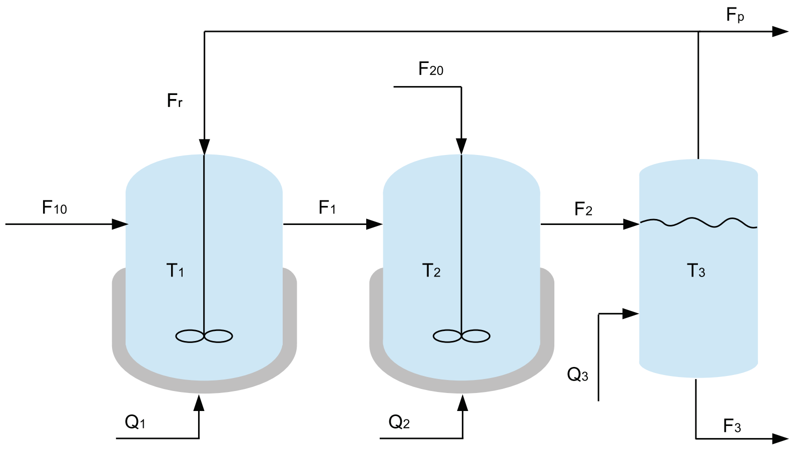


Figure 2.2: Reactor-separator process with a recycle stream.

matches in the interaction model, we would like to clarify that different uncertainties are actually included since 1) estimates of subsystem states are used in the interaction model in the design of the proposed DMHE, and 2) linear correction terms are used to compensate for nonlinear error dynamics.

Remark 13 *The results in Theorem 6 are conservative since the worst case scenario is considered. It might be possible to derive less conservative conditions if real-time (instead of the worst) interactions between the subsystems are considered. However, this will make the sufficient conditions for subsystems coupled together and may complicate the design process of the distributed estimation system.*

2.6 Application to a reactor-separator process

2.6.1 Process description and modeling

In this section, the proposed distributed state estimation approach is applied to a reactor-separator process which includes two connected continuous stirred tank reactors (CSTR) and one flash tank separator as shown in Fig. 2.2. Similar processes have been studied in [65] in the context of networked process control. The feed stream entering the first tank contains pure reactant A at flow rate F_{10} and temperature T_{10}. A is expected to become the product B and there is also a second reaction which converts B to the side product C: $A \rightarrow B$, $B \rightarrow C$. The effluent of CSTR 1 is fed into CSTR 2 at flow rate F_1 and temperature T_1. There is also another flow of pure A that is fed into CSTR 2 at flow rate F_{20} and temperature T_{20}. The same reactions take place in CSTR 2. A portion of the effluent of CSTR 2 is passed through a separator and recycled back to CSTR 1 at recycle flow rate F_r and temperature T_3. Each reactor is equipped with a jacket to provide/remove heat to/from the reactor. Based on standard modeling assumptions and mass and energy balances, nine ordinary differential equations can be obtained to describe the dynamics of the process:

$$\frac{dx_{A1}}{dt} = \frac{F_{10}}{V_1}(x_{A10} - x_{A1}) + \frac{F_r}{V_1}(x_{Ar} - x_{A1}) - k_1 e^{\frac{-E_1}{RT_1}} x_{A1} \tag{2.33a}$$

$$\frac{dx_{B1}}{dt} = \frac{F_{10}}{V_1}(x_{B10} - x_{B1}) + \frac{F_r}{V_1}(x_{Br} - x_{B1}) + k_1 e^{\frac{-E_1}{RT_1}} x_{A1} - k_2 e^{\frac{-E_2}{RT_1}} x_{B1} \tag{2.33b}$$

$$\frac{dT_1}{dt} = \frac{F_{10}}{V_1}(T_{10} - T_1) + \frac{F_r}{V_1}(T_3 - T_1) - \frac{\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_1}} x_{A1} - \frac{\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_1}} x_{B1} + \frac{Q_1}{\rho c_p V_1} \tag{2.33c}$$

$$\frac{dx_{A2}}{dt} = \frac{F_1}{V_2}(x_{A1} - x_{A2}) + \frac{F_{20}}{V_2}(x_{A20} - x_{A2}) - k_1 e^{\frac{-E_1}{RT_2}} x_{A2} \tag{2.33d}$$

$$\frac{dx_{B2}}{dt} = \frac{F_1}{V_2}(x_{B1} - x_{B2}) + \frac{F_{20}}{V_2}(x_{B20} - x_{B2}) + k_1 e^{\frac{-E_1}{RT_2}} x_{A2} - k_2 e^{\frac{-E_2}{RT_2}} x_{B2} \tag{2.33e}$$

$$\frac{dT_2}{dt} = \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_{20}}{V_2}(T_{20} - T_2) - \frac{\Delta H_1}{c_p} k_1 e^{\frac{-E_1}{RT_2}} x_{A2} - \frac{\Delta H_2}{c_p} k_2 e^{\frac{-E_2}{RT_2}} x_{B2} + \frac{Q_2}{\rho c_p V_2} \tag{2.33f}$$

$$\frac{dx_{A3}}{dt} = \frac{F_2}{V_3}(x_{A2} - x_{A3}) - \frac{(F_r + F_p)}{V_3}(x_{Ar} - x_{A3}) \tag{2.33g}$$

$$\frac{dx_{B3}}{dt} = \frac{F_2}{V_3}(x_{B2} - x_{B3}) - \frac{(F_r + F_p)}{V_3}(x_{Br} - x_{B3}) \tag{2.33h}$$

$$\frac{dT_3}{dt} = \frac{F_2}{V_3}(T_2 - T_3) + \frac{Q_3}{\rho c_p V_3} + \frac{(F_r + F_p)}{\rho c_p V_3}(x_{Ar}\Delta H_{vap1} + x_{Br}\Delta H_{vap2} + x_{Cr}\Delta H_{vap3}) \tag{2.33i}$$

It is assumed that there is a negligible amount of reaction taking place in the separator and that the relative volatility for each of the components remains constant within the operating temperature range. The algebraic equations modeling the composition of the overhead stream relative to composition of liquid in the flash tank is described as follows:

$$\begin{aligned} x_{Ar} &= \frac{\alpha_A x_{A3}}{\alpha_A x_{A3} + \alpha_B x_{B3} + \alpha_C x_{C3}} \\ x_{Br} &= \frac{\alpha_B x_{B3}}{\alpha_A x_{A3} + \alpha_B x_{B3} + \alpha_C x_{C3}} \\ x_{Cr} &= \frac{\alpha_C x_{C3}}{\alpha_A x_{A3} + \alpha_B x_{B3} + \alpha_C x_{C3}} \end{aligned} \tag{2.34}$$

The definition of process variables and values of these parameters are given in Table 2.1 and Table 2.2, respectively. Note that the reaction (1) and reaction (2) refer to the reactions $A \rightarrow B$ and $B \rightarrow C$, respectively.

The system is divided into three subsystems with respect to the three vessels in the process, and the states are noted by $x_i = [x_{Ai},\ x_{Bi},\ T_i]^T,\ i = 1,\ 2,\ 3$. For each subsystem,

Table 2.1: Process variables for the reactors.

$x_{A1},\ x_{A2},\ x_{A3}$	mass fractions of A in reactors 1, 2, 3
$x_{B1},\ x_{B2},\ x_{B3}$	mass fractions of B in reactors 1, 2, 3
$x_{C1},\ x_{C2},\ x_{C3}$	mass fractions of C in reactors 1, 2, 3
$x_{Ar},\ x_{Br},\ x_{Cr}$	mass fractions of $A,\ B,\ C$ in the recycle stream
$T_1,\ T_2,\ T_3$	temperatures in reactors 1, 2, 3
$T_{10},\ T_{20}$	feed stream temperatures to reactors 1 and 2
$F_1,\ F_2$	effluent flow rates from reactors 1 and 2
$F_{10},\ F_{20}$	steady-state feed stream flow rates to reactors 1 and 2
$F_r,\ F_p$	flow rates of the recycle and purge streams
$V_1,\ V_2,\ V_3$	volumes of reactors 1, 2, 3
$E_1,\ E_2$	activation energy for reactions (1) and (2)
$k_1,\ k_2$	pre-exponential values for reactions (1) and (2)
$\Delta H_1,\ \Delta H_2$	heats of reaction for reactions (1) and (2)
$\Delta H_{vap1},\ \Delta H_{vap2},\ \Delta H_{vap3}$	evaporating enthalpies for $A,\ B,\ C$
$\alpha_A,\ \alpha_B,\ \alpha_C$	relative volatilities of A, B, C
$Q_1,\ Q_2,\ Q_3$	heat inputs/removals into/from reactors 1, 2, 3
$c_p,\ R,\ \rho$	heat capacity, gas constant and solution density

Table 2.2: Process parameters for the reactors.

$F_{10} = 5.04\ m^3/h$	$\Delta H_1 = -6.0 \times 10^4\ KJ/kmol$
$F_{20} = 5.04\ m^3/h$	$\Delta H_2 = -7.0 \times 10^4\ KJ/kmol$
$F_r = 50.4\ m^3/h$	$\Delta H_{vap1} = -3.53 \times 10^4\ KJ/kmol$
$F_p = 5.04\ m^3/h$	$\Delta H_{vap2} = -1.57 \times 10^4\ KJ/kmol$
$V_1 = 1.0\ m^3/h$	$\Delta H_{vap3} = -4.068 \times 10^4\ KJ/kmol$
$V_2 = 0.5\ m^3/h$	$k_1 = 2.77 \times 10^3\ s^{-1}$
$V_3 = 1.0\ m^3/h$	$k_2 = 2.6 \times 10^3\ s^{-1}$
$\alpha_A = 3.5$	$c_p = 4.2\ KJ/kg \cdot K$
$\alpha_B = 1.0$	$R = 8.314\ KJ/kmol \cdot K$
$\alpha_C = 0.5$	$\rho = 1000.0\ kg/m^3$
$T_{10} = 300\ K$	$x_{A10} = 1$
$T_{20} = 300\ K$	$x_{B10} = 0$
$E_1 = 5.0 \times 10^4\ KJ/kmol$	$x_{A20} = 1$
$E_2 = 6.0 \times 10^4\ KJ/kmol$	$x_{B20} = 0$

there's an external input to the corresponding vessel: Q_1, Q_2 and Q_3. It is assumed that the measured outputs of the process are the temperatures (i.e., T_1, T_2 and T_3) and the measurements are subject to bounded noise. The bounded noise in the measurements is generated as normal distributed values with zero mean and standard deviation 1 but the values are restricted to be in the interval $[-2, 2]$. In addition to the measurement noise, bounded random disturbances are added to the right-hand-side of Eq. (2.33). The random disturbances added to the dynamics of the temperatures are generated as normal distributed values with zero mean and standard deviation 100 in the range $[-200, 200]$ while the disturbances added to the dynamics of the concentrations are generated as normal distributed values with zero mean and standard deviation 1 in the range $[-5, 5]$. The process has one unstable steady-state:

$$x_s = [0.1763,\ 0.6731,\ 480.3165\,K,\ 0.1965,\ 0.6536,\ 472.7863\,K,\ 0.0651,\ 0.6703,\ 474.8877\,K]^T$$

which is the desired operating point, corresponding to $Q_s = [2.9 \times 10^6\ KJ/h,\ 1.0 \times 10^6\ KJ/h,\ 2.9 \times 10^6\ KJ/h]^T$. The process is stabilized around this operating point by manipulating the three external inputs.

2.6.2 Local MHE design

First, a deterministic nonlinear observer is designed following [31] for each subsystem without considering the interactions between them. The nonlinear observers take the following form for $i = 1, 2, 3$:

$$\dot{\hat{x}}_i(t) = f_i(\hat{x}_i(t), 0) + G_i(\hat{x}_i(t))^{-1} K_{o,i}(y_i(t) - \hat{y}_i(t)) \tag{2.35}$$

where $\hat{x}_i$ denotes the state of the observer, $G_i = \dfrac{d\Phi_i(\hat{x}_i)}{d\hat{x}_i}$ with $\Phi_i(\hat{x}_i)$ defined as:

$$\Phi_i(\hat{x}_i) = [h_i(\hat{x}_i),\ \frac{\partial h_i(\hat{x}_i)}{\partial \hat{x}_i} f_i(\hat{x}_i),\ \frac{\partial(\partial h_i(\hat{x}_i)/\partial \hat{x}_i \cdot f_i(\hat{x}_i))}{\partial \hat{x}_i} f_i(\hat{x}_i)]^T$$

and $K_{o,i}$ is a gain matrix and its value is determined such that the eigenvalues of the matrix $A_{o,i} - K_{o,i}C_{o,i}$ with $A_{oi} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and $C_{oi} = [1\ 0\ 0]$ are placed at $-0.1 \pm i$ and -10. These nonlinear observers can asymptotically track the nominal subsystem states when the interactions between them are absent.

The above designed nonlinear observers are used in the design of the subsystem MHEs. As shown in (2.8), these nonlinear observers are augmented with the interaction models and the correction terms. From the process model of Eq. (2.33), it can be seen that $\mathbb{I}_1 = \{3\}$,

$\mathbb{I}_2 = \{1\}$, $\mathbb{I}_3 = \{2\}$. Following (2.10), the gains of these correction terms are determined as follows:

$$K_{1,3} = [0,\ 0,\ 50.4]^T,\ K_{2,1} = [0,\ 0,\ 110.88]^T,\ K_{3,2} = [0,\ 0,\ 60.48]^T.$$

In the design of the local MHEs, the sampling time is $\Delta = 0.005\ h$, moving horizon is $N = 3$, the parameters κ_i $(i = 1, 2, 3)$ are $\kappa_1 = \kappa_2 = \kappa_3 = 0.5$ determined based on extensive offline simulations. The bounds on subsystem states x_i are determined as $0 < x_{Ai} < 1,\ 0 < x_{Bi} < 1,\ 350K < T_i < 650K$. The weighting matrices in the cost function of each subsystem MHE are $Q_i = \text{diag}([1\ 1\ 10^4])$, $R_i = 1$ for $i = 1, 2, 3$. An extended Kalman filtering approach is used to approximate the arrival cost in the subsystem MHEs [53]. These subsystem MHEs are implemented following Algorithm 2 to estimate the entire system state in a distributed fashion.

2.6.3 Simulation results

In this section, the proposed DMHE is compared with different estimation techniques to illustrate its performance. Specifically, the proposed DMHE will be compared with 1) the deterministic nonlinear observers in the form (2.8e) implemented the same as the proposed DMHE, 2) a decentralized MHE in which the subsystem MHEs do not communicate and the interactions between subsystems are compensated for using their steady-state values, and 3) the proposed DMHE with the correction gains in (2.8e) being zero vectors (i.e., $K_{1,3} = K_{2,1} = K_{3,2} = [0,\ 0,\ 0]^T$).

First, we compare the performance of the proposed DMHE with the deterministic nonlinear observers of Eq. (2.8e) implemented following Algorithm 2 as well. The key difference between the two approaches is that in the proposed DMHE, optimality considerations are taken into account. The initial condition for the process is as follows:

$$x_0 = [0.1939,\ 0.7404,\ 528.3482\,K,\ 0.2162,\ 0.7190,\ 520.0649\,K,\ 0.0716,\ 0.7373,\ 522.3765\,K]^T$$

and initial guess for the proposed DMHE and the nonlinear observers of Eq. (2.8e) are the same:

$$\hat{x}_0 = [0.1675,\ 0.7,\ 500.3\ K,\ 0.18,\ 0.67,\ 500\ K,\ 0.06,\ 0.68,\ 500\ K]^T.$$

Figure 2.3 shows the trajectories of the estimates given by the proposed DMHE and the nonlinear observers of Eq. (2.8e). The corresponding trajectories of the norm of the estimation error is shown in Fig. 2.4. From these figures, it can be seen that both the proposed DMHE and the nonlinear observers of Eq. (2.8e) can track the actual system state. However, the proposed DMHE drives the estimates to a small region around the actual system

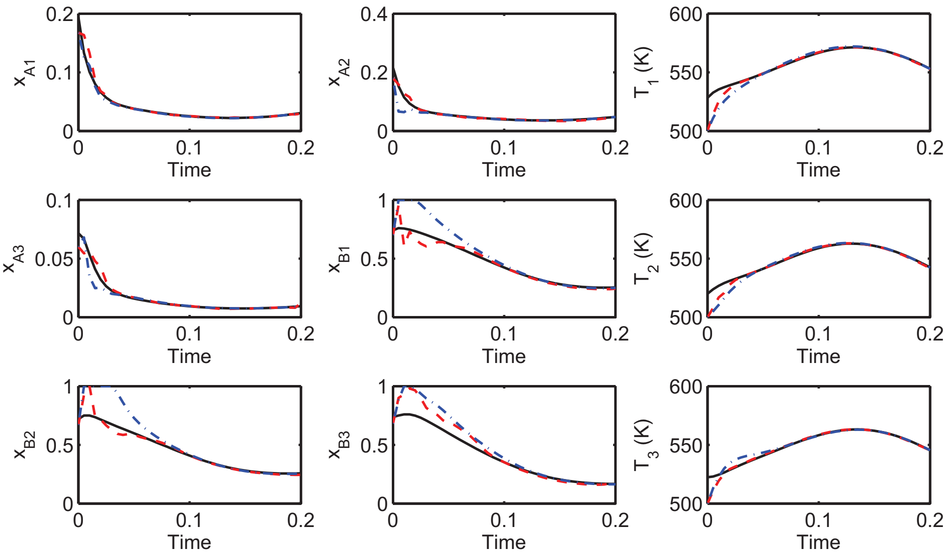


Figure 2.3: Trajectories of the actual system state (solid lines), the estimates given by the proposed DMHE (dashed lines) and the nonlinear observers of Eq. (2.8e) implemented following Algorithm 2 (dash-dotted lines).

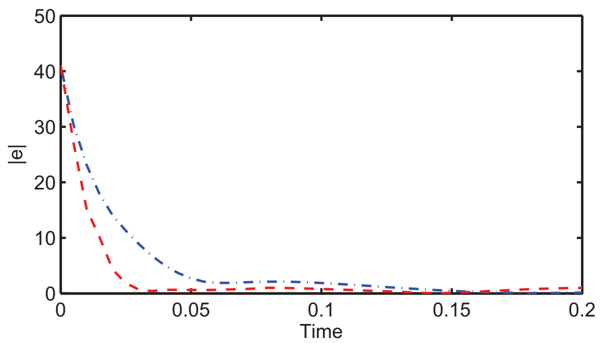


Figure 2.4: Trajectories of the estimation error norm of the proposed DMHE (dashed line) and of the nonlinear observers of Eq. (2.8e) implemented following Algorithm 2 (dash-dotted line).

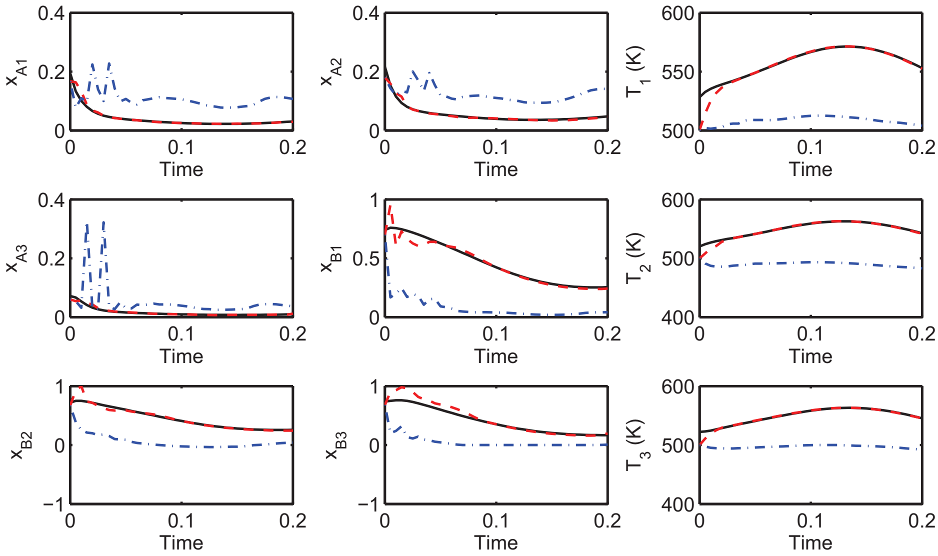


Figure 2.5: Trajectories of the actual system state (solid lines), the estimates given by the proposed DMHE (dashed lines) and a decentralized MHE in which the subsystem MHEs do not communicate and the interactions between subsystems are compensated for using their steady-state values (dash-dotted lines).

state much faster. This can be seen clearly from Fig. 2.4. This is because that even though the stability property of the proposed DMHE is essentially inherited from the nonlinear observers of Eq. (2.8e), in the proposed DMHE optimality considerations are taken into account.

Next, we compare the performance of the proposed DMHE with a decentralized MHE in which the subsystem MHEs do not communicate and the interactions between subsystems are compensated for using their steady-state values. Figures 2.5 and 2.6 show the trajectories of the estimates and the corresponding estimation error norms. From these figures, it can be seen that the proposed DMHE can track the actual system state very well while the decentralized MHE gives very unreliable estimates. This is due to the fact that in the proposed DMHE, the interactions between the subsystems are compensated for using latest estimates of the subsystem states communicated between the MHEs as well as additional correction terms to compensate for the errors in the estimates. This strategy can significantly improve the interaction compensation performance compared with the case that a constant steady-state value is used for the interactions.

In this set of simulations, we compare the performance of the proposed DMHE with the same proposed DMHE but with the correction gains in (2.8e) being zero vectors; that is,

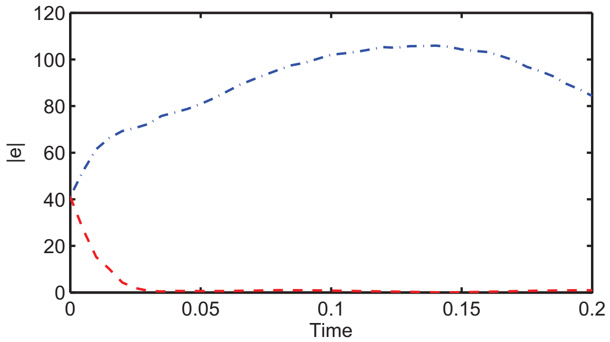


Figure 2.6: Trajectories of the estimation error norm of the proposed DMHE (dashed line) and of a decentralized MHE in which the subsystem MHEs do not communicate and the interactions between subsystems are compensated for using their steady-state values (dash-dotted lines).

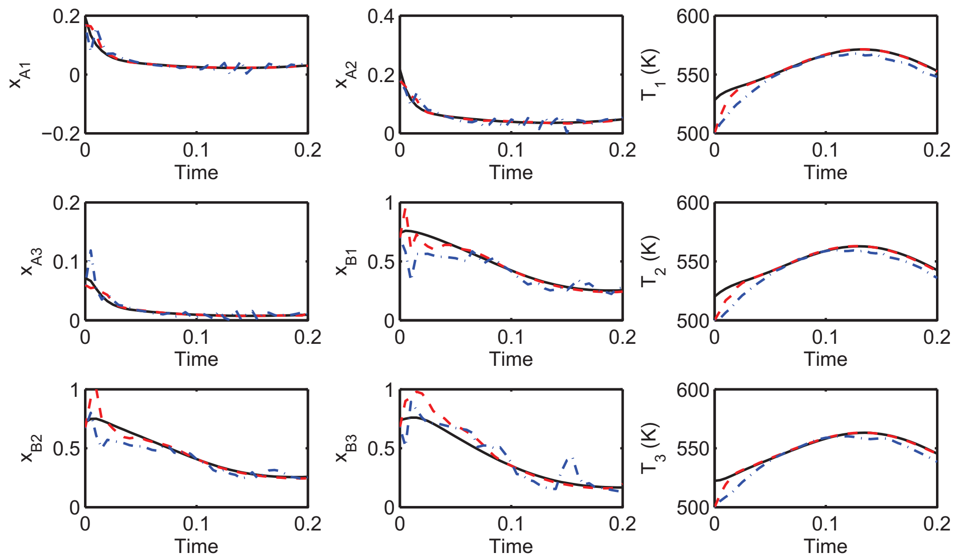


Figure 2.7: Trajectories of the actual system state (solid lines), the estimates given by the proposed DMHE (dashed lines) and the proposed DMHE with the correction gains in (2.8e) being zero vectors (dash-dotted lines).

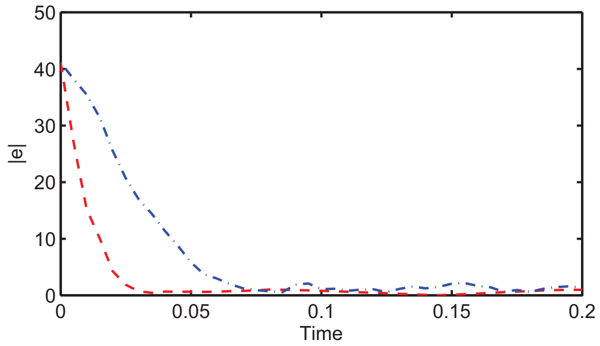


Figure 2.8: Trajectories of the estimation error norm of the proposed DMHE (dashed line) and of the proposed DMHE with the correction gains in (2.8e) being zero vectors (dash-dotted line).

$K_{1,3} = K_{2,1} = K_{3,2} = [0,\ 0,\ 0]^T$. Figures 2.7 and 2.8 show the trajectories of the estimates given by the two approaches and the corresponding estimation error norms, respectively. From Figs. 2.7 and 2.8, it can be seen that both the two approaches can track the actual system states. However, the estimates given by the proposed DMHE with the correction terms converge to the actual system states much faster than the estimates given by the approach without the correction terms. Moreover, the estimates given by the proposed DMHE without the correction terms have relatively larger fluctuations around the actual system states. This set of simulations illustrate that if we use the correction terms to compensate for the estimation errors, faster convergence rate and improved estimates can be obtained.

Finally, we demonstrate the robustness of the proposed DMHE to the uncertainties in model parameters. In this set of simulation, we consider that there are uncertainties in the inlet reactant concentrations in flow F_{10}. Specifically, we consider that the actual values of the reactant concentrations are $x_{A10} = 0.9$ and $x_{B10} = 0.1$ but in the DMHE design $x_{A10} = 1$ and $x_{B10} = 0$ are used; that is, in the DMHE design, pure A is thought to be contained in F_{10}. Figure 2.9 shows the simulation results. It can be seen from this figure that the proposed DMHE can track the actual system states. This set of simulations illustrate the robustness of the proposed DMHE with respect to model parameter uncertainties.

2.7 Conclusions

In this chapter, we developed a distributed state estimation approach in the framework of moving horizon estimation for a class of nonlinear systems. In particular, we focused on a

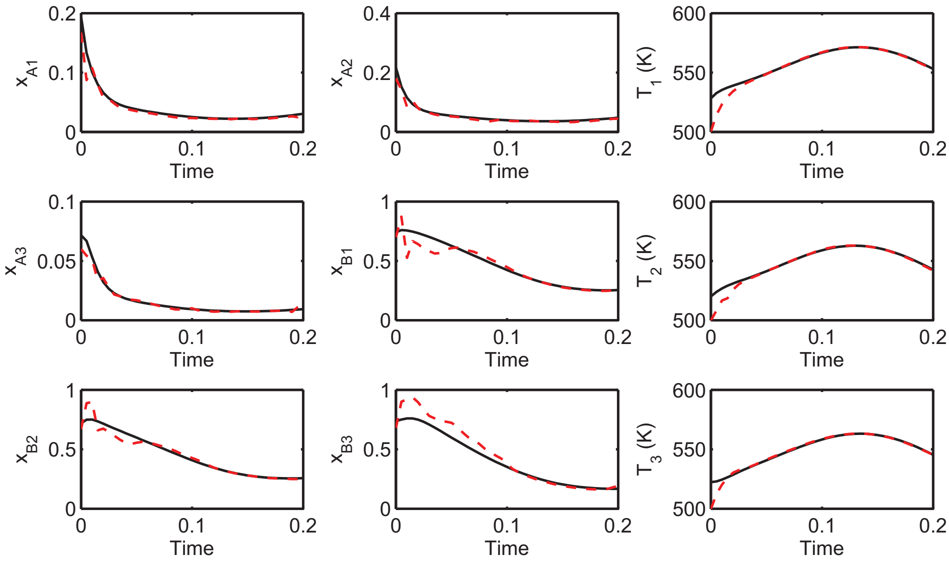


Figure 2.9: Trajectories of the actual system state (solid lines) and the estimates given by the proposed DMHE (dashed lines) subject to model parameter uncertainties.

class of nonlinear systems that are composed of several subsystems and the subsystems interact with each other via their subsystem states. First, a distributed estimation algorithm was proposed which specifies how the different subsystem MHEs collaborate. Subsequently, a local MHE scheme was designed for each subsystem. In the design of each subsystem MHE, an auxiliary nonlinear deterministic observer that can asymptotically track the corresponding nominal subsystem state when the subsystem interactions are absent was taken advantage of. For each subsystem, the nonlinear deterministic observer together with an error correction term was used to calculate a confidence region for the subsystem state every sampling time. Within the confidence region, the subsystem MHE was allowed to optimize its estimate. Sufficient conditions under which the proposed DMHE scheme gives bounded estimation errors in the case of bounded measurement noise and bounded process disturbances were derived. The performance of the proposed DMHE was illustrated via the application to a reactor-separator chemical process by comparing it with three other distributed estimation approaches.

Chapter 3

Two Triggered Communication Algorithms for Distributed Moving Horizon State Estimation*

3.1 Introduction

In Chapter 2, we developed a DMHE strategy for a class of nonlinear system, which requires information exchange every sampling time. In that strategy, for each subsystem, a nonlinear observer of the form (2.5) is augmented with a correction term based on the information communicated at the previous sampling time to calculate a confidence region at each sampling time. Within the confidence region, the corresponding local MHE is allowed to optimize its state estimates. The DMHE developed in Chapter 2 was proved to give decreasing and ultimately bounded estimation error. However, the results in Chapter 2 were obtained based on information communication every sampling time. The frequent information transmission requirement may impede the application of the DMHE to processes that have a shared communication network with limited capacity. Moreover, extensive information exchanging may reduce the robustness of the system due to data dropouts in the communication network. Motivated by the above observations, in this chapter, we propose two algorithms to reduce the number of information transmissions between subsystem based on the DMHE framework developed in Chapter 2 via event-triggered approaches. Event-triggered approaches have been widely used in the design of control systems that have shared communication and computation resources (e.g., [66, 67, 68]). When a triggered strategy is used to reduce the frequency of communication of the distributed state estimation system, the implementation algorithm and local MHE design in Chapter 2 need

*This chapter is a version of "J. Zhang and J. Liu. Two triggered information transmission algorithms for distributed moving horizon state estimation. Systems & Control Letters, 65:1-12, 2014."

to be redesigned to account for triggered communication between the subsystems in order to achieve boundedness of the estimation error.

3.2 Modeling of measurements

In this chapter, we consider that the outputs of the m subsystems, y_i, $i = 1, \ldots, m$, are sampled synchronously and periodically at time instants $\{t_{k\geq 0}\}$ such that $t_k = t_0 + k\Delta$ with $t_0 = 0$ the initial time, Δ a fixed sampling time interval and k positive integers. It's also assumed that the measurements of the time derivatives of the outputs, $\dot{y}_i$, …, $y_i^{(n-1)}$, $i = 1, \ldots, m$, are available at each sampling time.

Note that the availability of the output time derivatives is only required in the design of one of the two proposed triggered communication algorithms of the proposed DMHE. For the other proposed triggered communication algorithm, it only requires the availability of the output measurements (i.e., $y_i(t)$). The difference of the two triggering conditions will be discussed in Section 3.5.

In the distributed state estimation scheme, subsystem i is assumed to have direct and immediate access to the corresponding local output and its derivatives. The subsystems are assumed to be able to communicate with each other bi-directionally to exchange their subsystem state estimates and measurements when necessary.

3.3 The DMHE scheme with triggered communications

In this chapter, we will discuss the proposed DMHE design with triggered communication for the nonlinear system of Eq. (2.4) to minimize the communication cost of the distributed state estimation system. The structure of this design is shown in Figure 3.1. In this scheme, each subsystem has an MHE estimator and an communication trigger which determines if the information of the subsystem should be sent out to other subsystems at a sampling time. This implies that a subsystem does not necessarily send out information at each sampling time, which can reduce the communication load of the distributed state estimation system.In the first algorithm, the triggering condition is designed based on the difference between the current sate estimate and the last sent state estimate. In the second algorithm, the triggering condition is based on the difference between the sampled current output as well as its derivatives and the last sent output as well as its derivatives. The remainder of this chapter is organized as follows: first, an implementation algorithm for the DMHE with triggered communication is presented which is followed by the first triggering condition; then, the

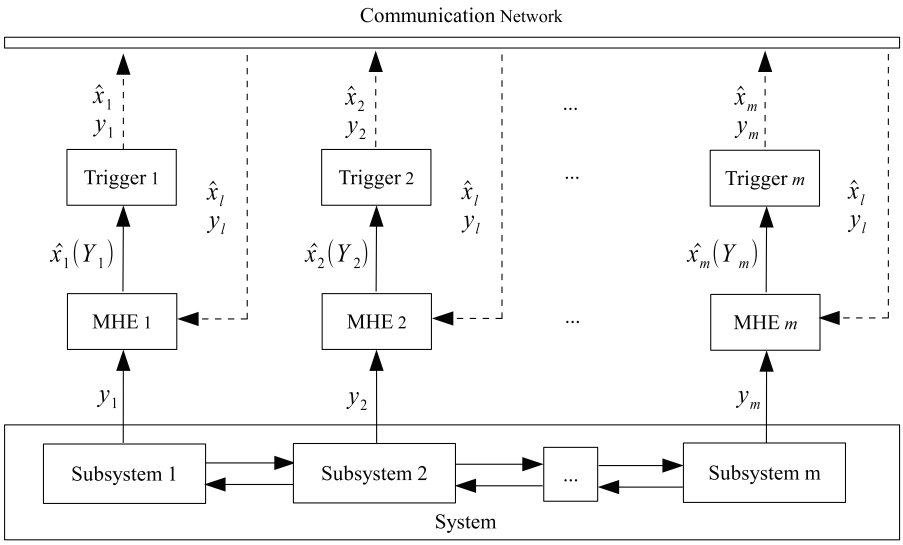


Figure 3.1: Scheme of the proposed DMHE design with triggered communication.

design of local MHE taking into account triggered communication explicitly is introduced; subsequently, the stability properties of the proposed DMHE with triggered communication are analyzed based on the first triggering condition; finally, the second triggering condition is proposed and the corresponding DMHE stability properties are proved.

3.4 DMHE with the first communication trigger

3.4.1 Implementation algorithm

In this section, we will discuss the proposed DMHE with the first triggered communication approach in which the trigger of a subsystem is designed based on the difference between the current and the previous sent state estimates of the subsystem. First, we propose the implementation algorithm for this strategy in the following Algorithm 14:

Algorithm 14 *Distributed state estimation algorithm 1*

1. *At the initial sampling time $t_0 = 0$, MHE i, $i \in \mathbb{I}$, is initialized with $y_i(t_0)$, $\hat{x}_i(t_0)$ as well as $y_l(t_0)$, $\hat{x}_l(t_0)$ for $l \in \mathbb{I}_i$.*

2. *At the current sampling time $t_k > 0$, MHE i, $i \in \mathbb{I}$, and its associated trigger carry out the following steps:*

   2.1. *MHE i receives the corresponding output measurement $y_i(t_k)$.*

   2.2. *MHE i calculates the current state estimate $\hat{x}_i(t_k)$ based on the local measurements $y_i(t_{k+s-N})$, $s = 0, 1, \ldots, N$, and the latest received information $\hat{x}_l(t_q^l)$ and*

$y_l(t_q^l)$ for $l \in \mathbb{I}_i$, where t_q^l is the last time instant that MHE l sends information to MHE i.

2.3. MHE i sends $\hat{x}_i(t_k)$ and $y_i(t_k)$ to its trigger and the trigger checks the triggering condition. If the condition is satisfied, the trigger updates $t_q^i = t_k$ and sends $\hat{x}_i(t_q^i)$ as well as $y_i(t_q^i)$ to subsystem j for all $j \in \mathbb{C}_i$. If the triggering condition is not satisfied, no information is transmitted to subsystem j from subsystem i and subsystem j continues to use the last updated information $\hat{x}_i(t_q^i)$ and $y_i(t_q^i)$.

3. Go to Step 2 at the next sampling time t_{k+1}.

In the above algorithm, the communication occurs in Step 2.3 at the end of each sampling time, implies that the current state estimate of a subsystem, if it is sent out to other subsystems, will be used to compensate for the interactions between the subsystems at the next sampling time. Moreover, the communication between subsystem is designed in a parallel and non-iterative fashion.

3.4.2 The first triggering condition

From Step 2.3 of Algorithm 14, it can be seen that the triggering condition for each subsystem is checked every sampling time after MHE i calculates its latest state estimate. The first triggering condition is designed based on the difference between the current subsystem state estimate and the last subsystem state estimate sent to other subsystems. Specifically, the triggering condition of MHE i at time t_k is designed as follows:

$$S_i(t_k) = \begin{cases} 1, & \text{if } |\hat{x}_i(t_k) - \hat{x}_i(t_q^i)| \geq \epsilon_i \\ 0, & \text{if } |\hat{x}_i(t_k) - \hat{x}_i(t_q^i)| < \epsilon_i \end{cases} \tag{3.1}$$

where t_q^i is the last sampling time that MHE i sent information to other MHEs, $\hat{x}_i(t_k)$ is the current state estimate of MHE i and $\hat{x}_i(t_q^i)$ is the last sent state estimate of MHE i. ϵ_i is a pre-determined threshold. When $S_i(t_k) = 0$, the triggering condition is not satisfied and MHE i does not send out information so the other MHEs will continue to use $\hat{x}_i(t_q^i)$ and $y_i(t_q^i)$. When $S_i(t_k) = 1$, the triggering condition is satisfied and MHE i sends out $\hat{x}_i(t_k)$, $y_i(t_k)$ and updates $t_q^i = t_k$.

It can be seen from (3.1) that the triggering condition of a subsystem is independent from the states of other subsystems. This implies that the triggering conditions for different subsystems may be satisfied at different time instants.

3.4.3 Local MHE formulation

Due to the triggered communication of the DMHE, the subsystems may not transmit their latest state estimates and output measurements every sampling time. Thus the local MHE design developed in Chapter 2 which requires state estimates transmission every sampling time needs to be modified in order to ensure the robustness and stability of the distributed state estimation system under the proposed triggered implementation algorithms.

Before presenting the proposed local MHE design, we define the sampled state trajectory $z_{n,i}(t_k)$ which is obtained by integrating the following ordinary differential equation from t_{k-1} to t_k:

$$\begin{aligned}\dot{z}_{n,i}(t) = F_i(z_{n,i}(t), y_i(t_{k-1})) + \tilde{f}_i(\hat{X}_i(t_{k-1})) + \sum_{l\in\mathbb{I}_i, l\neq i} K_{i,l}(\hat{x}_l)(y_l(t_q^l) - h_l(\hat{x}_l(t_q^l))) \\ + \mu_i K_{i,i}(\hat{x}_i)(y_i(t_{k-1}) - h_i(\hat{x}_i(t_{k-1})))\end{aligned} \tag{3.2a}$$

$$z_{n,i}(t_{k-1}) = \hat{x}_i(t_{k-1}) \tag{3.2b}$$

where the values of t_q^l for $l \in I_i$ are the updated ones at time t_{k-1}. In (3.2a), μ_i is a variable used to indicate if x_i is included in X_i. If x_i is included in X_i, then $\mu_i = 1$; if x_i is not included in X_i, then $\mu_i = 0$. $\hat{X}_i(t_{k-1})$ is an approximation of $X_i(t)$ and is composed of $\hat{x}_l(t_q^l)$ ($l \in \mathbb{I}_i, l \neq i$) and/or $\hat{x}_i(t_{k-1})$. The evolution of $z_{n,i}(t)$ should be evaluated before the evaluation of MHE i at the beginning of each sampling time based on the previous state estimate and output measurement of subsystem i and the latest information received from other subsystems. The information from other subsystems is not available every sampling time, so the last updated information is used to approximate the information of the pervious time instant. Specifically, the term $\hat{x}_l(t_{k-1})$ and $y_l(t_{k-1})$ are approximated by $\hat{x}_l(t_q^l)$ and $y_l(t_q^l)$ for $l \in \mathbb{I}_i, l \neq i$, respectively.

In nonlinear observer (3.2a), the first term of the right hand side comes from nonlinear observer (2.5), the second term explicitly describe the interactions between subsystem i and its associated subsystems l based on the interaction model $\tilde{f}_i(X_i)$ and the last two terms compensate for the error in the interaction model.

The gains $K_{i,l}, l \in \mathbb{I}_i$ associated with the correction terms are determined as follows:

$$K_{i,l} = \begin{cases} \left.\dfrac{\partial \tilde{f}_i}{\partial x_l}\left(\dfrac{\partial h_l}{\partial x_l}\right)^+\right|_{x_l=\hat{x}_l(t_q^l)}, & \text{if } l \neq i \\ \left.\dfrac{\partial \tilde{f}_i}{\partial x_i}\left(\dfrac{\partial h_i}{\partial x_i}\right)^+\right|_{x_i=\hat{x}_i(t_{k-1})}, & \text{if } l = i \end{cases} \tag{3.3}$$

for $l \in \mathbb{I}_i$ and $i \in 1, \ldots, m$. The gains of the corrections terms are picked to compensate for the error in the interaction model via its linear approximation which will be made explicit in Section 3.4.4.

Based on $z_{n,i}(t_k)$, the proposed local MHE for subsystem i accounting for triggered communication at t_k is designed as follows:

$$\min_{\tilde{x}_i(t_{k-N}),\ldots,\tilde{x}_i(t_k)} \sum_{p=k-N}^{k-1} |w_i(t_p)|^2_{Q_i^{-1}} + \sum_{p=k-N}^{k} |v_i(t_p)|^2_{R_i^{-1}} + V_i(t_{k-N}) \tag{3.4a}$$

$$\text{s.t. } \dot{\tilde{x}}_i(t) = f_i(\tilde{x}_i(t), w_i(t_p)) + \tilde{f}_i(\hat{X}_i(t_p)),\ t \in [t_p, t_{p+1}],\ p = k-N, \ldots, k-1 \tag{3.4b}$$

$$v_i(t_p) = y_i(t_p) - h_i(\tilde{x}_i(t_p)),\ p = k-N, \ldots, k \tag{3.4c}$$

$$w_i(t_p) \in \mathbb{W}_i,\ v_i(t_p) \in \mathbb{V}_i,\ \tilde{x}_i(t_p) \in \mathbb{X}_i,\ p = k-N, \ldots, k \tag{3.4d}$$

$$|\tilde{x}_i(t_k) - z_{n,i}(t_k)| \leq \kappa_i |y_i(t_k) - h_i(z_{n,i}(t_k))| \tag{3.4e}$$

where N is the estimation horizon, Q_i and R_i are the covariance matrices of w_i and v_i respectively, $V_i(t_{k-N})$ denotes the arrival cost which summarizes past information up to t_{k-N}, $\tilde{x}_i$ is the predicted x_i in the above optimization problem, and κ_i is a design parameter.

Once the optimization problem (3.4) is solved, $\tilde{x}_i^*(t_{k-N}), \ldots, \tilde{x}_i^*(t_k)$, an optimal trajectory of the system states is obtained. The optimal estimate of the state of subsystem i at t_k is defined as:

$$\hat{x}_i(t_k) = \tilde{x}_i^*(t_k). \tag{3.5}$$

In optimization problem (3.4), constraint (3.4a) is the cost function that needs to be minimized and $V_i(t_{k-N})$ is the arrival cost summarizing all the past information out of the estimation horizon. Constraint (3.4b) is the model of subsystem i, $f_i(\hat{X}_i(t_p))$ is used to approximate the function $f_i(X_i(t))$, $t \in [t_p, t_{p+1}]$. Thus each MHE is expected to store the previously received information of other subsystems within the estimation horizon. The equation of constraint (3.4d) are constraints on process disturbances, measurement noise and system state.

Constraint (3.4e) creates a confidence region (i.e., $\kappa_i|y_i(t_k) - h_i(z_{n,i}(t_k))|$) taking advantage of the reference state estimate provided by nonlinear observer (3.2a) (i.e., $z_{n,i}(t_k)$) and the current output measurement (i.e., $y_i(t_k)$). The estimate of the MHE of Eq. (3.4) is only allowed to be optimized within the confidence region. This method guarantees that the proposed DMHE with triggered communication gives bounded estimation error when certain conditions are satisfied. The parameter κ_i is a design parameter whose value depends on the system and observer (2.5) properties. Guidelines for picking κ_i will be provided in Theorems 16 and 20.

3.4.4 Stability analysis

In this subsection, the stability property of the proposed DMHE with triggered communication based on the first triggering condition (i.e., MHEs of Eq. (3.4) implemented following Algorithm 14) is studied. First we study the boundedness of the estimation error of nonlinear observer (3.2a) with $K_{i,l}$ determined following (3.3) taking into account measurement noise and process disturbances. Subsequently we derive sufficient conditions under which the stability and ultimate boundedness of the estimation error of the proposed DMHE with triggered communication are guaranteed.The following Proposition 15 gives an upper bound on the deviation of $z_{n,i}$ from x_i in one sampling time.

Proposition 15 *Consider the nonlinear observer of Eq. (3.2a) for subsystem i, $i \in \mathbb{I}$, in the time interval $t \in [t_k, t_{k+1}]$ with initial condition $z_{n,i}(t_k) = \hat{x}_i(t_k)$, output measurement $y_i(t_k)$ and the values of t_q^l, $l \in I_i$, determined following the triggering condition (3.1). If $K_{i,l}$ for $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$ are determined as in (3.3) and $K_{i,l}$ are bounded, then the deviation of the observer state $z_{n,i}$ in one sampling time Δ (i.e., at t_{k+1}) from the actual subsystem state x_i is bounded for all $x_i \in \mathbb{X}_i$, $i \in \mathbb{I}$, as follows:*

$$|e_{z,i}(t_{k+1})| \leq \beta_i(|e_{z,i}(t_k)|, \Delta) + \gamma_i(\Delta) + L_i\Delta|e_{z,i}(t_k)|^2 + \sum_{l\in\mathbb{I}_i,\, l\neq i} \alpha_{i,l}(\Delta, e_{z,l}(t_k), e_{z,l}(t_q^l), \epsilon_l) \tag{3.6}$$

where $e_{z,i} = z_{n,i} - x_i$, $i \in \mathbb{I}$, and $\gamma_i(\tau) = L_{F_i}^{y_i} L_{h_i} M_i \tau^2/2 + L_{F_i}^{y_i}\theta_{v_i}\tau + L_{f_i}^{w_i}\theta_{w_i}\tau + \sum_{l\in\mathbb{I}_i} M_{K_{i,l}}\theta_{v_l}\tau + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \tau^2/2$ with $L_{F_i}^{y_i}$, L_{h_i}, $L_{f_i}^{w_i}$, and $L_{\tilde{f}_i}^{x_l}$ being the Lipschitz constants of F_i with respect to y_i, h_i with respect to x_i, f_i with respect to w_i, and $\tilde{f}_i$ with respect to x_l, respectively, and M_i, $M_{K_{i,l}}$, $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$, being constants that bound $\dot{x}_i$ in $\mathbb{X}_i$, and $K_{i,l}$ in $\mathbb{X}_l$, respectively, and $H_i^{\tilde{f}_i}$, $H_l^{h_l}$, and $H_i^{h_i}$ being positive constants that associated with the Taylor expansions of $\tilde{f}_i$, h_l and h_i, $L_i = \mu_i(M_{K_{i,l}} H_i^{h_i} + H_i^{\tilde{f}_i})$, and $\alpha_{i,l}(\tau, e_{z,l}(t_k), e_{z,l}(t_q^l), \epsilon_l) = K_{i,l} L_{h_l} \tau(|e_{z,l}(t_k)| + |e_{z,l}(t_q^l)| + \epsilon_l) + \tau(K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(|e_{z,l}(t_k)| + \epsilon_l)^2$.

Proof: We consider the nonlinear observer of Eq. (3.2a) and define $e_{z,i} = z_{n,i} - x_i$ where $z_{n,i}$ denotes the trajectory of observer (3.2a) and x_i is the state trajectory of the actual subsystem of Eq. (2.1). The time derivative of $e_{z,i}$ is evaluated as follows:

$$\begin{aligned}\dot{e}_{z,i}(t) = {} & F_i(z_{n,i}(t), y_i(t_k)) - f_i(x_i(t), w_i(t)) + \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) \\ & + \sum_{l\in\mathbb{I}_i, l\neq i} K_{i,l}(\hat{x}_l)(y_l(t_q^l) - h_l(\hat{x}_l(t_q^l)) + \mu_i K_{i,i}(\hat{x}_i)(y_i(t_k) - h_i(\hat{x}_i(t_k)))\end{aligned} \tag{3.7}$$

From the Lipschitz properties of F_i, f_i and h_i, the fact that $y_i(t_k) = h_i(x_i(t_k)) + v_i(t_k)$, and $|v_i(t_k)| \leq \theta_{v_i}$, and $|w_i(t)| \leq \theta_{w_i}$, the following inequality can be obtained from (3.7):

$$\begin{aligned} |\dot{e}_{z,i}(t)| \leq\ & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i} |x_i(t) - x_i(t_k)| + L_{F_i}^{y_i}\theta_{v_i} + L_{f_i}^{w_i}\theta_{w_i} \\ & + \left| \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l}(\hat{x}_l)(y_l(t_q^l) - h_l(\hat{x}_l(t_q^l))) + \mu_i K_{i,i}(\hat{x}_i)(y_i(t_k) - h_i(\hat{x}_i(t_k))) \right. \\ & \left. + \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) \right| \end{aligned} \tag{3.8}$$

where $L_{F_i}^{y_i}$, L_{h_i} and $L_{f_i}^{w_i}$ are the Lipschitz constants associated with F_i, h_i and f_i, respectively.

Using Taylor series expansion, the following inequalities can be obtained:

$$\begin{aligned} \tilde{f}_i(X_i(t_k)) &= \tilde{f}_i(\hat{X}_i(t_k)) + \sum_{l \in \mathbb{I}_i, l \neq i} \frac{\partial \tilde{f}_i}{\partial \hat{x}_l}(\hat{x}_l(t_q^l))(x_l(t_k) - \hat{x}_l(t_q^l)) \\ &\quad + \mu_i \frac{\partial \tilde{f}_i}{\partial \hat{x}_i}(\hat{x}_i(t_k))(x_i(t_k) - \hat{x}_i(t_k)) + H.O.T_i^{\tilde{f}_i} \\ h_l(x_l(t_k)) &= h_l(\hat{x}_l(t_q^l)) + \frac{\partial h_l}{\partial \hat{x}_l}(\hat{x}_l(t_q^l))(x_l(t_k) - \hat{x}_l(t_q^l)) + H.O.T_l^{h_l},\ \forall\, l \in \mathbb{I}_i, l \neq i \\ h_i(x_i(t_k)) &= h_i(\hat{x}_i(t_k)) + \frac{\partial h_i}{\partial \hat{x}_i}(\hat{x}_i(t_k))(x_i(t_k) - \hat{x}_i(t_k)) + H.O.T_i^{h_i} \end{aligned} \tag{3.9}$$

where $H.O.T_i^{\tilde{f}_i}$, $H.O.T_i^{h_i}$ and $H.O.T_l^{h_l}$ are high order terms associated with the expansions of $\tilde{f}_i$, h_i and $h_l, l \in \mathbb{I}_i, l \neq i$. These high order terms satisfy the following constraints:

$$\begin{aligned} H.O.T_i^{\tilde{f}_i} &\leq H_i^{\tilde{f}_i}|X_i(t_k) - \hat{X}_i(t_k)|^2, \quad H.O.T_i^{h_i} \leq H_i^{h_i}|x_i(t_k) - \hat{x}_i(t_k)|^2 \\ H.O.T_l^{h_l} &\leq H_l^{h_l}|x_l(t_k) - \hat{x}_l(t_q^l)|^2,\ \forall\, l \in \mathbb{I}_i, l \neq i \end{aligned} \tag{3.10}$$

for all $x_i \in \mathbb{X}_i$ and $x_l \in \mathbb{X}_l$, $l \neq i$ with $H_i^{\tilde{f}_i}$, $H_i^{h_i}$ and $H_l^{h_l}$, $l \in I_i$, $l \neq i$, are positive constants.

Define $A_i = \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t_k)) + \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l}(\hat{x}_l(t_q^l))(h_l(x_l(t_k)) - h_l(\hat{x}_l(t_q^l))) + \mu_i K_{i,i}(\hat{x}_i)(h_i(x_i(t_k)) - h_i(\hat{x}_i(t_k)))$. From (3.9), the following equation can be written:

$$\begin{aligned} A_i &= \mu_i \left(-\frac{\partial \tilde{f}_i}{\partial \hat{x}_i}(\hat{x}_i(t_k))(x_i(t_k) - \hat{x}_i(t_k)) + K_{i,i}\frac{\partial h_i}{\partial \hat{x}_i}(\hat{x}_i)(x_i(t_k) - \hat{x}_i(t_k)) \right) \\ &\quad + \sum_{l \in \mathbb{I}_i, l \neq i} \left(-\frac{\partial \tilde{f}_i}{\partial \hat{x}_l}(\hat{x}_l(t_q^l))(x_l(t_k) - \hat{x}_l(t_q^l)) + K_{i,l}\frac{\partial h_l}{\partial \hat{x}_l}(\hat{x}_l(t_q^l)) \left(x_l(t_k) - \hat{x}_l(t_q^l) \right) \right) \\ &\quad + \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l} H.O.T_l^{h_l} + \mu_i K_{i,i} H.O.T_i^{h_i} - H.O.T_i^{\tilde{f}_i} \end{aligned} \tag{3.11}$$

If $K_{i,l}$ is determined following (3.3), from (3.10) and (3.11), it can be obtained that:

$$|A_i| \leq \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l} H_l^{h_l}|x_l(t_k) - \hat{x}_l(t_q^l)|^2 + \mu_i K_{i,i} H_i^{h_i}|x_i(t_k) - \hat{x}_i(t_k)|^2 + H_i^{\tilde{f}_i}|X_i(t_k) - \hat{X}_i(t_k)|^2 \tag{3.12}$$

Using the triangle inequality $|x_l(t_k) - \hat{x}_l(t_q^l)| \leq |x_l(t_k) - \hat{x}_l(t_k)| + |\hat{x}_l(t_k) - \hat{x}_l(t_q^l)|$, and $|X_i(t_k) - \hat{X}_i(t_k)|^2 = \sum_{l \in \mathbb{I}_i, l \neq i} |x_l(t_k) - \hat{x}_l(t_q^l)|^2 + \mu_i |x_i(t_k) - \hat{x}_i(t_k)|^2$, (3.12) becomes:

$$
\begin{aligned}
|A_i| \quad \leq \quad & \sum_{l \in \mathbb{I}_i, l \neq i} (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(|x_l(t_k) - \hat{x}_l(t_k)| + |\hat{x}_l(t_k) - \hat{x}_l(t_q^l)|)^2 \\
& + \mu_i (K_{i,i} H_i^{h_i} + H_i^{\tilde{f}_i}) |x_i(t_k) - \hat{x}_i(t_k)|^2
\end{aligned}
\tag{3.13}
$$

From (3.8) and the Lipschitz property of $\tilde{f}_i$ with respect to x_l ($l \in \mathbb{I}_i$), the following inequality can be obtained:

$$
\begin{aligned}
|\dot{e}_{z,i}(t)| \quad \leq \quad & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i} |x_i(t) - x_i(t_k)| + \sum_{l \in \mathbb{I}_i} K_{i,l} \theta_{v_l} + L_{F_i}^{y_i} \theta_{v_i} \\
& + L_{f_i}^{w_i} \theta_{w_i} + \sum_{l \in \mathbb{I}_i} L_{\tilde{f}_i}^{x_l} |x_l(t) - x_l(t_k)| + \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l} L_{h_l} (|e_{z,l}(t_k)| + |e_{z,l}(t_q^l)| + |\hat{x}_l(t_k) - \hat{x}_l(t_q^l)|) \\
& + \sum_{l \in \mathbb{I}_i, l \neq i} (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(|e_{z,l}(t_k)| + |\hat{x}_l(t_k) - \hat{x}_l(t_q^l)|)^2 + \mu_i (K_{i,i} H_i^{h_i} + H_i^{\tilde{f}_i}) |e_{z,i}(t_k)|^2
\end{aligned}
\tag{3.14}
$$

with $e_{z,i}(t_k) = x_i(t_k) - \hat{x}_i(t_k)$, $e_{z,l}(t_k) = x_l(t_k) - \hat{x}_l(t_k)$ and $e_{z,l}(t_q^l) = x_l(t_q^l) - \hat{x}_l(t_q^l)$.

From the triggering condition of Eq. (3.1), it can be written for all $l \in \mathbb{I}_i$ that:

$$
|\hat{x}_l(t_k) - \hat{x}_l(t_q^l)| \leq \epsilon_l \tag{3.15}
$$

Using constraint (3.15), taking into account the boundedness of the system state and condition (2.6), integrating (3.14) from $t = t_k$ to $t = t_{k+1}$, the following inequality can be obtained:

$$
\begin{aligned}
|e_{z,i}(t_{k+1})| \quad \leq \quad & \beta_i(|e_{z,i}(t_k)|, \Delta) + L_{F_i}^{y_i} L_{h_i} M_i \Delta^2/2 + L_{F_i}^{y_i} \theta_{v_i} \Delta + L_{f_i}^{w_i} \theta_{w_i} \Delta + \sum_{l \in \mathbb{I}_i} M_{K_{i,l}} \theta_{v_l} \Delta \\
& + \sum_{l \in \mathbb{I}_i} L_{\tilde{f}_i}^{x_l} \Delta^2/2 + \sum_{l \in \mathbb{I}_i,\, l \neq i} K_{i,l} L_{h_l} \Delta (|e_{z,l}(t_k)| + |e_{z,l}(t_q^l)| + \epsilon_l) \\
& + \sum_{l \in \mathbb{I}_i,\, l \neq i} (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i}) \Delta (|e_{z,l}(t_k)| + \epsilon_l)^2 + \mu_i (K_{i,i} H_i^{h_i} + H_i^{\tilde{f}_i}) |e_{z,i}(t_k)|^2 \Delta
\end{aligned}
\tag{3.16}
$$

where M_i, $i \in \mathbb{I}$, are constants that bounds $\dot{x}_i$ in $\mathbb{X}_i$ (i.e., $|\dot{x}_i| \leq M_i$), and $M_{K_{i,l}}$, $l \in \mathbb{I}_i$, are constants that bounds $K_{i,l}$ in $\mathbb{X}_l$ (i.e., $|K_{i,l}| \leq M_{K_{i,l}}$). If $\gamma_i(\tau)$, L_i, and $\alpha_{i,l}$ are defined as in Proposition 15, (3.16) can be written in the form of Eq. (3.6). This proves Proposition 15. □

Proposition 15 provides an upper bound on the estimation error of a subsystem state between the nonlinear observer (3.2a) $z_{n,i}$ and the actual system state x_i. This upper bound is related to the accuracy of the initial estimate $|e_{z,i}(t_k)|$, Lipschitz properties of the system, sampling interval Δ, magnitudes of noise and process disturbances, subsystem interactions

$|e_{z,l}(t_k)|$, $|e_{z,l}(t_q^l)|$ and the triggering thresholds ϵ_l. From the formulation of the local MHE of Eq. (3.4), it can be seen that observer (3.2a) is used to generate a reference estimate $z_{n,i}(t_k)$. Based on the reference estimate, a confidence region is calculated for the optimal state estimate $\hat{x}_i(t_k)$. Theorem 16 below provides sufficient conditions for the convergence and ultimate boundedness of the estimation error of the proposed DMHE with the first triggered communication strategy.

Theorem 16 *Consider system (2.4) with the outputs of its subsystems y_i sampled at time instants $\{t_{k\geq 0}\}$. If the proposed DMHE implemented following Algorithm 14 based on the triggering condition (3.1) with subsystem MHE designed as in (3.4) based on deterministic nonlinear observers satisfying (2.6) and $K_{i,l}$ determined following (3.3), and if there exist concave functions $g_i(\cdot)$, $i \in \mathbb{I}$, such that:*

$$g_i(|e_i|) \geq \beta_i(|e_i|, \Delta) \tag{3.17}$$

for all $|e_i| \leq d_i$ and if there exist constants $d_{s,i}$, d_i such that $0 \leq d_{s,i} \leq d_i$ and positive constants $a_i \geq 1$, $b_i > 0$, and $\eta_i > 0$, such that:

$$d_{s,i} - a_i \left(g_i(d_{s,i}) + \gamma_i(\Delta) + L_i \Delta d_{s,i}^2 + \sum_{l \in \mathbb{I}_i,\, l \neq i} \alpha_{i,l}(\Delta, d_l, d_l, \epsilon_l) \right) - b_i \theta_{v_i} \geq \eta_i \tag{3.18}$$

for all $i \in \mathbb{I}$, and if κ_i for all $i \in \mathbb{I}$, are picked as follows:

$$0 \leq \kappa_i \leq \min\{(a_i - 1)/L_{h_i}, b_i\}, \tag{3.19}$$

then the estimation error $|e_i| = |\hat{x}_i - x_i|$ ($i \in \mathbb{I}$) is a decreasing sequences if $|e_i(0)| \leq d_i$ for all $i \in \mathbb{I}$ and is ultimately bounded as follows:

$$\lim_{t\to\infty} \sup |e_i(t)| \leq d_{i,\min} \tag{3.20}$$

for $i \in \mathbb{I}$ with $d_{i,\min} = \max\{|e_i(t + \Delta)| : |e_i(t)| \leq d_{s,i}\}$ for all $e_i(0) \leq d_i$ and $x_i \in \mathbb{X}_i$. This also implies that the entire system state estimation error is ultimately bounded.

Proof: We prove that the evolution of the estimation error of each subsystem state $|e_i| = |\hat{x}_i - x_i|$, $i \in \mathbb{I}$, under the proposed DMHE with the local MHE of Eq. (3.4) implemented following Algorithm 14 is a decreasing sequence and is ultimately bounded in a small region around zero. The decrease and ultimate boundedness of subsystem estimation errors imply the decrease and ultimate boundedness of the entire system state estimation error. Specifically, we first focus on MHE i, $i \in \mathbb{I}$, for the time interval from t_k to t_{k+1} and then extend to the general case. From constraint (3.4e) for MHE i, it can be written that:

$$|\hat{x}_i(t_{k+1}) - z_{n,i}(t_{k+1})| \leq \kappa_i |y_i(t_{k+1}) - h_i(z_{n,i}(t_{k+1}))| \tag{3.21}$$

From the Lipschitz property of h_i, the fact that $y_i = h_i(x_i) + v_i$ and $|v_i| \leq \theta_{v_i}$, it is obtained that:

$$|\hat{x}_i(t_{k+1}) - z_{n,i}(t_{k+1})| \leq \kappa_i L_{h_i} |x_i(t_{k+1}) - z_{n,i}(t_{k+1})| + \kappa_i \theta_{v_i} \tag{3.22}$$

where L_{h_i} is the Lipschitz constant of h_i as defined in Proposition 15. Using the triangle inequality $|\hat{x}_i - x_i| \leq |\hat{x}_i - z_{n,i}| + |z_{n,i} - x_i|$, it is obtained from (3.22) that:

$$|\hat{x}_i(t_{k+1}) - x_i(t_{k+1})| \leq (1 + \kappa_i L_{h_i}) |x_i(t_{k+1}) - z_{n,i}(t_{k+1})| + \kappa_i \theta_{v_i} \tag{3.23}$$

From Proposition 15 and (3.23), and noticing that $e_i(t_k) = e_{z,i}(t_k)$, the following inequality can be obtained:

$$\begin{aligned} |e_i(t_{k+1})| \quad \leq \quad & (1 + \kappa_i L_{h_i}) \left(\beta_i(|e_i(t_k)|, \Delta) + \gamma_i(\Delta) + L_i \Delta |e_i(t_k)|^2 + \sum_{l \in \mathbb{I}_i,\, l \neq i} \alpha_{i,l}(\Delta, e_{z,l}(t_k), e_{z,l}(t_q^l), \epsilon_l) \right) \\ & + \kappa_i \theta_{v_i} \end{aligned} \tag{3.24}$$

If condition (3.17) is satisfied, from (3.24), it can be obtained that:

$$\begin{aligned} |e_i(t_{k+1})| \quad \leq \quad & (1 + \kappa_i L_{h_i}) \left(g_i(|e_i(t_k)|) + \gamma_i(\Delta) L_i \Delta |e_i(t_k)|^2 + \sum_{l \in \mathbb{I}_i,\, l \neq i} \alpha_{i,l}(\Delta, e_{z,l}(t_k), e_{z,l}(t_q^l), \epsilon_l) \right) \\ & + \kappa_i \theta_{v_i} \end{aligned} \tag{3.25}$$

If there exists $d_{s,i}$ satisfy (3.18) and κ_i is picked following (3.19), then (3.18) holds for all $d_{s,i} \leq |e_i| \leq d_i$. taking into account that $g_i(\cdot)$ is a concave function; that is:

$$|e_i| - (1 + \kappa_i L_{h_i}) \left(g_i(|e_i|) + \gamma_i(\Delta) L_i \Delta |e_i|^2 + \sum_{l \in \mathbb{I}_i,\, l \neq i} \alpha_{i,l}(\Delta, d_l, d_l, \epsilon_l) + L_i \Delta |e_i|^2 \right) - \kappa_i \theta_{v_i} \geq \eta_i \tag{3.26}$$

for all $d_{s,i} \leq |e_i| \leq d_i$ and $|e_l| \leq d_l$ ($l \in \mathbb{I}_i$). From (3.25) and (3.26), it can be obtained that:

$$|e_i(t_{k+1})| \leq |e_i(t_k)| - \eta_i \tag{3.27}$$

for all $d_{s,i} \leq |e_i| \leq d_i$. If $|e_i| \geq d_{s,i}$ for all the time from 0 to t_k, using (3.27) recursively, it can be obtained that:

$$|e_i(t_k)| \leq |e_i(0)| - k\eta_i \tag{3.28}$$

for all $d_{s,i} \leq |e_i(t_k)| \leq d_i$. This implies that $|e_i|$ decreases every sampling time and will become smaller than $d_{s,i}$ in finite steps. Once $|e_{s,i}| < d_{s,i}$, it will remain to satisfy $|e_i(t)| \leq d_{i,\min}$ which is ensured by the definition of $d_{i,\min}$; that is, $\lim_{t\to\infty} \sup |e_i(t)| \leq d_{i,\min}$. Note that the above proof holds for all $i \in \mathbb{I}$.

The ultimate boundedness of subsystem state estimation errors implies the ultimate boundedness of the entire system state estimation error. This can be seen from the inequality $|e| \leq \sum_{i=1}^{m} |e_i|$ which implies that:

$$\lim_{t\to\infty} \sup |e| \leq \sum_{i=1}^{m} d_{i,\min}. \tag{3.29}$$

This proves Theorem 16. □

Remark 17 *Referring to the condition (3.18) in Theorem 16, the term $g_i(|e_i(t_k)|)$ is the upper bound of the estimation error for the nominal subsystem after one sampling time if the initial error term is $|e_i(t_k)|$ when the interactions between subsystems are absent; the term $\gamma_i(\Delta)$ denotes the effect of the sample-and-hold implementation of the nonlinear observer of Eq. (3.2a), and process disturbances and measurement noise; the term $\alpha_{i,l}$ is related to system interactions and triggering thresholds; and the term $b_i\theta_{v_i}$ characterizes the uncertainty introduced into the condition (3.18) due to measurement noise.*

3.5 DMHE with the second communication trigger

3.5.1 Implementation algorithm

In this section, we discuss the second design of the communication trigger as well as the associated distributed state estimation algorithm.

Algorithm 18 *Distributed state estimation algorithm 2*

1. *At $t_0 = 0$, MHE i is initialized with $Y_i(t_0)$, $\hat{x}_i(t_0)$ and $y_l(t_0)$, $\hat{x}_l(t_0)$ for $l \in \mathbb{I}_i$.*

2. *At $t_k > 0$, MHE i and its trigger carry out the following steps:*

   2.1. *MHE i receives the corresponding output and output time derivatives $Y_i(t_k)$.*

   2.2. *MHE i calculates $\hat{x}_i(t_k)$ based on $y_i(t_{k+s-N})$, $s = 0, 1, \ldots, N$, and the latest $\hat{x}_l(t_q^l)$ and $y_l(t_q^l)$ for $l \in \mathbb{I}_i$.*

   2.3. *MHE i sends $Y_i(t_k)$ to its trigger and the trigger checks the triggering condition. If the condition is satisfied, the trigger updates $t_q^i = t_k$ and sends $\hat{x}_i(t_q^i)$ and*

$y_i(t_q^i)$ to subsystem j for all $j \in \mathbb{C}_i$. If the triggering condition is not satisfied, no information is sent out from subsystem i.

3. *Go to Step 2 at the next sampling time t_{k+1}.*

Compared with Algorithm 14, Algorithm 18 has a similar algorithm structure but there are slight differences in Step 1, Step 2.1, and Step 2.3. In these steps, MHE i receives the output and its derivative measurements Y_i and sends Y_i to the corresponding trigger.

3.5.2 The second triggering condition

The second triggering condition for each subsystem is designed based on the difference between the measurement of the current output and its derivatives and the last sent measurement of the output and its derivatives. Specifically, the second triggering condition of MHE i is designed as follows:

$$S_{2,i}(t_k) = \begin{cases} 1, & \text{if } |Y_i(t_k) - Y_i(t_q^i)| \geq \epsilon_{2,i} \\ 0, & \text{if } |Y_i(t_k) - Y_i(t_q^i)| < \epsilon_{2,i} \end{cases} \tag{3.30}$$

where $Y_i(t_k)$ is the current measurement of the output and its derivatives of subsystem i and $Y_i(t_q^i)$ is the last sent information of Y_i. $\epsilon_{2,i}$ is a pre-determined threshold. When $S_{2,i}(t_k) = 0$, the triggering condition is not satisfied and MHE i does not send out information. When $S_{i,2}(t_k) = 1$, the triggering condition is satisfied and MHE i sends out $\hat{x}_i(t_k)$, $y_i(t_k)$ and updates $t_q^i = t_k$.

Note that the main difference between the two algorithms is the design of the triggering conditions. Even though the two triggering conditions may generate different sequences of $t_q^i, i = 1, \ldots, m$, the definition of $z_{n,i}$ and design of the local MHEs presented in Section 3.4.3 apply to both triggering conditions. However, the conditions derived in Theorem 16 are not sufficient to ensure the decrease and ultimately boundedness of the estimation error of the DMHE with the second communication trigger. In the next subsection, we derive another set of sufficient conditions for the DMHE with the second communication trigger.

3.5.3 Stability analysis

In this subsection, we study the stability property of the proposed DMHE implemented following Algorithm 18 and provide a set of sufficient conditions for the decrease and ultimate boundedness of the estimation error.

Proposition 19 *Consider the nonlinear observer of Eq. (3.2a) for subsystem i under the DMHE with output communication of condition (3.30), $i \in \mathbb{I}$, for $t \in [t_k, t_{k+1}]$ with $z_{n,i}(t_k) =$*

$\hat{x}_i(t_k)$, measurement $y_i(t_k)$ and the values of $t_q^l, l \in I_i$ determined following the triggering condition (3.30). If $K_{i,l}$ for $i \in \mathbb{I}$ and $l \in \mathbb{I}_i$ are determined as in (3.3) and $K_{i,l}$ are bounded, then the deviation of $z_{n,i}$ in one sampling time Δ from x_i is bounded for all $x_i \in \mathbb{X}_i$, $i \in \mathbb{I}$, as follows:

$$|e_{z,i}(t_{k+1})| \leq \beta_i(|e_{z,i}(t_k)|, \Delta) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i,\, l \neq i} \lambda_{i,l}(\Delta, |e_{z,l}(t_q^l)|, \delta_l) + L_i \Delta |e_{z,i}(t_k)|^2 \tag{3.31}$$

where $\lambda_{i,l}(\tau, |e_{z,l}(t_q^l)|, \delta_l) = K_{i,l} L_{h_l} \tau \delta_l + \tau (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(\delta_l + |e_{z,l}(t_q^l)|)^2$, $\delta_l = L_{\Phi,l}\,(\epsilon_{2,l} + 2\theta_{\Phi,l})$, and the other constants are defined as in Proposition 15.

Proof: Define $\dot{e}_{z,i}$ and A_i as in Proposition 15, the inequality (3.12) can be rewritten into

$$\begin{aligned} |A_i| \;\leq\; & \sum_{l \in \mathbb{I}_i, l \neq i} (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(|x_l(t_k) - x_l(t_q^l)| + |x_l(t_q^l) - \hat{x}_l(t_q^l)|)^2 \\ & + \mu_i (K_{i,i} H_i^{h_i} + H_i^{\tilde{f}_i}) |x_i(t_k) - \hat{x}_i(t_k)|^2 \end{aligned} \tag{3.32}$$

with the triangle inequality $|x_l(t_k) - \hat{x}_l(t_q^l)| \leq |x_l(t_k) - \hat{x}_l(t_k)| + |\hat{x}_l(t_k) - \hat{x}_l(t_q^l)|$, and $|X_i(t_k) - \hat{X}_i(t_k)|^2 = \sum_{l \in \mathbb{I}_i, l \neq i} |x_l(t_k) - \hat{x}_l(t_q^l)|^2 + \mu_i |x_i(t_k) - \hat{x}_i(t_k)|^2$.

From (3.8) and the Lipschitz property of $\tilde{f}_i$ with respect to x_l ($l \in \mathbb{I}_i$), it is obtained that:

$$\begin{aligned} |\dot{e}_{z,i}(t)| \;\leq\; & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i} |x_i(t) - x_i(t_k)| + L_{F_i}^{y_i} \theta_{v_i} + L_{f_i}^{w_i} \theta_{w_i} \\ & + \sum_{l \in \mathbb{I}_i} K_{i,l} \theta_{v_l} + \sum_{l \in \mathbb{I}_i} L_{\tilde{f}_i}^{x_l} |x_l(t) - x_l(t_k)| + \sum_{l \in \mathbb{I}_i, l \neq i} K_{i,l} L_{h_l} |x_l(t_k) - x_l(t_q^l)| \\ & + \sum_{l \in \mathbb{I}_i, l \neq i} (K_{i,l} H_l^{h_l} + H_i^{\tilde{f}_i})(|x_l(t_k) - x_l(t_q^l)| + |e_{z,l}(t_q^l)|)^2 + \mu_i (K_{i,i} H_i^{h_i} + H_i^{\tilde{f}_i}) |e_{z,i}(t_k)|^2 \end{aligned} \tag{3.33}$$

with $e_{z,i}(t_k) = x_i(t_k) - \hat{x}_i(t_k)$, $e_{z,l}(t_k) = x_l(t_k) - \hat{x}_l(t_k)$ and $e_{z,l}(t_q^l) = x_l(t_q^l) - \hat{x}_l(t_q^l)$.

Following the triggering condition of Eq. (3.30), it can be written for all $l \in \mathbb{I}_i$ that:

$$|Y_l(t_k) - Y_l(t_q^l)| \leq \epsilon_{2,l} \tag{3.34}$$

From the definition of Y in Section 3.2, it can be obtained that:

$$x_l(t_k) = \Phi^{-1}(Y_l(t_k) - \phi_l(t_k)) \tag{3.35}$$

Based on the Lipschitz property of Φ, the following inequality can be derived:

$$|x_l(t_k) - x_l(t_q^l)| \leq L_{\Phi,l} \left(|Y_l(t_k) - Y_l(t_q^l)| + |\phi_l(t_k) - \phi_l(t_q^l)| \right) \tag{3.36}$$

where $L_{\Phi,l}$ is the Lipschitz constant of Φ^{-1}. Taking into account the boundedness of $\phi(t_k)$ and constraint (3.34), it's obtained that:

$$|x_l(t_k) - x_l(t_q^l)| \leq \delta_l \tag{3.37}$$

with $\delta_l = L_{\Phi,l}\,(\epsilon_{2,l} + 2\theta_{\Phi,l})$.

Applying constraint (3.37), taking into account the boundedness of the system state and condition (2.6), integrating (3.33) from $t = t_k$ to $t = t_{k+1}$, the following inequality can be obtained:

$$\begin{aligned} |e_{z,i}(t_{k+1})| \;\leq\; & \beta_i(|e_{z,i}(t_k)|, \Delta) + L_{F_i}^{y_i} L_{h_i} M_i \Delta^2/2 + L_{F_i}^{y_i}\theta_{v_i}\Delta + L_{f_i}^{w_i}\theta_{w_i}\Delta + \sum_{l\in\mathbb{I}_i} M_{K_{i,l}}\theta_{v_l}\Delta \\ & + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l}\Delta^2/2 + \sum_{l\in\mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \tau^2/2 + \sum_{l\in\mathbb{I}_i,\, l\neq i} K_{i,l} L_{h_l}\Delta\delta_l \\ & + \sum_{l\in\mathbb{I}_i,\, l\neq i} \Delta(K_{i,l}H_l^{h_l} + H_i^{\tilde{f}_i})(\delta_l + |e_{z,l}(t_q^l)|)^2 + L_i|e_{z,i}(t_k)|^2\Delta \end{aligned} \tag{3.38}$$

if M_i, $M_{K_{i,l}}$, $\gamma_i(\tau)$, L_i are defined as in Proposition 15 and 19. (3.38) can be written in the form of Eq. (3.31), which proves Proposition 19. □

Theorem 20 *Consider system (2.4) with Y_i sampled at $\{t_{k\geq 0}\}$. If the proposed DMHE implemented following Algorithm 2 based on the triggering condition (3.30) with subsystem MHE designed as in (3.4) based on deterministic nonlinear observers satisfying (2.6) and $K_{i,l}$ determined following (3.3), and if there exist concave functions $g_i(\cdot)$, $i \in \mathbb{I}$, as defined in (3.17) for all $|e_i| \leq d_{2,i}$, constants $d_{s_2,i}$, and $d_{2,i}$ such that $0 \leq d_{s_2,i} \leq d_{2,i}$, and positive constants $a_{2,i} \geq 1$, $b_{2,i} > 0$, and $\eta_{2,i} > 0$, such that:*

$$d_{s_2,i} - a_{2,i}\left(g_i(d_{s_2,i}) + \gamma_i(\Delta) + \sum_{l\in\mathbb{I}_i,\, l\neq i} \lambda_{i,l}(\Delta, d_l, \delta_l) + L_i\Delta d_{s_2,i}^2\right) - b_{2,i}\theta_{v_i} \geq \eta_{2,i} \tag{3.39}$$

for all $i \in \mathbb{I}$, and if κ_i for all $i \in \mathbb{I}$, are picked as follows:

$$0 \leq \kappa_i \leq \min\{(a_{2,i} - 1)/L_{h_i}, b_{2,i}\}, \tag{3.40}$$

then the estimation error $|e_i| = |\hat{x}_i - x_i|$ ($i \in \mathbb{I}$) is a decreasing sequences if $|e_i(0)| \leq d_{2,i}$ for all $i \in \mathbb{I}$ and the whole system is ultimately bounded as follows:

$$\lim_{t\to\infty} \sup |e| \leq \sum_{i=1}^{m} d_{i,\min_2} \tag{3.41}$$

for $i \in \mathbb{I}$ with $d_{i,\min_2} = \max\{|e_i(t + \Delta)| : |e_i(t)| \leq d_{s_2,i}\}$ for all $e_i(0) \leq d_{2,i}$ and $x_i \in \mathbb{X}_i$.

Proof: Similar to the proof in Theorem 16, to ensure the ultimate boundedness of the estimation error for the whole system, we only need to prove that there exists a constraint that makes the estimation error of each subsystem is decreasing and ultimately bounded in a small region (i.e., $|e_i(t_k)| \leq |e_i(0)| - k\eta_{2,i}$). Considering MHE i, $i \in \mathbb{I}$, for $t \in [t_k,\ t_{k+1}]$, from Proposition 19 and (3.23), and if condition (3.17) is satisfied, it can be obtained that:

$$|e_i(t_{k+1})| \leq (1+\kappa_i L_{h_i}) \left(g_i(|e_i(t_k)|) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i,\, l \neq i} \lambda_{i,l}(\Delta, |e_{z,l}(t_q^l)|, \delta_l) + L_i \Delta |e_i(t_k)|^2 \right) + \kappa_i \theta_{v_i} \tag{3.42}$$

If there exists $d_{s_{2,i}}$ satisfy (3.39) and κ_i is picked following (3.40), then (3.39) holds for all $d_{s_{2,i}} \leq |e_i| \leq d_{2,i}$. taking into account that $g_i(\cdot)$ is a concave function; that is:

$$|e_i| - (1+\kappa_i L_{h_i}) \left(g_i(|e_i|) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i,\, l \neq i} \lambda_{i,l}(\Delta, |e_l|, \delta_l) + L_i \Delta |e_i|^2 \right) - \kappa_i \theta_{v_i} \geq \eta_{2,i} \tag{3.43}$$

for all $d_{s_{2,i}} \leq |e_i| \leq d_{2,i}$ and $|e_l| \leq d_{2,l}$ ($l \in \mathbb{I}_i$). From (3.42) and (3.43), we get:

$$|e_i(t_{k+1})| \leq |e_i(t_k)| - \eta_{2,i} \tag{3.44}$$

for all $d_{s_{2,i}} \leq |e_i| \leq d_{2,i}$. Following the proving procedure in Theorem 16, it proves Theorem 20. □

Remark 21 *Considering the condition (3.39) in Theorem 20, the terms $g_i(|e_i(t_k)|)$, $\gamma_i(\Delta)$, $b_i\theta_{v_i}$ in the left hand side have already been explained in Remark 17; the term $\lambda_{i,l}$ is related to system interactions and triggering thresholds $\epsilon_{2,l}$.*

Remark 22 *Referring to the triggering conditions provided in Proposition 15 and Proposition 19, they are both used in bounding the initial estimation errors of subsystem i as well as its associated subsystems. However, their influences are quite different: condition (3.1) provides an upper bound on the difference between the current and last updated state estimate, while condition (3.30) bounds the deviation of the current augmented output measurement from its last sent one, which can be reduced to the actual state difference during two consecutive communication instants and is not directly related to the estimation error.*

3.6 Application to the reactor-separator process

3.6.1 Simulation settings

A detailed modeling of the reactor-separator process can be found in Section 2.6. Based on the estimates of the local MHEs, we design the first triggering condition as discussed in

Section 3.4.2. In order to design the second triggering condition, we need time derivatives of the temperature measurements. In the simulations, we use a finite difference method to approximate the change of temperatures. Specifically, in the design of the second triggering condition, $Y_i(t_k)$, is evaluated as follows:

$$Y_i(t_k) = [T_i(t_k),\ T_i(t_k) - T_i(t_{k-1}),\ T_i(t_k) - 2T_i(t_{k-1}) + T_i(t_{k-2})]^T$$

where $i = 1, 2, 3$.

3.6.2 Simulation results

In this section, the proposed DMHE with the two different communication triggers are compared to illustrate their performances from a communication cost point of view. First we carried out a set of simulations when the the thresholds of the two triggering conditions are $\epsilon_i = 1.0$ and $\epsilon_{2,i} = 1.0$, $i = 1, 2, 3$, respectively. In this set of simulations, the initial state of the process is

$$x_0 = [0.1939; 0.7404; 528.3482\ K; 0.2162; 0.7190; 520.0649\ K; 0.0716; 0.7373; 522.3765\ K]^T$$

and the initial guess in MHEs is

$$\hat{x}_0 = [0.1675; 0.7; 500.3\ K; 0.18; 0.67; 500\ K; 0.06; 0.68; 500\ K]^T$$

with input $Q_i = Q_{s,i} + 10^9 e^{-0.01t} sin(0.1t)$ as shown in Figure 3.2. This type of inputs is used to excite different change rates in the process state trajectories.

Figures 3.3 - 3.7show the simulation results. Figures 4 and 6 show the estimated state trajectories obtained under the two different communication triggering conditions. Figures 5 and 7 show the time instants that the subsystems send out information under the two triggering conditions. Figure 8 shows the evolution of the estimation errors under the two triggering conditions. It can be seen from these figures that: 1) the proposed DMHE with the two triggered communication approaches can track the actual system states very well (Figs. 3.3 and 3.5); 2) the estimation errors decrease to values close to zero quickly and maintain close to zero (Fig. 3.7); 3) the number of information exchange between the subsystems is significantly reduced under both the two triggering conditions (Figs. 3.4 and 3.6); 4) the communication between the subsystem is more often when the states of the process change fast than when the states of the process have small changes.

We also conducted another set of simulations to compare the two communication algorithms in terms of the mean performance and number of communications between subsystems with varying triggering thresholds. Specifically, simulations were carried out under

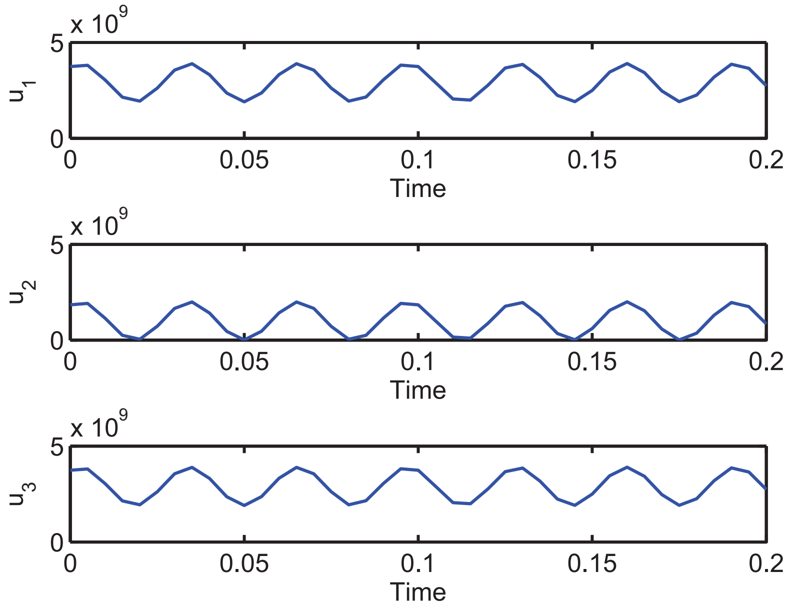


Figure 3.2: Damped sinusoidal inputs to the three subsystems.

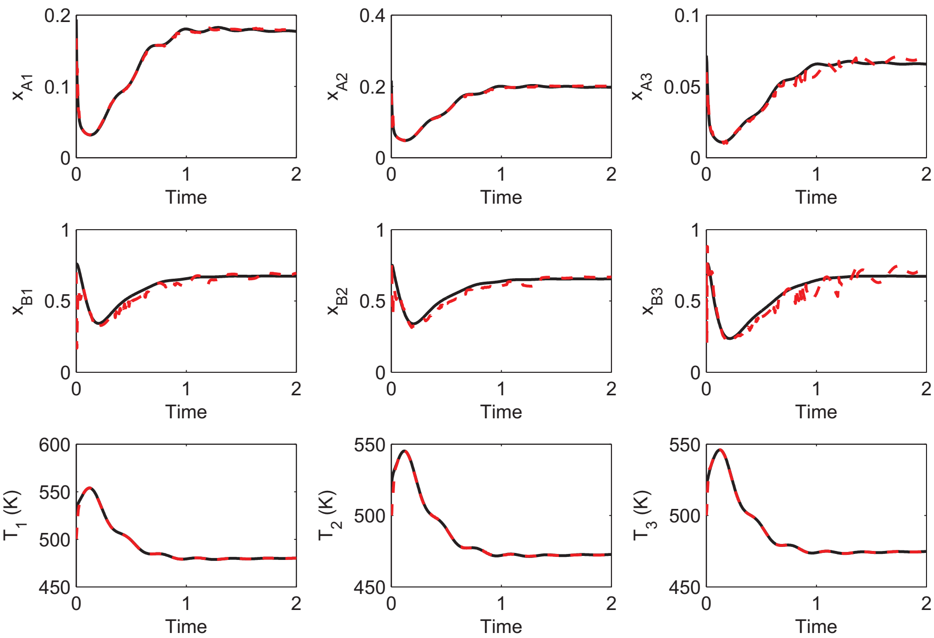


Figure 3.3: State trajectories of the actual system state(solid lines) and the state estimates given by the proposed DMHE implemented following Algorithm 14 based on triggering condition (3.1) with $\epsilon_i = 1.0, i = 1, 2, 3$ (dashed lines).

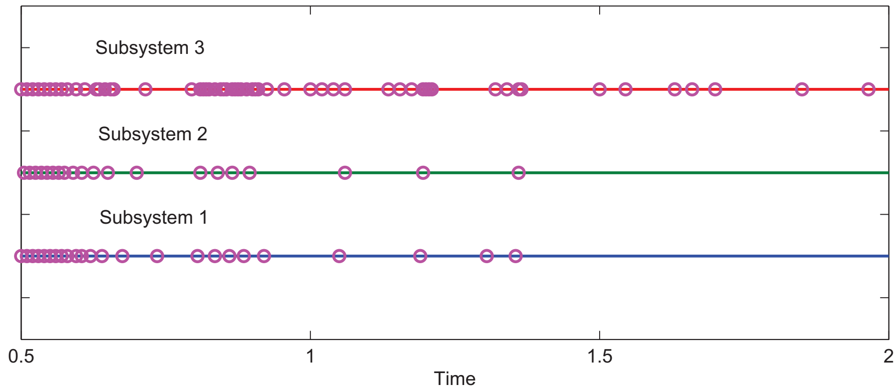


Figure 3.4: Time instants when subsystem i, $i = 1, 2, 3$, sent out its information by the proposed DMHE implemented following Algorithm 14 based on triggering condition (3.1) .

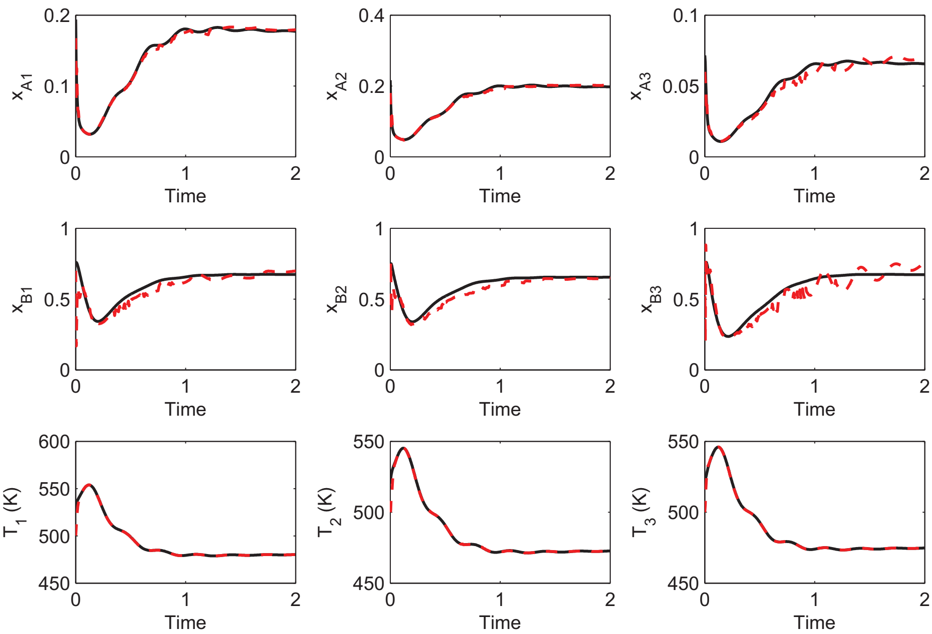


Figure 3.5: State trajectories of the actual system state(solid lines) and the state estimates given by the proposed DMHE implemented following Algorithm 18 based on triggering condition (3.30) with $\epsilon_{2,i} = 1.0, i = 1, 2, 3$ (dashed lines).

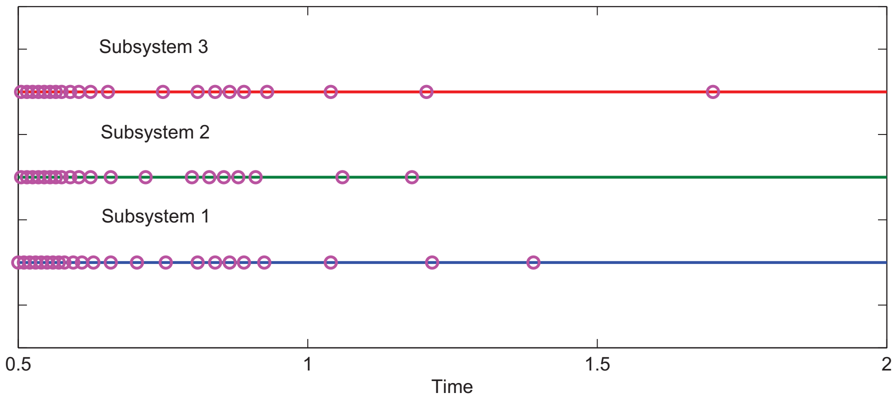


Figure 3.6: Time instants when subsystem i, $i = 1, 2, 3$, sent out its information by the proposed DMHE implemented following Algorithm 18 based on triggering condition (3.30) .

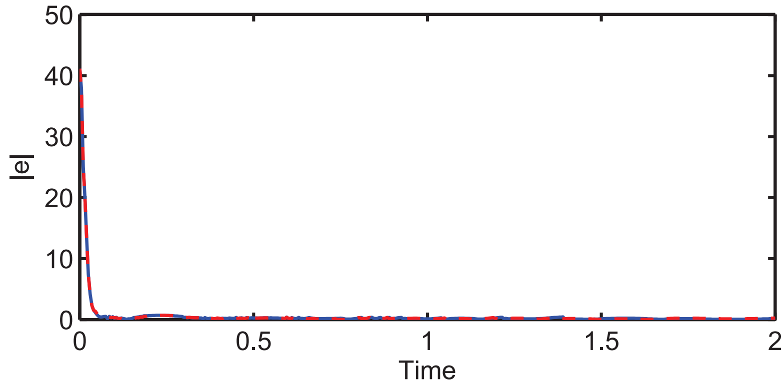


Figure 3.7: Trajectories of the estimation error norm of the proposed DMHE implemented following Algorithm 14 (dashed lines) based on triggering condition (3.1) and Algorithm 18 based on triggering condition (3.30) (solid lines).

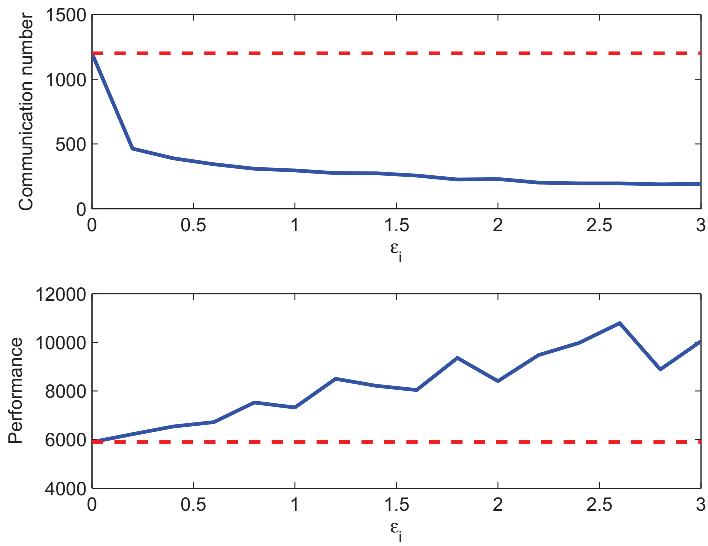


Figure 3.8: Average number of communications and performance index of the proposed DMHE implemented following Algorithm 14 based on triggering condition (3.1) with $\epsilon_i, i = 1, 2, 3$, varying from 0 to 3 (solid lines) and the dashed lines denote the number of communications and performance of the proposed DMHE with the subsystems exchanging information every sampling time.

different settings (initial conditions, random noise sequences and triggering thresholds) and the performance index for MHE i is designed as follows:

$$J_i = \sum_{k=0}^{M} |\hat{x}_i(t_k) - x_i(t_k)|^2_{Q_i^*} \tag{3.45}$$

where $i = 1, 2, 3$, $t_0 = 0$ is the initial simulation time and $t_M = 2.0\ h$ is the end of the simulation time. The parameters Q_i^*, $i = 1, 2, 3$, are the factors to compensate for the different orders of magnitude of the states, and $Q_i^* = \text{diag}[10^3, 10^3, 1], i = 1, 2, 3$. The overall performance is measured by $J = J_1 + J_2 + J_3$. For each value of the triggering threshold, 10 simulation runs were used to calculate the average performance and number of communications between subsystems. Figures 3.8 and 3.9 show the simulation results. From these figures, it can be seen that: 1) as the threshold increases, the number of communications between the subsystems gets worse; 2) as the threshold increases, the performance of the DMHE also decreases; 3) both of the two triggering strategies give similar trends. These results imply that a balance between the number of communication and the estimation performance should be reached for a specific application. From the above simulation results, it can be seen that the two algorithms both can lead to significant reduction in the number of communications and maintain the estimation performance close to the case when

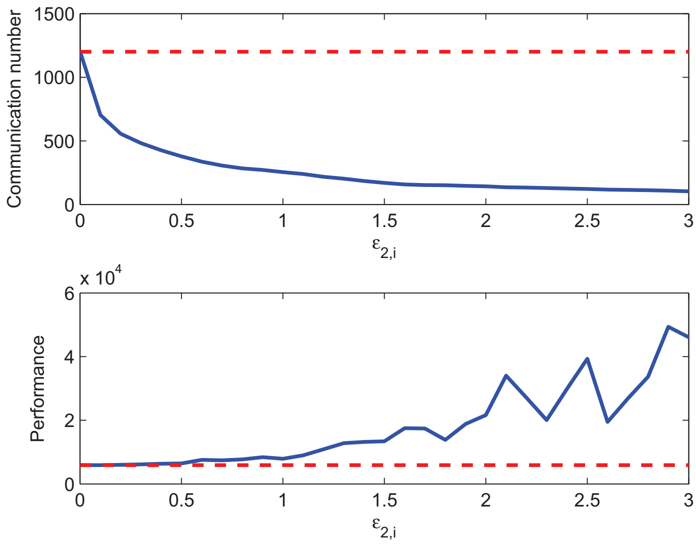


Figure 3.9: Average number of communications and performance index of the proposed DMHE implemented following Algorithm 18 based on triggering condition (3.30) with $\epsilon_{2,i}, i = 1, 2, 3$, varying from 0 to 3 (solid lines) and the dashed lines denote the number of communications and performance of the proposed DMHE with the subsystems exchanging information every sampling time.

information is exchanged between the subsystem every sampling time if the thresholds are chosen properly. In this application, the second algorithm based on triggering condition (3.30) shows slightly superior than the first algorithm based on triggering condition (3.1) in terms of reducing the number of communications while maintaining the performance (see Figs. 3.4, 3.6, 3.8, 3.9). However, this approach requires the availability of the time derivatives of the output measurements which may be difficult or expensive to obtain in certain applications.

3.7 Conclusions

In this chapter, we presented two triggered communication algorithms for a DMHE scheme to reduce the communication cost via decreasing the number of communications between subsystems. In the first algorithm, the communication between subsystems is triggered by the difference between the current subsystem state estimates and last transmitted state estimations; in the second algorithm, the communication is triggered by the difference between the current measurement of subsystem outputs and output derivatives and corresponding last transmitted values. Sufficient conditions under which the convergence and boundedness of the estimation error for the whole system were derived. The proposed two algorithms were compared from a performance and number of communication times point of view via the application to a chemical process. Both proposed communication algorithms are capable of maintaining the estimation performance as well as greatly reducing the number of communications with proper triggering thresholds.

Chapter 4

Distributed Moving Horizon State Estimation Subject to Communication Delays*

4.1 Introduction

In Chapter 2, an observer-enhanced DMHE design was developed for a class of nonlinear systems with bounded process uncertainties. In this DMHE, each subsystem MHE communicates with subsystems that it interacts with every sampling time. In the design of each subsystem MHE, an auxiliary deterministic nonlinear observer is taken advantage of to calculate a confidence region that contains the actual system state every sampling time. The subsystem MHE is only allowed to optimize its state estimate within the confidence region. This strategy was demonstrated to guarantee the convergence and ultimate boundedness properties of the estimation error. However, the results in Chapter 2 were derived under the assumption that the communication between subsystems is flawless and there is no delay in the information transmission. In practice, this assumption may not hold especially when shared wireless communication network is used. Issues brought into the design by communication need to be carefully addressed [69]. Motivated by the above considerations, in this chapter we proposed a DMHE scheme that is able to handle time-varying communication delays in the DMHE network of Chapter 2. In the proposed design, a nonlinear observer-enhanced MHE is designed for each subsystem and the distributed MHEs are assumed to be able to communicate and exchange information with each other via a shared communication network which may introduce communication delays. To handle time-varying delays in the communication, the implementation algorithm and local MHE design in Chapter 2 need

*This chapter is a version of "J. Zhang and J. Liu. Observer-enhanced distributed moving horizon state estimation subject to communication delays. Journal of Process Control, in press."

to be revised to account for the communication delays between the subsystems in order to achieve boundedness of the estimation error.

4.2 Modeling of measurements and communications

In this chapter, it's assumed that each subsystem observer has direct and immediate access to its local output measurements. We also assume that the subsystems can communicate to exchange information via a shared communication network and the information transmitted in the communication network is subject to time-varying communication delays. It is assumed that output measurements of the m subsystems, y_i, $i \in \mathbb{I}$ are sampled synchronously and periodically at time instants $t_k = t_0 + k\Delta$, where t_0 is the initial sampling time, Δ is a fixed sampling interval. To model delays in the communication, an auxiliary variable $d_{i,j}(t_k)$, is introduced to indicate the delay associated with the information of subsystem j available to subsystem i at time instants t_k. The variable $d_{i,j}(t_k)$ takes values that are positive integers. For example, if at time t_k, the latest information of subsystem j received by subsystem i was sent at t_{k-q}, then $d_{i,j}(t_k) = q$. All the information are time-labeled, so the delays are known in the communication network. In order to study the deterministic stability property of the proposed distributed state estimation scheme, we assume that there is an upper bound D on the delay $d_{i,j}(t_k)$. Since delays are time-varying, it is possible that no new information is provided within two consecutive sampling periods. A subsystem stores all the received information that was sent within a time period of $D\Delta$ from the current time instant. At time t_k, if the latest information of subsystem j received by subsystem i was sent at $t_{k-d_{i,j}(t_k)}$, a data package containing newer information about subsystem j will be received by $t_{k+D-d_{i,j}(t_k)}$ because the maximum possible delay is $D\Delta$.

4.3 The DMHE scheme subject to communication delays

In this chapter, we present the proposed DMHE subject to communication delays. A schematic of the proposed design is shown in Figure 4.1. In this design, each subsystem has its own local state estimator and sends information to all the other subsystems when the current state estimate and output measurement are available. However, due to communication delays in information transmission, a subsystem may not receive the information from other subsystems synchronously. To address this issue, a predictor is embedded in each subsystem to predict the states of other subsystems in open-loop based on previously received information.

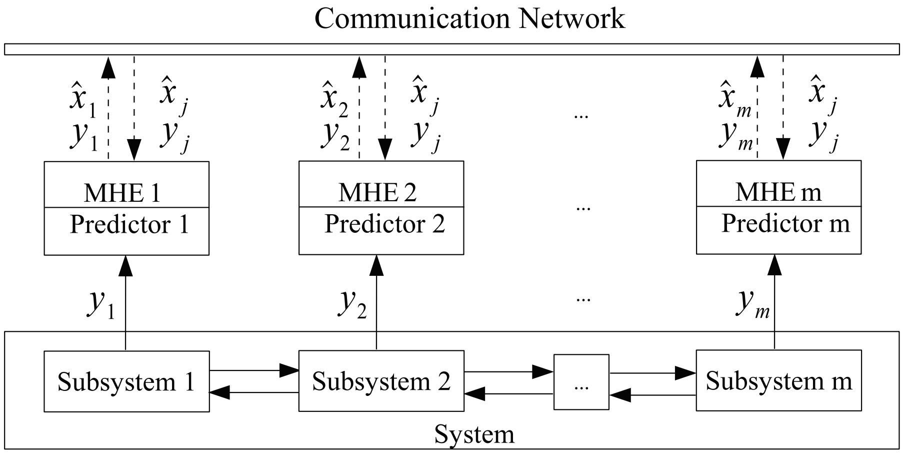


Figure 4.1: Scheme of the proposed DMHE design considering communication delays.

In the rest of this chapter, we first present a distributed state estimation algorithm accounting for communication delays; and then we discuss the design of the open-loop predictor for each subsystem; subsequently, the calculation of confidence regions for the state estimates is described which is followed by the design of local MHEs taking into account communication delays explicitly; finally, the stability properties of the proposed DMHE are analyzed.

4.3.1 Distributed state estimation algorithm

The proposed DMHE that is capable of handling communication delays is implemented using the following algorithm:

Algorithm 23 *Distributed state estimation algorithm*

1. *At* $t_0 = 0$*, MHE* i*,* $i \in \mathbb{I}$*, is initialized with system measurement at the initial time* $y(t_0)$*, and the initial system state guess* $\hat{x}(t_0)$*.*
2. *At* $t_k > 0$*, carry out the following steps:*

    2.1. *MHE* i *receives the local output measurement* $y_i(t_k)$*.*

    2.2. *If subsystem* i *receives any new data packages between the time interval* t_{k-1} *and* t_k*, it checks if they provide more recent information. If it does, update* $d_{i,j}(t_k)$ *values and store the data packages ; Else, store the data packages.*

    2.3. *If any* $d_{i,j}(t_k)$ *is greater than 1, the corresponding predictor of MHE* i *predicts the system state* $x^{p,i}(t_{k-1})$ *based on previously received information. Based on the* *local measurement $y_i(t_k)$, the state prediction $x^{p,i}(t_{k-1})$ and the received information from other subsystems, a reference state estimate for subsystem i, $z_{n,i}(t_k)$, is calculated.*

2.4. MHE i calculates current state estimate $\hat{x}_i(t_k)$ based on $z_{n,i}(t_k)$ and sends the current information $\hat{x}_i(t_k)$ to all the other subsystems and $y_i(t_k)$ to subsystem j for all $j \in \mathbb{I}_i$, $j \neq i$.

3. Go to Step 2 at the next sampling time t_{k+1}.

From the above algorithm, it can be seen that information is sent out at Step 2.4 and the information will be used at a future sampling time. This implies that the delay of one sampling time is unavoidable, i.e., $d_{i,j}(t_k) > 0$; and data transmission may cause additional delays. If the total delay is greater than one sampling time (i.e., $d_{i,j}(t_k) > 1$), the predictor will be used to calculate a prediction of the state at t_{k-1} and based on the prediction, a subsystem MHE calculates the current subsystem state estimate. Compared with Algorithm 2 in Chapter 2, Algorithm 23 accounts for communication delays explicitly. Another important difference between Algorithm 23 and Algorithm 2 in Chapter 2 is that in Algorithm 23, an all-to-all communication is required whereas in Algorithm 2 an estimator only has to send information to a subset of the subsystems. The all-to-all communication is needed because a centralized model is used in the design of the open-loop predictors.

4.3.2 State prediction

Consider subsystem i, $i \in \mathbb{I}$, at time t_k, when $d_{i,l}(t_k) > 1$ for any $l \in \mathbb{I}_i$ (i.e., $\hat{x}_l(t_{k-1})$ has not been received by subsystem i due to communication delays), the nominal centralized system model of Eq. (2.4) is used to generate a prediction of the entire system state at t_{k-1}, $x^{p,i}(t_{k-1})$. In the notation $x^{p,i}$, the superscript 'p' means prediction while 'i' means that the prediction is calculated in subsystem i. The unreceived subsystem states will be approximated by the subsystem states in $x^{p,i}(t_{k-1})$. The value of $x^{p,i}(t_{k-1})$ is calculated as follows:

$$\dot{x}^{p,i}(t) = f(x^{p,i}(t), 0) + \tilde{f}(x^{p,i}(t)),\ t \in [t_s, t_{s+1}] \tag{4.1}$$

with the initial condition $x^{p,i}(t_q) = \hat{x}(t_q)$, where $s = q, q+1, ..., k-2$ and t_q is the latest time instant that all of the other subsystem state estimates have been received so that $\hat{x}(t_q)$ is available. In the worst case, $t_q = t_{k-D}$ which is ensured because the maximum transmission delay is D. Note that in evaluating $x^{p,i}$ recursively, its value should be updated with any received state estimates between t_q and t_{k-1} in order to get better state prediction. For example, at time t_s, $q < s < k-1$, once $x^{p,i}(t_s)$ is obtained from (4.1), we should check if there are any received state estimates at t_s of all the subsystems. If there are any, $x^{p,i}(t_s)$ should be updated with those state estimates and the updated $x^{p,i}(t_s)$ will be used to evaluate the prediction at the next sampling time. Note that the predictions calculated using (4.1) are obtained in open-loop in the sense that no output measurements have been used in the calculation.

4.3.3 Reference state estimate calculation

For subsystem i, $i \in \mathbb{I}$, an augmented observer based on observer (2.5) is designed. In this design, estimates received without delays or predictions of the subsystem states involved in X_i will be used to calculate an approximation of the interaction between subsystem i and other subsystems. We assume that at t_k, the information sent at t_{k-1} by subsystem g, $g \in \mathbb{Z}_i(t_k) \subset \mathbb{I}_i$, is received without delay (i.e., $d_{i,g}(t_k) = 1$, $g \in \mathbb{Z}_i(t_k)$) and the information sent at t_{k-1} by subsystem l, $l \in \mathbb{I}_i \backslash \mathbb{Z}_i(t_k)$ is delayed. Note that the set $\mathbb{Z}_i(t_k)$ is a function of time. At t_k, the main purpose of this augmented observer is to calculate a reference state estimate $z_{n,i}(t_k)$ for subsystem i. The observer is designed as follows:

$$\dot{z}_{n,i}(t) = F_i(z_{n,i}(t), y_i(t_{k-1})) + \tilde{f}_i(\hat{X}_i(t_{k-1})) + \sum_{g \in \mathbb{Z}_i(t_k)} K_{i,g}(\hat{x}_g(t_{k-1}))(y_g(t_{k-1}) - h_g(\hat{x}_g(t_{k-1}))) \tag{4.2a}$$

$$z_{n,i}(t_{k-1}) = \hat{x}_i(t_{k-1}) \tag{4.2b}$$

where $z_{n,i}(t)$ is the state of this augmented observer, $\hat{X}_i(t_{k-1})$ is an approximation of $X_i(t)$ for $t \in [t_{k-1}, t_k)$ and is composed of $x_l^{p,i}(t_{k-1})$ for $l \in \mathbb{I}_i \backslash \mathbb{Z}_i(t_k)$ and $\hat{x}_g(t_{k-1})$ for $g \in \mathbb{Z}_i(t_k)$. Note that $x_l^{p,i}$ is the portion of $x^{p,i}$ that corresponds to subsystem l. Moreover, $K_{i,g}$ for $g \in \mathbb{Z}_i(t_k)$ are gain matrices. Observer (4.2) should be evaluated before the evaluation of MHE i at t_k to generate a reference state estimate as specified in Step 2.3 in Algorithm 23.

In nonlinear observer (4.2a), the first term of the right hand side comes from nonlinear observer (2.5), the second term explicitly describe the interactions between subsystem i and other subsystems based on the interaction model $\tilde{f}_i(X_i)$ and the last term contains corrections used to compensate for the error in the interaction model. Note that only measurements received without delay are used in the corrections. The gains $K_{i,g}$, $g \in \mathbb{Z}_i(t_k)$, in nonlinear observer (4.2a) are determined as follows:

$$K_{i,g}(\hat{x}_g(t_{k-1})) = \left.\frac{\partial \tilde{f}_i}{\partial x_g}\left(\frac{\partial h_g}{\partial x_g}\right)^{+}\right|_{x_g = \hat{x}_g(t_{k-1})}, \ \forall g \in \mathbb{Z}_i(t_k) \tag{4.3}$$

The correction gains are picked to compensate for the interaction model mismatch (i.e., the difference between $\tilde{f}_i(\hat{X}_i(t_{k-1}))$ and $\tilde{f}_i(X_i(t)))$ via its linear approximation obtained by Taylor series expansion. This point will be made clear in the proof of Proposition 26.

Remark 24 *In the evaluation of the reference state estimate, $z_{n,i}(t_k)$, in Eq. (4.2a), the output measurements and subsystem state estimates/predictions at t_{k-1} are used to approximate the output and the subsystem states over the time period $t \in [t_{k-1}, t_k)$ because system output measurements and state estimates are only available at periodic sampling time instants. For example, the local output $y_i(t)$ is approximated by $y_i(t_{k-1})$ for $t \in [t_{k-1},\ t_k)$.*

4.3.4 Subsystem MHE design

Based on the reference state estimate $z_{n,i}(t_k)$ provided by observer (4.2), the proposed design of local MHE for subsystem i at time instant t_k subject to communication delays is formulated as follows:

$$\min_{\tilde{x}_i(t_{k-N}),\ldots,\tilde{x}_i(t_k)} \sum_{q=k-N}^{k-1} |w_i(t_q)|^2_{Q_i^{-1}} + \sum_{q=k-N}^{k} |v_i(t_q)|^2_{R_i^{-1}} + V_i(\tilde{x}_i(t_{k-N})) \tag{4.4a}$$

$$\text{s.t. } \dot{\tilde{x}}_i(t) = f_i(\tilde{x}_i(t), w_i(t_q)) + \tilde{f}_i(\hat{X}_i(t_q)),\ t \in [t_q, t_{q+1}],\ q = k-N, \ldots, k-1 \tag{4.4b}$$

$$v_i(t_q) = y_i(t_q) - h_i(\tilde{x}_i(t_q)),\ q = k-N, \ldots, k \tag{4.4c}$$

$$w_i(t_q) \in \mathbb{W}_i,\ v_i(t_q) \in \mathbb{V}_i,\ \tilde{x}_i(t_q) \in \mathbb{X}_i,\ q = k-N, \ldots, k-1 \tag{4.4d}$$

$$|\tilde{x}_i(t_k) - z_{n,i}(t_k)| \leq \kappa_i |y_i(t_k) - h_i(z_{n,i}(t_k))| \tag{4.4e}$$

where $\tilde{x}_i$ is the predicted x_i in the above optimization problem, Q_i and R_i are positive definite covariance matrices of w_i and v_i respectively, $V_i(\tilde{x}_i(t_{k-N}))$ is the arrival cost that summarizes past information up to t_{k-N}, $\hat{X}_i$ is the best estimate of X_i at previous time instants, N is the estimation horizon, and κ_i is a positive constant. The roles of these parameters will be made clear in the following discussion. The optimal solution to problem (4.4) is denoted as $\tilde{x}_i^*(t_{k-N}), \ldots, \tilde{x}_i^*(t_k)$, and only the last element $\tilde{x}_i^*(t_k)$ is used as the current optimal estimate of the state of subsystem i at t_k and is denoted as $\hat{x}_i(t_k)$. That is,

$$\hat{x}_i(t_k) = \tilde{x}_i^*(t_k). \tag{4.5}$$

To ensure the optimization problem (4.4) is a finite dimensional one, w_i and v_i are assumed to be piece-wise constant variables between two consecutive time instants.

In the above design, (4.4a) is the cost function to be minimized. Constraints (4.4b)-(4.4c) are from the subsystem model of Eq. (2.1), and $f_i(\hat{X}_i(t_q))$ is used to approximate the function $f_i(\hat{X}_i(t)),\ t \in [t_q, t_{q+1}]$. To this end, each MHE should be capable of storing the previously received information of other subsystems within the estimation horizon. Constraint (4.4d) contains constraints on process disturbances, measurement noise and system state. Constraint (4.4e) is used to calculate a confidence region (i.e., $\kappa_i|y_i(t_k) - h_i(z_{n,i}(t_k))|$) by taking advantage of the current output measurement $y_i(t_k)$ and the reference state estimate $z_{n,i}(t_k)$. The estimate of the current subsystem state is only allowed to be optimized within the confidence region.

Note that constraint (4.4e) is imposed at t_k only and constraint (4.4d) is imposed from t_{k-N} to t_{k-1}. Since in the calculation of the confidence region the boundedness properties of the process disturbance, measurement noise and the system state are taken into account, it is not necessary to impose constraint (4.4d) at t_k. Essentially, constraints (4.4d) and (4.4e) do not conflict with each other. The reference state estimate, $z_{n,i}(t_k)$, provided by nonlinear observer (4.2) is always a feasible solution to constraint (4.4e).

4.3.5 Stability analysis

In this section, we study the stability properties of the proposed DMHE subject to communication delays implemented following Algorithm 23. Proposition 25 below provides an upper bound on the deviation of the state trajectory obtained with the nominal system model from the actual state trajectory.

Proposition 25 *Consider the following state trajectories:*

$$\begin{aligned} \dot{x}_a(t) &= f(x_a(t), w(t)) + \tilde{f}(x_a(t)) \\ \dot{x}_b(t) &= f(x_b(t), 0) + \tilde{f}(x_b(t)) \end{aligned} \tag{4.6}$$

then the following inequality holds for all $x_a(t), x_b(t) \in \mathbb{X}$, $w(t) \in \mathbb{W}$:

$$|x_a(t) - x_b(t)| \leq f_W(t - t_0, |x_a(0) - x_b(0)|) \tag{4.7}$$

where $f_W(\tau, |e(0)|) = \dfrac{L_f^w \theta_w}{L_f^x + L_{\tilde{f}}^x}(e^{(L_f^x + L_{\tilde{f}}^x)\tau} - 1) + |e(0)|e^{(L_f^x + L_{\tilde{f}}^x)\tau}$ with L_f^w, L_f^x being Lipschitz constants of f with respect to w and x respectively and $L_{\tilde{f}}^x$ being a Lipschitz constant of $\tilde{f}$ with respect to x.

Proof: Define the error term $e(t) = x_a(t) - x_b(t)$. The time derivative of the error is given by:

$$\dot{e}(t) = f(x_a(t), w(t)) - f(x_b(t), 0) + \tilde{f}(x_a(t)) - \tilde{f}(x_b(t)) \tag{4.8}$$

Applying the Lipschitz property of f, $\tilde{f}$ and the boundeness of w such that $|w| \leq \theta_w$, it is obtained that:

$$|\dot{e}(t)| \leq L_f^w \theta_w + (L_f^x + L_{\tilde{f}}^x)|e(t)| \tag{4.9}$$

for all $x_a(t), x_b(t) \in \mathbb{X}$, $w(t) \in \mathbb{W}$. Integrating Eq. (4.9) with the initial condition $|e(0)| = |x_a(0) - x_b(0)|$, the following inequality holds:

$$|e(t)| \leq \frac{L_f^w \theta_w}{L_f^x + L_{\tilde{f}}^x}(e^{(L_f^x + L_{\tilde{f}}^x)(t-t_0)} - 1) + |x_a(0) - x_b(0)|e^{(L_f^x + L_{\tilde{f}}^x)(t-t_0)} \tag{4.10}$$

This implies that (4.7) holds if f_W is defined as in Proposition 25. □

In the following Proposition 26, we study the evolution of the estimation error given by nonlinear observer (4.2) with $K_{i,g}$ determined following (4.3) in one sampling time (i.e., Δ). Proposition 26 provides an upper bound on the estimation error given by nonlinear observer (4.2) taking into account model uncertainties and communication delays. In order to proceed with the presentation, we define $\gamma_i(\tau)$ for $i \in \mathbb{I}$ as follows:

$$\gamma_i(\tau) = L_{F_i}^{y_i} L_{h_i}^{x_i} M_i \tau^2/2 + L_{F_i}^{y_i} \theta_{v_i} \tau + L_{f_i}^{w_i} \theta_{w_i} \tau + \sum_{l \in \mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \tau^2/2 + \sum_{l \in \mathbb{I}_i} M_{K_{i,l}} \theta_{v_l} \tau \tag{4.11}$$

where $L_{F_i}^{y_i}$ is the Lipschitz constant of F_i defined in (2.5) with respect to its second argument, $L_{h_i}^{x_i}$ is the Lipschitz constant of h_i with respect to its argument, M_i is a constant such that $|\dot{x}_i| \leq M_i$ for all $x_i \in \mathbb{X}_i$ with $i \in \mathbb{I}$, $L_{f_i}^{w_i}$ is the Lipschitz constant of f_i of its second argument, $L_{\tilde{f}_i}^{x_l}$ is the Lipschitz constant of $\tilde{f}_i$ with respect to x_l (noting that x_l is one part of X_i for $l \in \mathbb{I}_i$), $M_{K_{i,l}}$ is a constant such that $|K_{i,l}| \leq M_{K_{i,l}}$ for all $x_i \in \mathbb{X}_i$. Note that in the definition of γ_i, it is assumed that $|K_{i,l}|$ ($l \in \mathbb{I}_i$) is upper bounded which will be formally assumed in Proposition 26.

Proposition 26 *Consider the nonlinear observer of Eq. (4.2) for subsystem i, $i \in \mathbb{I}$, during the time interval $t \in [t_k, t_{k+1}]$ with initial condition $z_{n,i}(t_k) = \hat{x}_i(t_k)$ and output measurement $y_i(t_k)$. If $K_{i,l}$ are determined as in (4.3) and is bounded such that $|K_{i,l}| \leq M_{K_{i,l}}$ for all $x_i \in \mathbb{X}_i$, $l \in \mathbb{I}_i$, and $\frac{\partial \tilde{f}_i}{\partial x_l}$, $l \in \mathbb{I}_i$, are bounded such that $|\frac{\partial \tilde{f}_i}{\partial x_l}| \leq M_{\tilde{f}_i}^l$ for all $x_l \in \mathbb{X}_l$, then the deviation of the observer state $z_{n,i}$ in one sampling time Δ (i.e., at t_{k+1}) from the actual subsystem state x_i is bounded for all $x_i \in \mathbb{X}_i$, $i \in \mathbb{I}$, as follows:*

$$|e_{z,i}(t_{k+1})| \leq \beta_i(|e_{z,i}(t_k)|, \Delta) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta |e_{z,l}(t_k)|^2 + \Delta \alpha_i((D-1)\Delta, |e_z(t_{k-D+1})|) \tag{4.12}$$

where $e_{z,i} = z_{n,i} - x_i$, $L_{i,l} = M_{K_{i,l}} H_l^{h_l} + H_i^{\tilde{f}_i}$ and $\alpha_i(\tau, s) = \sum_{l \in \mathbb{I}_i, l \neq i} M_{\tilde{f}_i}^l f_W(\tau, s) + H_i^{\tilde{f}_i} f_W(\tau, s)^2$ with $H_l^{h_l}$, $H_i^{\tilde{f}_i}$ being positive constants that associated with the Taylor expansions of h_l, $\tilde{f}_i$, respectively.

Proof: We focus on the time interval $t \in [t_k, t_{k+1}]$ and define $e_{z,i} = z_{n,i} - x_i$ where $z_{n,i}$ denotes the trajectory of observer (4.2) and x_i is the actual state trajectory of subsystem i. Based on the subsystem model (2.1) and the expression of observer (4.2),the time derivative of $e_{z,i}$ is evaluated as follows:

$$\begin{aligned} \dot{e}_{z,i}(t) &= \dot{z}_{n,i}(t) - \dot{x}_i(t) \\ &= F_i(z_{n,i}(t), y_i(t_k)) - f_i(x_i(t), w_i(t)) + \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) \\ &\quad + \sum_{g \in \mathbb{Z}_i(t_{k+1})} K_{i,g}(\hat{x}_g(t_k))(y_g(t_k) - h_g(\hat{x}_g(t_k))). \end{aligned} \tag{4.13}$$

Note that the state trajectory of observer (4.2), $z_{n,i}$ for $t \in [t_k, t_{k+1}]$ is evaluated at time t_{k+1}. Therefore, $\mathbb{Z}_i(t_{k+1})$ is known in the above evaluation. From the Lipschitz properties of F_i, f_i and h_i, the fact that $y_i(t_k) = h_i(x_i(t_k)) + v_i(t_k)$, and $|v_i(t_k)| \leq \theta_{v_i}$, and $|w_i(t)| \leq \theta_{w_i}$, the following inequality can be obtained from (4.13):

$$\begin{aligned} |\dot{e}_{z,i}(t)| &\leq |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L_{F_i}^{y_i} L_{h_i}^{x_i} |x_i(t) - x_i(t_k)| + L_{F_i}^{y_i}\theta_{v_i} + L_{f_i}^{w_i}\theta_{w_i} \\ &\quad + \left| \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t)) + \sum_{g \in \mathbb{Z}_i(t_{k+1})} K_{i,g}(\hat{x}_g(t_k))(y_g(t_k) - h_g(\hat{x}_g(t_k))) \right| \end{aligned} \tag{4.14}$$

where $L_{F_i}^{y_i}$, $L_{h_i}^{x_i}$ and $L_{f_i}^{w_i}$ are the Lipschitz constants associated with F_i, h_i and f_i, respectively.

In order to find an upper bound on the right-hand-side of (4.14), we first expand the interaction term at t_k, $\tilde{f}_i(X_i(t_k))$, around its estimate $\hat{X}_i(t_k)$ and also expand $h_g(x_i(t_k))$ around its estimate $\hat{x}_i(t_k)$. Recalling that $\hat{X}_i(t_k)$ is composed of the estimate of the state of subsystem g for $g \in \mathbb{Z}_i(t_{k+1})$ and the predictions of the state of subsystem l with $l \in \mathbb{I}_i \backslash \mathbb{Z}_i(t_{k+1})$, the following expansion can be obtained using Taylor series expansion:

$$\begin{aligned} \tilde{f}_i(X_i(t_k)) &= \tilde{f}_i(\hat{X}_i(t_k)) + \sum_{l \in \mathbb{I}_i \backslash \mathbb{Z}_i(t_{k+1})} \frac{\partial \tilde{f}_i}{\partial x_l}(x_l^{p,i}(t_k))(x_l(t_k) - x_l^{p,i}(t_k)) \\ &\quad + \sum_{g \in \mathbb{Z}_i(t_{k+1})} \frac{\partial \tilde{f}_i}{\partial x_g}(\hat{x}_g(t_k))(x_g(t_k) - \hat{x}_g(t_k)) + H.O.T_i^{\tilde{f}_i} \\ h_g(x_g(t_k)) &= h_g(\hat{x}_g(t_k)) + \frac{\partial h_g}{\partial x_g}(\hat{x}_g(t_k))(x_g(t_k) - \hat{x}_g(t_k)) + H.O.T_g^{h_g} \end{aligned} \tag{4.15}$$

where $H.O.T_i^{\tilde{f}_i}$ and $H.O.T_g^{h_g}$ are high order terms associated with the expansions of $\tilde{f}_i$ and h_g. These high order terms satisfy the following constraints:

$$H.O.T_i^{\tilde{f}_i} \leq H_i^{\tilde{f}_i} |X_i(t_k) - \hat{X}_i(t_k)|^2, \; H.O.T_g^{h_g} \leq H_g^{h_g} |x_g(t_k) - \hat{x}_g(t_k)|^2 \tag{4.16}$$

for all $x_i \in \mathbb{X}_i$, $i = \mathbb{I}_i$ with $H_i^{\tilde{f}_i}$ and $H_g^{h_g}$ being positive constants. Let us define $A_i = \tilde{f}_i(\hat{X}_i(t_k)) - \tilde{f}_i(X_i(t_k)) + \sum_{g \in \mathbb{Z}_i(t_{k+1})} K_{i,g}(\hat{x}_g(t_k))(h_g(x_g(t_k)) - h_g(\hat{x}_g(t_k)))$. Applying the expansions in (4.15), A_i can be written in the following form:

$$\begin{aligned} A_i \;=\; & \sum_{g\in\mathbb{Z}_i(t_{k+1})}\left(-\frac{\partial\tilde{f}_i}{\partial x_g}(\hat{x}_g(t_k))(x_g(t_k)-\hat{x}_g(t_k))+K_{i,g}(\hat{x}_g(t_k))\frac{\partial h_g}{\partial x_g}(\hat{x}_g(t_k))(x_g(t_k)-\hat{x}_g(t_k))\right) \\ & -\sum_{l\in\mathbb{I}_i\backslash\mathbb{Z}_i(t_{k+1})}\frac{\partial\tilde{f}_i}{\partial x_l}(x_l^{p,i}(t_k))(x_l(t_k)-x_l^{p,i}(t_k))+\sum_{g\in\mathbb{Z}_i(t_{k+1})}K_{i,g}(\hat{x}_g(t_k))H.O.T_g^{h_g}-H.O.T_i^{\tilde{f}_i} \end{aligned} \tag{4.17}$$

If $K_{i,g}$ for $g\in\mathbb{Z}_i(t_{k+1})$ is determined following (4.3), the two terms in the first summation of the right-hand-side of (4.17) cancel with each other. Applying (4.16) to the last two terms in (4.17), it can be obtained that:

$$\begin{aligned} |A_i| \;\leq\; & \left|\sum_{l\in\mathbb{I}_i\backslash\mathbb{Z}_i(t_{k+1})}\frac{\partial\tilde{f}_i}{\partial x_l}(x_l^{p,i}(t_k))(x_l(t_k)-x_l^{p,i}(t_k))\right| \\ & +\sum_{g\in\mathbb{Z}_i(t_{k+1})}|K_{i,g}(\hat{x}_g(t_k))|\,H_g^{h_g}|x_g(t_k)-\hat{x}_g(t_k)|^2+H_i^{\tilde{f}_i}|X_i(t_k)-\hat{X}_i(t_k)|^2 \end{aligned} \tag{4.18}$$

Considering that $|x_l(t_k)-x_l^{p,i}(t_k)|\leq|x(t_k)-x^{p,i}(t_k)|^\dagger$, $l\in\mathbb{I}_i\backslash\mathbb{Z}_i(t_{k+1})$, the following inequality can be obtained by applying Proposition 25:

$$|x_l(t_k)-x_l^{p,i}(t_k)|\leq f_W\left(t_k-t_q,|x(t_q)-\hat{x}(t_q)|\right) \tag{4.19}$$

where t_q is the latest time instant that all of the other subsystem state estimates have been received by subsystem i so that $\hat{x}(t_q)$ is available. The worst case is that $t_q=t_{k-D+1}$ since the maximum possible delay is $D\Delta$. This implies that:

$$|x_l(t_k)-x_l^{p,i}(t_k)|\leq f_W\left((D-1)\Delta,|x(t_{k-D+1})-\hat{x}(t_{k-D+1})|\right). \tag{4.20}$$

If $\left|\frac{\partial\tilde{f}_i}{\partial x_l}(x_l^{p,i}(t_k))\right|$ is bounded such that $\left|\frac{\partial\tilde{f}_i}{\partial x_l}(x_l^{p,i}(t_k))\right|\leq M_{\tilde{f}_i}^l$ with $M_{\tilde{f}_i}^l$ a positive constant for all $x_l\in\mathbb{X}_l$, $l\in\mathbb{I}$, and if $|K_{i,g}(\hat{x}_g)|$ is bounded such that $|K_{i,g}(x_g)|\leq M_{K_{i,g}}$ with $M_{K_{i,g}}$ a positive constant for all $x_g\in\mathbb{X}_g$, $g\in\mathbb{I}$, (4.18) can be written as follows:

$$\begin{aligned} |A_i| \;\leq\; & \sum_{l\in\mathbb{I}_i\backslash\mathbb{Z}_i(t_{k+1})}M_{\tilde{f}_i}^l|x_l(t_k)-x_l^{p,i}(t_k)| \\ & +\sum_{g\in\mathbb{Z}_i(t_{k+1})}M_{K_{i,g}}H_g^{h_g}|x_g(t_k)-\hat{x}_g(t_k)|^2+H_i^{\tilde{f}_i}|X_i(t_k)-\hat{X}_i(t_k)|^2 \end{aligned} \tag{4.21}$$

Recalling that X_i is composed of two types of subsystem state elements: subsystem states predicted by the open-loop predictor and estimated subsystem states received from other

†Note that we use the difference between the entire system state and its prediction as an upper bound on the different between a subsystem state and its prediction. This approach is conservative but simplifies the proof.

subsystems without delay. It can be written that $|X_i(t_k) - \hat{X}_i(t_k)|^2 \leq \sum_{g\in\mathbb{Z}_i(t_{k+1})} |x_g(t_k) - \hat{x}_g(t_k)|^2 + |x(t_k) - x^{p,i}(t_k)|^2$. Therefore, (4.21) becomes:

$$\begin{aligned} |A_i| \;\leq\; & \sum_{l\in\mathbb{I}_i\setminus\mathbb{Z}_i(t_{k+1})} M^l_{\tilde{f}_i}|x_l(t_k) - x^{p,i}_l(t_k)| \\ & + \sum_{g\in\mathbb{Z}_i(t_{k+1})} (M_{K_{i,g}} H^{h_g}_g + H^{\tilde{f}_i}_i)|x_g(t_k) - \hat{x}_g(t_k)|^2 + H^{\tilde{f}_i}_i|x(t_k) - x^{p,i}(t_k)|^2 \end{aligned} \tag{4.22}$$

From (4.14) and the definition of A_i and $|K_{i,g}(\hat{x}_g(t_k))v_g(t_k)| \leq M_{K_{i,g}}\theta_{v_i}$, it can be obtained that:

$$\begin{aligned} |\dot{e}_{z,i}(t)| \;\leq\; & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L^{y_i}_{F_i}L^{x_i}_{h_i}|x_i(t) - x_i(t_k)| + L^{y_i}_{F_i}\theta_{v_i} + L^{w_i}_{f_i}\theta_{w_i} \\ & + |A_i| + |\tilde{f}_i(X_i(t)) - \tilde{f}_i(X_i(t_k))| + \sum_{g\in\mathbb{Z}_i(t_{k+1})} M_{K_{i,g}}\theta_{v_i} \end{aligned} \tag{4.23}$$

From (4.22) and (4.23) and the Lipschitz property of $\tilde{f}_i$ as well as the fact that X_i is composed of x_l for $l \in \mathbb{I}_i$, we have:

$$\begin{aligned} |\dot{e}_{z,i}(t)| \;\leq\; & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L^{y_i}_{F_i}L^{x_i}_{h_i}|x_i(t) - x_i(t_k)| + L^{y_i}_{F_i}\theta_{v_i} + L^{w_i}_{f_i}\theta_{w_i} \\ & + \sum_{l\in\mathbb{I}_i} L^{x_l}_{\tilde{f}_i}|x_l(t) - x_l(t_k)| + \sum_{g\in\mathbb{Z}_i(t_{k+1})} M_{K_{i,g}}\theta_{v_i} + \sum_{l\in\mathbb{I}_i\setminus\mathbb{Z}_i(t_{k+1})} M^l_{\tilde{f}_i}|x_l(t_k) - x^{p,i}_l(t_k)| \\ & + \sum_{g\in\mathbb{Z}_i(t_{k+1})} (M_{K_{i,g}} H^{h_g}_g + H^{\tilde{f}_i}_i)|x_g(t_k) - \hat{x}_g(t_k)|^2 + H^{\tilde{f}_i}_i|x(t_k) - x^{p,i}(t_k)|^2 \end{aligned} \tag{4.24}$$

Within one sampling time (e.g., $t \in [t_k, t_{k+1}]$), $|x_i(t) - x_i(t_k)| \leq M_i(t - t_k)$, $i \in \mathbb{I}$, where M_i is a positive constant that bounds the time derivative of x_i in $\mathbb{X}_i$ such that $|\dot{x}_i| \leq M_i$. With the initial condition $z_{n,i}(t_k) = \hat{x}_i(t_k)$ and the definition of $e_{z,i}$, it is known that $e_{z,i}(t_k) = x_i(t_k) - \hat{x}_i(t_k)$, $i \in \mathbb{I}$. Based on these results and (4.19) and Proposition 25, the following upper bound on $|\dot{e}_{z,i}(t)|$ can be written:

$$\begin{aligned} |\dot{e}_{z,i}(t)| \;\leq\; & |F_i(z_{n,i}(t), h_i(x_i(t))) - f_i(x_i(t), 0)| + L^{y_i}_{F_i}L^{x_i}_{h_i}M_i(t - t_k) + L^{y_i}_{F_i}\theta_{v_i} + L^{w_i}_{f_i}\theta_{w_i} \\ & + \sum_{l\in\mathbb{I}_i} L^{x_l}_{\tilde{f}_i}M_l(t - t_k) + \sum_{g\in\mathbb{Z}_i(t_{k+1})} M_{K_{i,g}}\theta_{v_i} + \sum_{g\in\mathbb{Z}_i(t_{k+1})} (M_{K_{i,g}} H^{h_g}_g + H^{\tilde{f}_i}_i)|e_{z,g}(t_k)|^2 \\ & + \sum_{l\in\mathbb{I}_i\setminus\mathbb{Z}_i(t_{k+1})} M^l_{\tilde{f}_i}f_W\left((D-1)\Delta, |e_z(t_{k-D+1})|\right) + H^{\tilde{f}_i}_i f_W\left((D-1)\Delta, |e_z(t_{k-D+1})|\right)^2 \end{aligned} \tag{4.25}$$

where $e_{z,i}(t_{k-D+1}) = x(t_{k-D+1}) - \hat{x}(t_{k-D+1})$. Integrating (4.25) from $t = t_k$ to $t = t_{k+1}$, the following inequality can be obtained:

$$
\begin{aligned}
|e_{z,i}(t_{k+1})| \quad \leq \quad & \beta_i(|e_{z,i}(t_k)|, \Delta) + L_{F_i}^{y_i} L_{h_i}^{x_i} M_i \Delta^2/2 + L_{F_i}^{y_i} \theta_{v_i} \Delta + L_{f_i}^{w_i} \theta_{w_i} \Delta + \sum_{l \in \mathbb{I}_i} L_{\tilde{f}_i}^{x_l} M_l \Delta^2/2 \\
& + \sum_{g \in \mathbb{Z}_i(t_{k+1})} M_{K_{i,g}} \theta_{v_i} \Delta + \sum_{g \in \mathbb{Z}_i(t_{k+1})} (M_{K_{i,g}} H_g^{h_g} + H_i^{\tilde{f}_i}) |e_{z,g}(t_k)|^2 \Delta \\
& + \sum_{l \in \mathbb{I}_i \setminus \mathbb{Z}_i(t_{k+1})} M_{\tilde{f}_i}^{l} f_W \left((D-1)\Delta, |e_z(t_{k-D+1})|\right) \Delta \\
& + H_i^{\tilde{f}_i} f_W \left((D-1)\Delta, |e_z(t_{k-D+1})|\right)^2 \Delta
\end{aligned}
\tag{4.26}
$$

Note that when integrating (4.25) from $t = t_k$ to $t = t_{k+1}$, the first term on the right-hand-side of (4.25) leads to the first term on the right-hand-side of (4.26) which is from the assumed property of observer (2.5); the second and fifth terms on the right-hand-side of (4.25) are linear in time which lead to second order in time in (4.26); and the other terms can be considered as constants when integrating with respect to time.

In (4.26), the set $\mathbb{Z}_i(t_{k+1})$ is a time-varying set. The two extreme cases are $\mathbb{Z}_i(t_{k+1}) = \{i\}$ which corresponds to the case that no other interacting subsystems' information is received without delay and $\mathbb{Z}_i(t_{k+1}) = \mathbb{I}_i$ which corresponds to the case that all interacting subsystems' information is received without delay. If γ_i, $L_{i,l}$, and α_i ($i \in \mathbb{I}$, $l \in \mathbb{I}_i$) are defined as in (4.11) and Proposition 26, (4.26) can be written in the form of (4.12). This proves Proposition 26. □

In Proposition 26, the estimation error of a subsystem state between the nonlinear observer (4.2a) $z_{n,i}$ and the actual system state x_i is shown to be bounded and the upper bound is associated with the Lipschitz properties of the system, the accuracy of the initial estimate $|e_{z,i}(t_k)|$, sampling interval Δ, magnitudes of noise and process disturbances, subsystem interactions and maximum delay D. Theorem 27 below provides sufficient conditions for the convergence and ultimate boundedness of the estimation error of the proposed DMHE with communication delays.

Theorem 27 *Consider system (2.4) with the outputs of its subsystems y_i sampled at time instants $\{t_{k\geq 0}\}$. If the proposed DMHE implemented following Algorithm 23 with subsystem MHE designed as in (4.4) based on deterministic nonlinear observers satisfying (2.6) and the assumptions in Proposition 26 are satisfied, and if there exist concave functions $g_i(\cdot)$, $i \in \mathbb{I}$, such that:*

$$g_i(|e_i|) \geq \beta_i(|e_i|, \Delta) \tag{4.27}$$

for all $|e_i| \leq d_i$ and if there exist constants $d_{s,i}$, d_i such that $0 \leq d_{s,i} \leq d_i$, $D > 0$ and *positive constants $a_i \geq 1$, $b_i > 0$, and $\eta_i > 0$, such that:*

$$d_{s,i} - a_i \left(g_i(d_{s,i}) + \gamma_i(\Delta) + \sum_{l \in \mathbb{I}_i} L_{i,l} \Delta d_l^2 + \Delta \alpha_i((D-1)\Delta, d) \right) - b_i \theta_{v_i} \geq \eta_i \tag{4.28}$$

where $d = [d_1, \ldots, d_m]^T$ for all $i \in \mathbb{I}$, and if κ_i for all $i \in \mathbb{I}$, are picked as follows:

$$0 \leq \kappa_i \leq \min\{(a_i - 1)/L_{h_i}^{x_i}, b_i\}, \tag{4.29}$$

then the estimation error $|e_i| = |\hat{x}_i - x_i|$ is a decreasing sequence if $d_{s,i} \leq |e_i(0)| \leq d_i$ for all $i \in \mathbb{I}$ and is ultimately bounded as follows:

$$\lim_{t \to \infty} \sup |e_i(t)| \leq d_{i,\min} \tag{4.30}$$

for $i \in \mathbb{I}$ with $d_{i,\min} = \max\{|e_i(t + \Delta)| : |e_i(t)| \leq d_{s,i}\}$ for all $e_i(0) \leq d_i$ and $x_i \in \mathbb{X}_i$. This also implies that the entire system state estimation error is ultimately bounded.

Proof: We prove that the evolution of the estimation error of each subsystem state $|e_i| = |\hat{x}_i - x_i|$, $i \in \mathbb{I}$, under the proposed DMHE with the local MHE of Eq. (4.4) implemented following Algorithm 1 is a decreasing sequence and is ultimately bounded in a small region around zero. The decrease and ultimate boundedness of subsystem estimation errors imply the decrease and ultimate boundedness of the entire system state estimation error. We first focus on MHE i, $i \in \mathbb{I}$, for the time interval from t_k to t_{k+1} and then extend it to the general case. From constraint (4.4e) for MHE i, it can be written that:

$$|\hat{x}_i(t_{k+1}) - z_{n,i}(t_{k+1})| \leq \kappa_i |y_i(t_{k+1}) - h_i(z_{n,i}(t_{k+1}))| \tag{4.31}$$

Note that $\hat{x}_i$ denotes the final optimal estimate obtained by MHE i. From the Lipschitz property of h_i, the fact that $y_i = h_i(x_i) + v_i$ and $|v_i| \leq \theta_{v_i}$, (4.31) becomes:

$$|\hat{x}_i(t_{k+1}) - z_{n,i}(t_{k+1})| \leq \kappa_i L_{h_i}^{x_i} |x_i(t_{k+1}) - z_{n,i}(t_{k+1})| + \kappa_i \theta_{v_i} \tag{4.32}$$

where $L_{h_i}^{x_i}$ is the Lipschitz constant of h_i with respect to x_i as defined before Proposition 26. Using the triangle inequality $|\hat{x}_i - x_i| \leq |\hat{x}_i - z_{n,i}| + |z_{n,i} - x_i|$, it is obtained from (4.32) that:

$$|e_i(t_{k+1})| \leq \left(1 + \kappa_i L_{h_i}^{x_i}\right) |e_{z,i}(t_{k+1})| + \kappa_i \theta_{v_i} \tag{4.33}$$

with $e_i(t_{k+1}) = x_i(t_{k+1}) - \hat{x}_i(t_{k+1})$ and $e_{z,i}(t_{k+1}) = x_i(t_{k+1}) - z_{n,i}(t_{k+1})$. From the design of observer (4.2), it can be seen that in order to calculate a reference state estimate for t_{k+1}, it is initialized with $z_{n,i}(t_k) = \hat{x}_i(t_k)$. This implies that $e_i(t_k) = e_{z,i}(t_k)$. Based on the upper bound on $|e_{z,i}(t_{k+1})|$ obtained in Proposition 26 and (4.33), the following inequality can be written:

$$\begin{aligned} |e_i(t_{k+1})| \quad \leq \quad & \left(1+\kappa_i L_{h_i}^{x_i}\right)(\beta_i(|e_i(t_k)|,\Delta)+\gamma_i(\Delta)+\sum_{l\in\mathbb{I}_i} L_{i,l}\Delta|e_l(t_k)|^2 \\ & +\Delta\alpha_i((D-1)\Delta,|e(t_{k-D+1})|))+\kappa_i\theta_{v_i} \end{aligned} \tag{4.34}$$

If condition (4.27) is satisfied, $\beta_i(|e_i(t_k)|,\Delta)$ in (4.34) can be replaced by $g_i(|e_i(t_k)|)$ and the following inequality can be obtained:

$$\begin{aligned} |e_i(t_{k+1})| \quad \leq \quad & \left(1+\kappa_i L_{h_i}^{x_i}\right)(g_i(|e_i(t_k)|)+\gamma_i(\Delta)+\sum_{l\in\mathbb{I}_i} L_{i,l}\Delta|e_l(t_k)|^2 \\ & +\Delta\alpha_i((D-1)\Delta,|e(t_{k-D+1})|))+\kappa_i\theta_{v_i} \end{aligned} \tag{4.35}$$

When examining condition (4.28) in Theorem 27, the left-hand-side can be considered as a function of $d_{s,i}$. The first term on the left-hand-side has a slope of 1 and the slope of g_i decreases with the increase of $d_{s,i}$ since g_i is a concave function. Given that $g_i(0) \geq 0$ (which is ensured by (4.27)), if there exists $d_{s,i}$ satisfying (4.28) which means the left-hand-side has a positive value, then (4.28) holds for all $d_{s,i} \leq |e_i| \leq d_i$. That is,

$$|e_i| - a_i\left(g_i(|e_i|)+\gamma_i(\Delta)+\sum_{l\in\mathbb{I}_i} L_{i,l}\Delta|e_l|^2+\Delta\alpha_i((D-1)\Delta,|e|)\right)-b_i\theta_{v_i} \geq \eta_i \tag{4.36}$$

for all $d_{s,i} \leq |e_i| \leq d_i$, $|e_l| \leq d_l$, $l \in \mathbb{I}_i$ and $|e| \leq d$. If κ_i is picked following (4.29), it is ensured that $1+\kappa_i L_{h_i}^{x_i} \leq a_i$ and $\kappa_i \leq b_i$. This further implies that:

$$|e_i| - \left(1+\kappa_i L_{h_i}^{x_i}\right)\left(g_i(|e_i|)+\gamma_i(\Delta)+\sum_{l\in\mathbb{I}_i} L_{i,l}\Delta|e_l|^2+\Delta\alpha_i((D-1)\Delta,|e|)\right)-\kappa_i\theta_{v_i} \geq \eta_i \tag{4.37}$$

for all $d_{s,i} \leq |e_i| \leq d_i$, $|e_l| \leq d_l$, $l \in \mathbb{I}_i$ and $|e| \leq d$. Rearrange (4.37) and put it back into (4.35), it can be obtained that:

$$|e_i(t_{k+1})| \leq |e_i(t_k)| - \eta_i \tag{4.38}$$

for all $d_{s,i} \leq |e_i(t_k)| \leq d_i$, $|e_l(t_k)| \leq d_l$, $l \in \mathbb{I}_i$ and $|e| \leq d$. If $|e_i| \geq d_{s,i}$ for all the time from 0 to t_k, using (4.38) recursively, it can be obtained that:

$$|e_i(t_k)| \leq |e_i(0)| - k\eta_i \tag{4.39}$$

for all $d_{s,i} \leq |e_i(t_k)| \leq d_i$. This implies that $|e_i|$ decreases every sampling time and will become smaller than $d_{s,i}$ in finite steps. Once $|e_i(t)| < d_{s,i}$, there is no longer a guarantee that the estimation error will decrease. However, it will remain to prove that $|e_i(t)| \leq d_{i,\min}$

given the definition of $d_{i,\min}$; that is, $\lim_{t\to\infty} \sup |e_i(t)| \leq d_{i,\min}$. Note that the above proof holds for all $i \in \mathbb{I}$.

The ultimate boundedness of subsystem state estimation errors implies the ultimate boundedness of the entire system state estimation error. This can be seen from the inequality $|e| \leq \sum_{i=1}^{m} |e_i|$ which implies that:

$$\lim_{t\to\infty} \sup |e| \leq \sum_{i=1}^{m} d_{i,\min}. \tag{4.40}$$

This proves Theorem 27. □

Remark 28 *Referring to the inequality (4.28) in Theorem 27, it characterizes the interplay between different parameters. The first term in the bracket $g_i(d_{s,i})$ is the upper bound of the estimation error for the nominal subsystem i without interactions after one sampling time if the initial error term is $|e_i(t_k)|$; the term $\gamma_i(\Delta)$ indicates the effect of the sample-and-hold implementation of the nonlinear observer of Eq. (4.2), process disturbances and measurement noise; the term $\sum_{l\in\mathbb{I}_i} L_{i,l}\Delta d_l^2$ bounds the model mismatches in the interaction model due to periodic measurements and information transmission; the term $\Delta\alpha_i((D-1)\Delta, d)$ characterizes the contribution of other subsystems in the interactions considering the maximum delay D; and the term $b_i\theta_{v_i}$ represents the effect of measurement noise. The inequality (4.28) essentially requires that the nonlinear observer (2.5) converges fast enough and the interactions between the subsystems are well compensated for such that the contribution of the nonlinear observer (2.5) to the decrease of the estimation error dominates the effects caused by other factors that contribute to the increase of the estimation error.*

Remark 29 *Note that the stability of the proposed DMHE with communication delays is essentially from the deterministic nonlinear observers. The subsystem MHE design (4.4) integrates the deterministic nonlinear observer into the framework of MHE via constraint (4.4e). This integration brings some very interesting features into the design [50]: 1) characterizable boundedness of the estimation error for bounded uncertainties which is a difficult task for the traditional MHE; and 2) a potentially tunable convergence rate of the estimate to the actual system state (via the tuning of the nonlinear observer) which in general is not straightforward (or not possible) in the traditional MHE for nonlinear systems. These features are very important from an output feedback control view point. Besides these features, the MHE design adopted in this work essentially maintains the properties of the traditional MHE such as optimality of the estimate since the analysis used to establish the optimality* *of the traditional MHE based on the cost function and the approximation of the arrival cost may be applied to the MHE design in this work as well. The focus of our work is more on the stability of the distributed MHE scheme subject to communication delays which relies on the deterministic observers.*

4.4 Application to the reactor-separator process

4.4.1 Simulation settings

In this section, the proposed DMHE design is demonstrated by the reactor-separator process used in Chapter 2 and 3. A detailed description of the modeling of the process can be found in Section 2.6. The temperature measurements (i.e., T_i, $i = 1, 2, 3$) are subject to bounded noise. The measurement noise is generated as normal distributed values with zero mean and standard deviation 1 but the values are bounded in $[-1, 1]$. Besides measurement noise, random disturbances are also introduced to the right-hand-side of dynamic equations of the process, which are generated as normal distributed values with zero mean and standard deviation 100 in the range $[-10, 10]$ for temperatures and normal distributed values with zero mean and standard deviation 1 in the range $[-1, 1]$ for species fractions.

The deterministic nonlinear observer (2.5) without considering interactions between subsystems are as in Section 2.6.2:

$$\dot{z}_i(t) = f_i(z_i(t), 0) + G_i(z_i(t))^{-1}K_{o,i}(y_i(t) - h_i(z_i(t))) \tag{4.41}$$

the gain matrices $K_{o,i}$ is determined such that the eigenvalues of the matrix $A_{o,i} - K_{o,i}C_{o,i}$ are placed at $-0.1 \pm i$ and -5 with $A_{oi} = [0\ 1\ 0;\ 0\ 0\ 1;\ 0\ 0\ 0]$ and $C_{o,i} = [1\ 0\ 0]$. These nonlinear observers are used in the design of observer (4.2) and in the design of the DMHE schemes. In all the local MHE designs, the estimation horizon is $N = 10$. All the other parameters can be referred in Section 2.6.

4.4.2 Simulation results

First, the proposed DMHE is compared with the DMHE scheme in Chapter 2 in which communication delays are not taken into account explicitly as well as the deterministic nonlinear observer (4.2) implemented following Algorithm 23. In this set of simulations, we consider an extreme scenario in which the communication delays between subsystems always equal to the maximum possible delay D with $D = 5$. In the proposed DMHE scheme as well as in the nonlinear observer (4.2), when there are communication delays, the predictors are used to generate predictions to minimize the effects of delays. While in the DMHE scheme

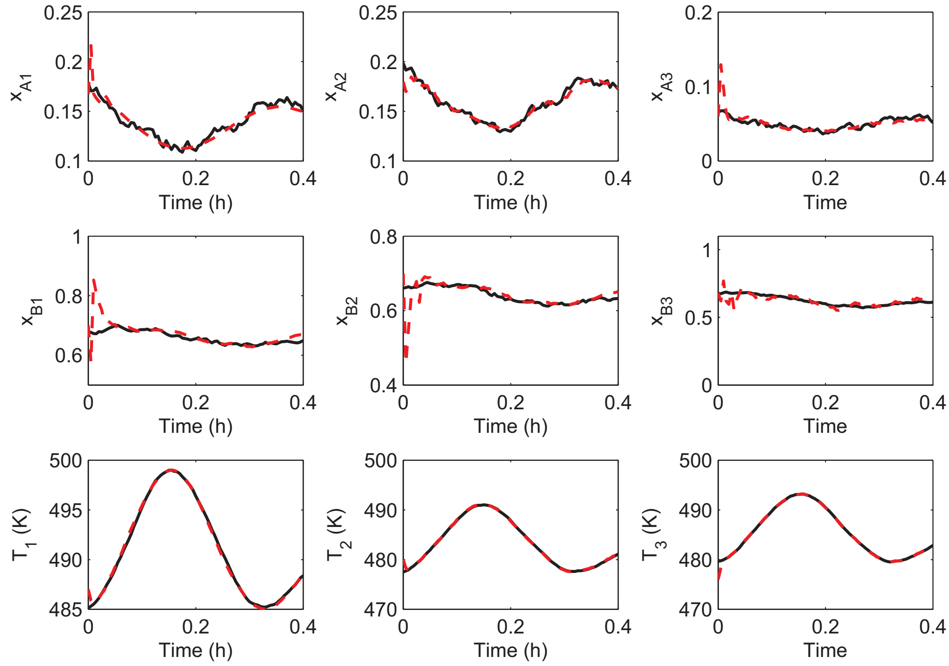


Figure 4.2: Trajectories of the actual process states (solid lines) and the estimates given by the proposed DMHE (dashed lines) when the communication delays between subsystems always equal to the maximum possible delay D with $D = 5$.

in Chapter 2, when there is communication delays, the latest received information is used to approximate the current information.

In this set of simulations, the same initial conditions, disturbances, noise sequences and heat inputs are used in the two DMHE schemes and in nonlinear observer (4.2). The initial condition of the process is:

$$x_0 = [0.178; 0.680; 485.120\ K; 0.199; 0.660; 477.514\ K; 0.066; 0.677; 479.637\ K]^T$$

and the initial guess in the two DMHE schemes and observer (4.2) is

$$\hat{x}_0 = [0.168; 0.700; 487.000\ K; 0.180; 0.700; 480.000\ K; 0.060; 0.680; 476.000\ K]^T$$

with input $Q_i = Q_{s,i} + 10^6 e^{-0.01t} sin(0.1t)\ KJ$. This type of inputs is used to excite different changing rates in the process state trajectories. Note that when process inputs are present, the inputs need to be taken into account in the design of the auxiliary nonlinear observer as well as the DMHE schemes and should be assumed to be known.

The simulation results are shown in Figs. 4.2, 4.3, 4.4 and 4.5. Figures 4.2, 4.3, and 4.4 show the trajectories of the estimates given by the proposed DMHE, the DMHE in Chapter

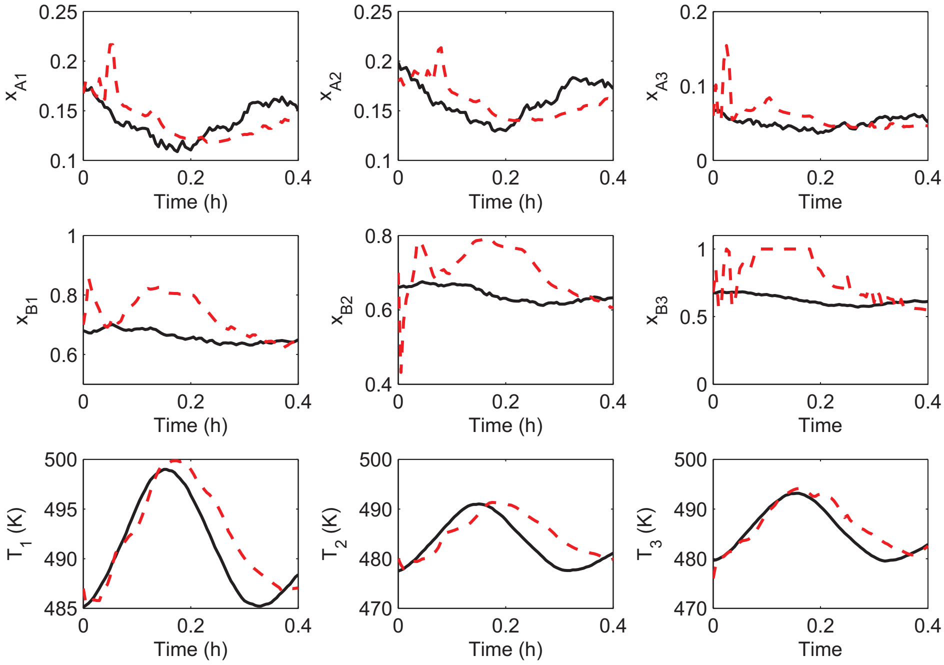


Figure 4.3: Trajectories of the actual process states (solid lines) and the estimates given by the DMHE (dashed lines) in Chapter 2 when the communication delays between subsystems always equal to the maximum possible delay D with $D = 5$.

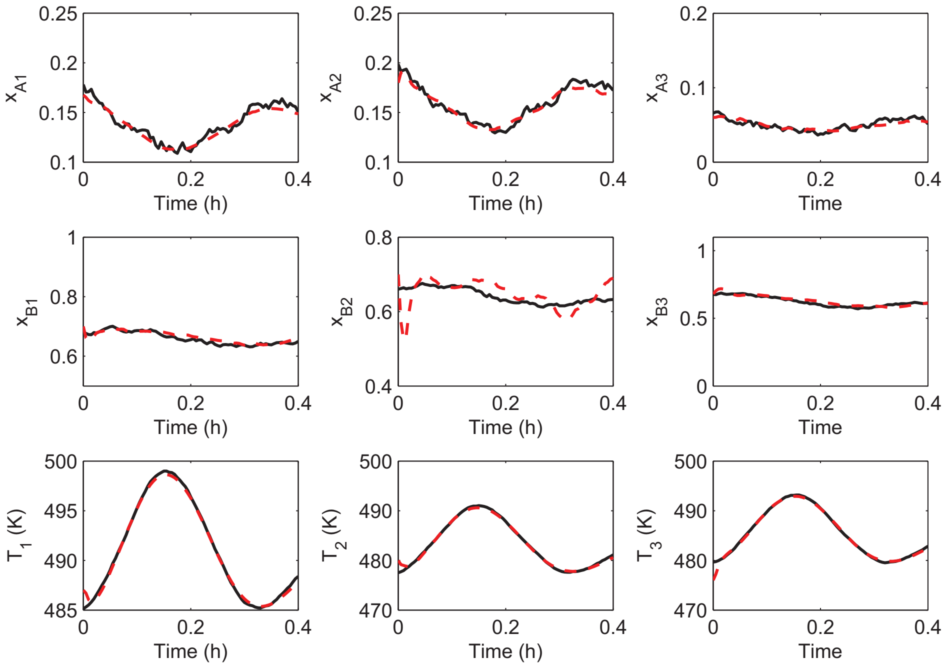


Figure 4.4: Trajectories of the actual process states (solid lines) and the estimates given by nonlinear observer (4.2) (dashed lines) when the communication delays between subsystems always equal to the maximum possible delay D with $D = 5$.

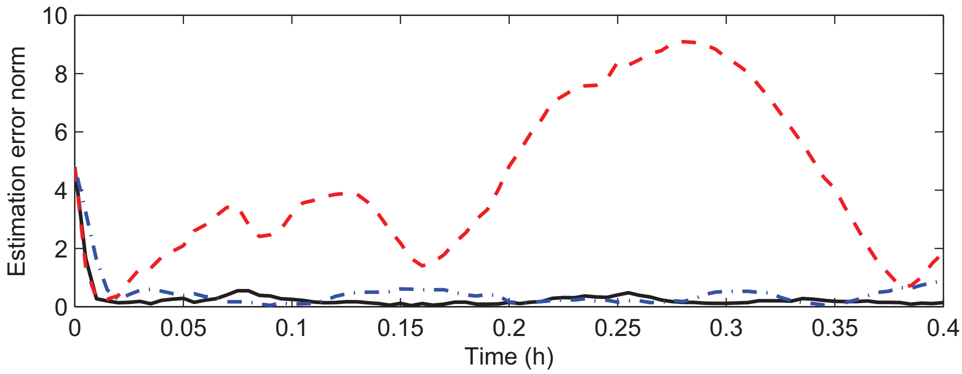


Figure 4.5: Trajectories of the norm of the estimation errors of the proposed DMHE (solid line) and the DMHE in Chapter 2 (dashed line) and nonlinear observer (4.2) (dash-dotted line) when the communication delays between subsystems always equal to the maximum possible delay D with $D = 5$.

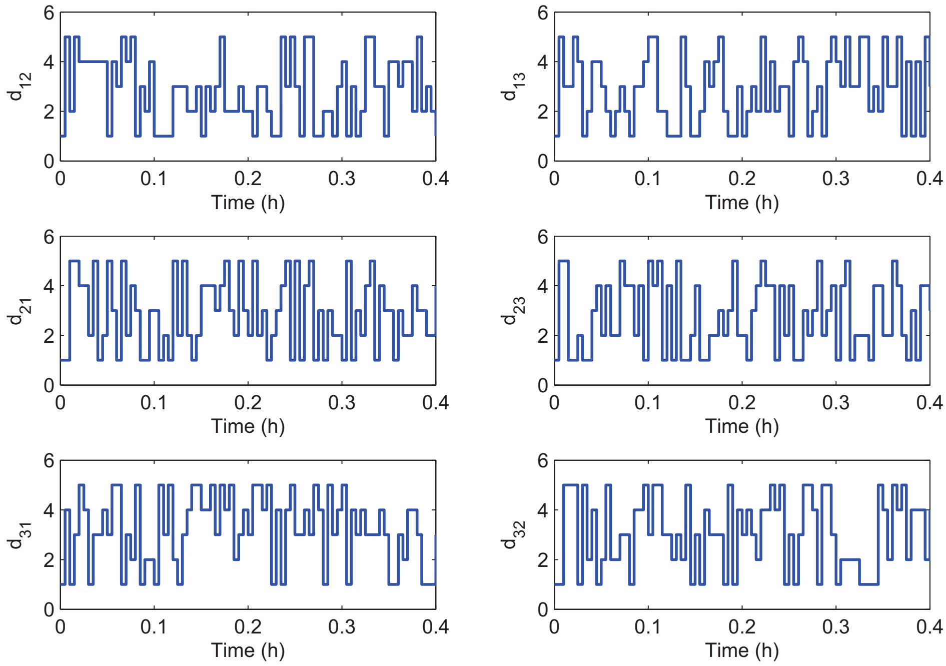


Figure 4.6: Communication delay sequences.

2 and nonlinear observer (4.2), respectively. From these figures, it can be seen that the both the proposed DMHE and nonlinear observer (4.2) are able to track the actual system states well while the DMHE in in Chapter 2 gives much poorer performance in tracking the actual system states. Figure 4.5 shows the trajectories of the norm of the estimation errors given by the three different schemes. From Fig. 4.5, it can be seen that the proposed DMHE and nonlinear observer (4.2) is able to drive the estimation error to a small value quickly and maintains the error in a small region close to zero while the estimation error of the DMHE in Chapter 2 varies significantly. Even though both the proposed DMHE and nonlinear observer (4.2) are able to tracking the actual states well, the proposed DMHE gives a much smaller average error norm (0.2778) compared with the one (0.4561) given by nonlinear observer (4.2). The average error norm reduction obtained by using the proposed DMHE is about 40% in this set of simulations.

Next, a set of simulations is carried out to consider a normal scenario in which the communication delays between subsystems are generated as random integers between 1 and D with $D = 5$. Figure 4.6 shows the communication delays among the subsystems in the simulations.

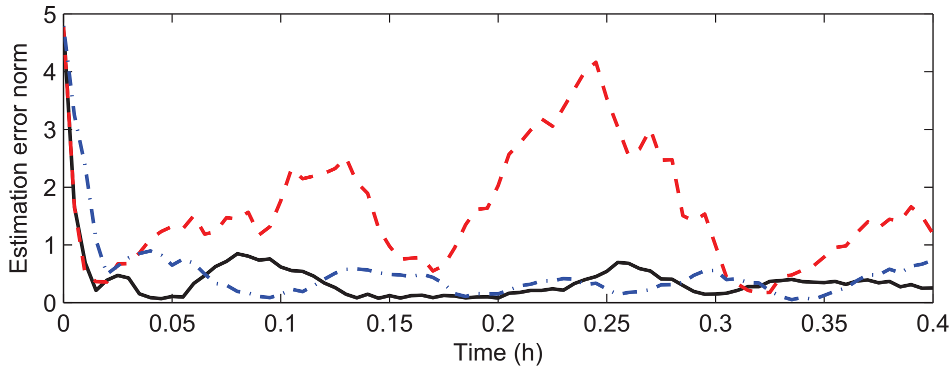
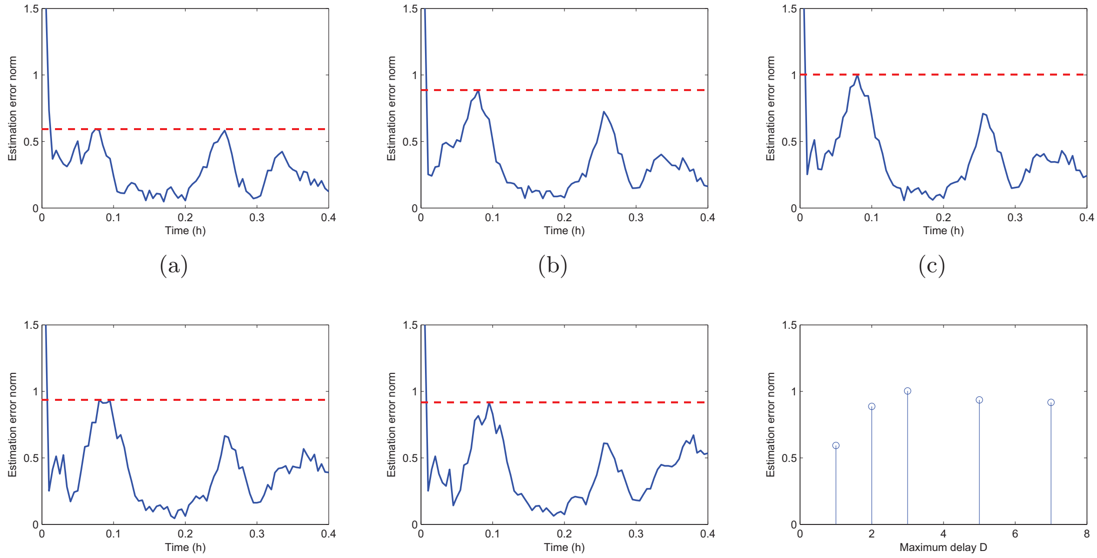


Figure 4.7: Trajectories of the norm of the estimation errors of the proposed DMHE (solid line) and the DMHE in Chapter 2 (dashed line) and nonlinear observer (4.2) (dash-dotted line).

Figure 4.7 shows the trajectories of the norm of the estimation errors given by the two DMHE schemes and nonlinear observer (4.2). Similar conclusions to the first set of simulations can be concluded. That is, the proposed DMHE as well as nonlinear observer (4.2) are able to drive the estimation error to a small region around zero quickly while the estimation error of the DMHE in Chapter 2 varies significantly. Note that in this set of simulations, the average error norm given by the proposed DMHE is 0.3972 while the one given by nonlinear observer (4.2) is 0.5018. About 20% average error norm reduction is achieved by using the proposed DMHE.

In another set of simulations, the effect of the maximum delay D on the size of the set that ultimately bounds the estimation error is investigated. In this set of simulations, the proposed DMHE scheme is simulated with different maximum communication delays. In particular, the maximum delay $D = 1, 2, 3, 5, 7$ are considered. Figures 4.8(a)-(e) show the corresponding trajectories of the norm of the estimation error for $D = 1, 2, 3, 5, 7$, respectively. In each of these figures, the flat line is the approximated bound that ultimately bounds the estimation error. Figure 4.8(e) shows these bounds in one figure for easy comparison. From these figures, it can be seen that (i) when $D = 1$ which corresponds to the case that there is no communication delay, the estimation error is maintained in a relatively smaller set compared with the other cases that there are communication delays; (ii) when there is communication delay (i.e., $D > 1$), the value of the communication delay has a relatively small effect on the size of the set that ultimately bounds the estimation error. The superior performance when $D = 1$ is primarily due to the use of the correction terms in observer (4.2a) for all the interacting subsystems. Note that when there is communication



(a) (b) (c) (d) (e) (f)

Figure 4.8: Trajectories of the norm of the estimation error of the proposed DMHE (solid lines) and the approximated bound that ultimately bounds the estimation error (dashed lines) when (a) $D = 1$, (b) $D = 2$, (c) $D = 3$, (d) $D = 5$, (e) $D = 7$; and (f) the bounds in one figure.

Table 4.1: Mean evaluation times of the predictor, observer (4.2) and local MHE for each subsystem.

Subsystem	Predictor (sec)	Observer (4.2) (sec)	Local MHE (sec)
# 1	6.23×10^{-4}	1.79×10^{-3}	1.26
# 2	5.31×10^{-4}	1.42×10^{-3}	1.54
# 3	6.55×10^{-4}	1.42×10^{-3}	1.23

delay, some of the correction terms will become inactive.

Finally, the mean evaluation times of the predictor, observer (4.2) and the local MHE for each subsystem are evaluated. The simulations are carried out in MATLAB using an Intel Core i7 Computer at 3.4GHz with 8Gb RAM. The mean values shown in Table 4.1 are obtained from 100 simulations with estimation window horizon $N = 10$. From these results, it can be seen that the evaluation times of the predictor for state prediction and nonlinear observer (4.2) for reference state estimate calculation are negligible compared with the evaluation times of local MHE schemes.

4.5 Conclusions

In this chapter, we developed a DMHE scheme for a class of nonlinear systems in the presence of delays in the communication network. In the proposed design, an all-to-all communication is required and at each sampling time for each subsystem, an open-loop predictor is used to generate predictions of delayed subsystem states which are used to calculate a reference state estimate as well as a confidence region for the local MHE. Under the assumption that there is an upper bound on the communication delay, the proposed approach ensures the convergence of the estimate to the actual system state and the ultimate boundedness of the estimation. An chemical process example was used to illustrate the performance of the proposed DMHE scheme compared with an existing DMHE without considering communication delays explicitly as well as a deterministic nonlinear observer. From the simulations, it was demonstrated that the proposed DMHE had a superior performance than the DMHE without considering the delays explicitly and had improved performance compared with the deterministic observer.

Chapter 5

Conclusions

In this thesis, distributed moving horizon estimation taking into account uncertainty, communication delays and triggered implementation are presented for nonlinear systems. Moreover, its effectiveness and applicability are illustrated by a chemical process example.

Specifically, Chapter 2 proposed an observer-enhanced DMHE design for a class of nonlinear systems with bounded process uncertainties. In this DMHE, each subsystem MHE communicates with subsystems that it interacts with every sampling time. In the design of each subsystem MHE, an auxiliary deterministic nonlinear observer is taken advantage of to calculate a confidence region that contains the actual system state every sampling time. The subsystem MHE is only allowed to optimize its state estimate within the confidence region. This strategy was demonstrated to guarantee the convergence and ultimate boundedness properties of the estimation error.

However, this approach requires the subsystems to communicate and exchange information every sampling time. The frequent communication requirement may impede the application of the DMHE to processes that have a shared communication network with limited capacity. Moreover, extensive information exchanging may reduce the robustness of the system due to data dropouts in the communication network. Currently, there is no systematic approaches available to reduce the communication burden of DMHE schemes. To address the above issues, in Chapter 3 event-triggered approach is adopted to reduce the number of communication between subsystems. Specifically, in the first proposed algorithm, a subsystem sends out its current information when a triggering condition based on the difference between the current state estimate and a previously transmitted state estimate is satisfied; in the second proposed algorithm, the transmission of information from a subsystem to other subsystems is triggered by the difference between the current measurement of the output and its derivatives and a previously transmitted measurement of the output and its derivatives. Because of the triggered communication, a subsystem may not have the latest information of the other subsystems. In order to ensure the convergence and ultimately boundedness of the estimation error, the local MHE of a subsystem also needs to be redesigned to account for the possible lack of state updates from other subsystems. Sufficient conditions for the proposed DMHE implemented following the two algorithms to ensure the convergence and ultimately boundedness of the estimation error are derived.

The results in Chapter 2 were derived under the assumption that the communication between subsystems is flawless and there is no delay in the information transmission. In practice, this assumption may not hold especially when shared wireless communication network is used. Motivated by the above considerations, in Chapter 4 we proposed a DMHE scheme that is able to handle time-varying communication delays. In the proposed design, a nonlinear observer-enhanced MHE is designed for each subsystem and the distributed MHEs are assumed to be able to communicate and exchange information with each other via a shared communication network which may introduce communication delays. To handle time-varying delays in the communication, an open-loop state predictor is designed for each subsystem to provide predictions of unavailable subsystem states. In the design of each predictor, the centralized system model is used. Based on the state predictions, an auxiliary nonlinear observer is used to generate a reference subsystem state estimate for each subsystem every sampling time. Based on the reference subsystem state estimate as well as the local output measurement, a confidence region is constructed for the actual state of a subsystem. A subsystem MHE is only allowed to optimize its state estimate within the corresponding confidence region at a sampling time. The proposed DMHE is proved to give decreasing and ultimately bounded estimation errors under the assumption that there is an upper bound on the time-varying delay.

Bibliography

[1] D. Viegas, P. Batista, P. Oliveira, and C. Silvestre. Decentralized H_2 observers for position and velocity estimation in vehicle formations with fixed topologies. *Systems & Control Letters*, 61:443–453, 2012.

[2] M. V. Subbotin and R. S. Smith. Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies. *Automatica*, 45:2491–2501, 2009.

[3] B. Chen and W. Wang. Robust stabilization of nonlinearly perturbed large-scale systems by decentralized observer-controller compensators. *Automatica*, 26:1305–1041, 1990.

[4] G. Antonelli, F. Arrichiello, F. Caccavale, and A. Marino. A decentralized controller-observer scheme for multi-agent weighted centroid tracking. *IEEE Transactions on Automatic Control*, 58:1310–1316, 2013.

[5] 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.

[6] 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.

[7] 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.

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

[9] 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.

[10] 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.

[11] N. R. Sandell, P. Varajya, M. Athans, and M. C. Safonov. Survey of decentralized control methods for large scale systems. *IEEE Transactions on Automatic Control*, 23:108–128, 1978.

[12] L. Bakule. Decentralized control: An overview. *Annual Reviews in Control*, 32:87–98, 2008.

[13] J. B. Rawlings and B. T. Stewart. Coordinating multiple optimization-based controllers: New opportunities and challenges. *Journal of Process Control*, 18:839–845, 2008.

[14] R. Scattolini. Architectures for distributed and hierarchical model predictive control - A review. *Journal of Process Control*, 19:723–731, 2009.

[15] P. D. Christofides, J. Liu, and D. Muñoz de la Peña. *Networked and Distributed Predictive Control: Methods and Nonlinear Process Network Applications*. Advances in Industrial Control Series. Springer-Verlag, London, England, 2011. (230 pages).

[16] A. Richards and J. P. How. Robust distributed model predictive control. *International Journal of Control*, 80:1517–1531, 2007.

[17] E. Camponogara, D. Jia, B. H. Krogh, and S. Talukdar. Distributed model predictive control. *IEEE Control Systems Magazine*, 22:44–52, 2002.

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

[19] W. B. Dunbar and R. M. Murray. Distributed receding horizon control for multi-vehicle formation stabilization. *Automatica*, 42:549–558, 2006.

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

[21] 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.

[22] A. N. Venkat, J. B. Rawlings, and S. J. Wright. Stability and optimality of distributed model predictive control. In *Proceedings of the 44th IEEE Conference on Decision and Control and the European Control Conference ECC 2005*, pages 6680–6685, Seville, Spain, 2005.

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

[24] B. T. Stewart, S. J. Wright, and J. B. Rawlings. Cooperative distributed model predictive control for nonlinear systems. *Journal of Process Control*, 21:698–704, 2011.

[25] 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.

[26] J. Liu, X. Chen, D. Muñoz de la Peña, and P. D. Christofides. Sequential and iterative architectures for distributed model predictive control of nonlinear process systems. *AIChE Journal*, 56:2137–2149, 2010.

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

[28] E. A. Misawa and J. K. Hedrick. Nonlinear observers - A state-of-the-art survey. *Transactions of the ASME*, 111:344–352, 1989.

[29] J. Tsinias. Observer design for nonlinear control systems. *Systems & Control Letters*, 13:135–142, 1989.

[30] J. P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems: applications to bioreactors. *IEEE Transactions on Automatic Control*, 37:875–880, 1992.

[31] G. Ciccarella, M. Dalla Mora, and A. Germani. A Luenberger-like observer for nonlilear systems. *International Journal of Control*, 57:537–556, 1993.

[32] N. Kazantzis and C. Kravaris. A nonlinear Luenberger-type observer with application to catalytic activity estimation. In *Proceedings of American Control Conference*, pages 1756–1760, Seattle, WA, USA, 1995.

[33] M. Soroush. Nonlinear state observer design with application to reactors. *Chemical Engineering Science*, 52:387–404, 1997.

[34] M. Soroush. State and parameter estimations and their applications in process control. *Computers & Chemical Engineering*, 23:229–245, 1998.

[35] N. Kazantzis and C. Kravaris. Nonlinear observer design using Lyapunov's auxiliary theorem. *Systems & Control Letters*, 34:241–247, 1998.

[36] D. Dochain. State and parameter estimation in chemical and biochemical processes: A tutorial. *Journal of Process Control*, 13:801–818, 2003.

[37] I. Karafyllis and C. Kravaris. Global exponential observers for two classes of nonlinear systems. *Systems & Control Letters*, 61:797–806, 2012.

[38] T. Ahmed-Ali, I. Karafyllis, and F. Lamnabhi-Lagarrigue. Global exponential sampled-data observers for nonlinear systems with delayed measurements. *Systems & Control Letters*, 62:539–549, 2013.

[39] H. K. Khalil and F. Esfandiari. Semiglobal stabilization of a class of nonlinear systems using output feedback. *IEEE Transactions on Automatic Control*, 38:1412–1415, 1993.

[40] N. A. Mahmoud and H. K. Khalil. Asymptotic regulation of minimum phase nonlinear systems using output feedback. *IEEE Transactions on Automatic Control*, 41:1402–1412, 1996.

[41] A. Teel and L. Praly. Tools for semi-global stabilization by partial state and output feedback. *SIAM Journal on Control and Optimization*, 33:1443–1488, 1995.

[42] D. Muñoz de la Peña and P. D. Christofides. Stability of nonlinear asynchronous systems. *Systems & Control Letters*, 57:465–473, 2008.

[43] M. Heidarinejad, J. Liu, and P. D. Christofides. State estimation-based economic model predictive control of nonlinear systems. *Systems & Control Letters*, 61:926–935, 2012.

[44] M. Du and P. Mhaskar. Isolation and handling of sensor faults in nonlinear system. In *Proceedings of the American Control Conference*, pages 6667–6672, Montreal, QC, Canada, 2012.

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

[46] C. V. Rao, J. B. Rawlings, and D. Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48:246–258, 2003.

[47] J. B. Rawlings and L. Ji. Optimization-based state estimation: Current status and some new results. *Journal of Process Control*, 22:1439–1444, 2012.

[48] R. Huang, L. T. Biegler, and S. C. Patwardhan. Fast offset-free nonlinear model predictive control based on moving horizon estimation. *Industrial & Engineering Chemistry Research*, 49:7882–7890, 2010.

[49] K. R. Muske, J. B. Rawlings, and J. H. Lee. Receding horizon recursive state estimation. In *Proceedings of the American Control Conference*, pages 900–904, San Francisco, CA, 1993.

[50] J. Liu. Moving horizon state estimation for nonlinear systems with bounded uncertainties. *Chemical Engineering Science*, 93:376–386, 2013.

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

[52] 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.

[53] C. V. Rao and J. B. Rawlings. Constrained process monitoring: Moving-horizon approach. *AIChE J.*, 48:97–109, 2002.

[54] C. C. Qu and J. Hahn. Computation of arrival cost for moving horizon estimation via unscented Kalman filtering. *Journal of Process Control*, 19:358–363, 2009.

[55] S. Ungarala. Computing arrival cost parameters in moving horizon estimation using sampling based filters. *Journal of Process Control*, 19:1576–1588, 2009.

[56] R. Lopez-Negrete, S. C. Patwardhan, and L. T. Biegler. Constrained particle filter approach to approximate the arrival cost in moving horizon estimation. *Journal of Process Control*, 21:909–919, 2011.

[57] F. Thau. Observing the state of nonlinear dynamic systems. *International Journal of Control*, 17:471–479, 1973.

[58] A. J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3:47–52, 1983.

[59] N. H. El-Farra and P. D. Christofides. Robust near-optimal output feedback control of nonlinear systems. *Int. J. Contr.*, 74:133–157, 2001.

[60] N. H. El-Farra and P. D. Christofides. Bounded robust control of constrained multivariable nonlinear processes. *Chemical Engineering Science*, 58:3025–3047, 2003.

[61] E. A. Garcia and P. M. Frank. Deterministic nonlinear observer-based approaches to fault diagnosis: A survey. *Contr. Eng. Prac.*, 5:663–670, 1997.

[62] A. Isidori. *Nonlinear Control Systems: An Introduction*. Springer-Verlag, Berlin-Heidelberg, third edition, 1995.

[63] J. H. Ahrens and H. K. Khalil. High-gain observers in the presence of measurement noise: A switched-gain approach. *Automatica*, 45:936–943, 2009.

[64] J. Zhang and J. Liu. Lyapunov-based MPC with robust moving horizon estimation and its triggered implementation. *AIChE Journal*, 59:4273–4286, 2013.

[65] J. Liu, D. Muñoz de la Peña, B. J. Ohran, P. D. Christofides, and J. F. Davis. A two-tier architecture for networked process control. *Chemical Engineering Science*, 63:5394–5409, 2008.

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

[67] 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.

[68] P. Millan Gata, L. Orihuela Espina, D. Muñoz de la Peña, C. Vivas, and F. R. Rubio. Self-triggered sampling selection based on quadratic programming. In *Proceedings of the 18th IFAC World Congress*, pages 8896–8901, Milan, Italy, 2011.

[69] N. H. El-Farra M. Baldea and B. E. Ydstie. Dynamics and control of chemical process networks: Integrating physics, communication and computation. *Computers and Chemical Engineering*, 51:42–54, 2013.