Simultaneous state and parameter estimation of not fully measured systems: a distributed approach

Siyu Liu [*,***] **Xunyuan Yin** [**] **Jinfeng Liu** [***]

[*] *School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China*
[**] *School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, Singapore, 637459*
[***] *Department of Chemical and Materials Engineering, University of Alberta, Edmonton AB, Canada T6G 1H9 (e-mail: jinfeng@ualberta.ca).*

Abstract: In this paper, we present a distributed approach to simultaneously estimate state and parameter for a class of nonlinear systems, for which the augmented models comprising both the states and parameters are not fully observable. Specifically, we first discuss how sensitivity analysis (SA) can be used to select the best subset of states and parameters for estimation. Then, the entire system is decomposed into several interconnected subsystems as the basis for distributed estimation. Subsequently, a local moving horizon estimator (MHE) is developed based on the corresponding subsystem model, and the local estimators communicate with each other to exchange their estimates. Finally, an SA-based distributed MHE scheme is proposed. The effectiveness of the proposed approach is illustrated using a chemical process consisting of four connected reactors.

Keywords: Subsystem decomposition; sensitivity analysis; distributed estimation; moving horizon estimation, nonlinear systems.

1. INTRODUCTION

Large-scale complex chemical processes are becoming the rule rather than the exception due to their economic benefits. Therefore, most of the existing state estimation methods developed under the centralized framework are not favorable for such processes. Distributed state estimation has been developed as a more flexible and scalable framework for complex processes (Yin et al. (2019); Mahmoud (2016)). Distributed moving horizon estimation (DMHE) is a popular distributed state estimation scheme (Zhang and Liu (2013); Battistelli (2019)). For example, a DMHE was derived for two-time-scale nonlinear systems, which was decomposed into a fast system and several slow subsystems (Yin and Liu (2017)). If all the parameters of a process are known, the state estimation problem can be solved directly by using the algorithms mentioned above and other existing distributed state estimation methods. However, in most applications, there are unknown or uncertain parameters a priori and the states and parameters are confounded in the dynamics. The knowledge of the parameters plays an important role in many process control related activities. Hence, to simultaneously identify the unknown parameters and the unknown states is necessary.

There have been many results on simultaneous state and parameter estimation for both linear systems (Stojanovic et al. (2020)) and nonlinear systems (Liu et al. (2020)). Particularly, simultaneous state-parameter estimation is of importance for fault diagnosis and control. For example, the joint robust estimation algorithm was proposed for the stochastic linear systems with sensor, component and parameter faults (Stojanovic et al. (2020)). However, most of the existing results have a prominent feature, that is, they require the entire system to be observable. In Liu et al. (2021), Liu and co-authors studied the case where the entire system is not fully observable, and proposed to perform simultaneous estimation by selecting the most appropriate state and parameter variables based on the sensitivity analysis. In this work, we extend the results in Liu et al. (2021) to the distributed framework so that large-scale complex processes can be considered.

For distributed simultaneous state and parameter problems, the system observability also plays an important role. However, it is challenging to directly test the observability of the nonlinear system, which involves the calculation of higher-order Lie derivative. Thus, some approximations can be used such as the linearization of the nonlinear system (Nahar et al. (2019)) and the sensitivity analysis of the nonlinear system (Fysikopoulos et al. (2019); Grubben and Keesman (2018)).

Motivated by the above consideration, in this work, we mainly study distributed simultaneous state and parameter estimation when the entire system is not fully observable based on sensitivity analysis. The sensitivity analysis

[⋆] The second author Xunyuan Yin gratefully acknowledges the financial support from Nanyang Technological University, Singapore (Start-Up Grant).

is performed to select which states and parameters should be estimated so that the estimation performance is the best based on the available measurements. After that, the subsystem models for distributed state and parameter estimation is established. Although there are some achievements in distributed state estimation, to our knowledge, this paper is the first work that considers the distributed simultaneous state and parameter estimation. Finally, we develop a DMHE algorithm to achieve our goal based on the results of the sensitivity analysis. It is worth noting that the proposed method is not a trivial extension of the existing results. The contributions of this paper are as follows:

- Simultaneous state and parameter estimation is addressed in a distributed way, when the entire system is not fully observable.
- The distributed moving horizon estimation algorithm which takes advantage of the sensitivity analysis results is presented for distributed simultaneous estimation.

2. PROBLEM STATEMENT AND PRELIMINARIES

In this paper, we consider a class of general nonlinear systems described by

$$\boldsymbol{x}(t+1) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}), \tag{1}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{\theta}), \tag{2}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$, $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$ and $\boldsymbol{y}(t) \in \mathbb{R}^{n_y}$ denote the state, input and output vectors of the nonlinear system, respectively, $\boldsymbol{\theta} \in \mathbb{R}^{n_p}$ is the unknown parameter vector, and $\boldsymbol{f}(\cdot)$ and $\boldsymbol{h}(\cdot)$ represent the nonlinear state and output equations, respectively.

The main objective of this work is to estimate the parameters and states of the system in (1) and (2) in a distributed way. Compared to centralized estimation, it can reduce the computational burden, increase the fault tolerance and improve the maintenance flexibility. This includes a few subobjectives: 1) to determine the most estimable state and parameter subset based on the given measurements when not all the variables can be estimated simultaneously; and 2) to design the distributed estimation scheme and present the distributed moving horizon estimation (DMHE).

3. PROPOSED APPROACH

To achieve the objective, the parameters are first appended to the state vector to construct the augmented system. If the augmented system is observable, the procedure executes the DMHE directly. If not, the sensitivity matrix is constructed and the identifiable states and parameters are selected according to the orthogonalization method based on the sensitivity analysis. After the estimable states and parameters are determined, the original system is decomposed into several subsystems. Finally, the distributed simultaneous estimation of the states and parameters is achieved based on the distributed moving horizon estimation.

3.1 Augmented system construction

In order to achieve simultaneous parameters and state estimation, we consider to append the parameters to the state vector. The following augmented system can be obtained:

$$\boldsymbol{x}_\theta(t+1) = \begin{bmatrix} \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{\theta}(t)) \\ \boldsymbol{\theta}(t) \end{bmatrix} := \boldsymbol{f}_\theta(\boldsymbol{x}_\theta(t), \boldsymbol{u}(t)), \tag{3}$$

$$\boldsymbol{y}_\theta(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{\theta}(t)) := \boldsymbol{h}_\theta(\boldsymbol{x}_\theta(t)), \tag{4}$$

where $\boldsymbol{x}_\theta(t) := [\boldsymbol{x}^{\mathrm{T}}(t), \boldsymbol{\theta}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{n_x+n_p}$ is the augmented state vector, and $\boldsymbol{f}_\theta(\cdot)$ and $\boldsymbol{h}_\theta(\cdot)$ denote the augmented state equation and output equation, respectively.

By estimating the augmented state $\boldsymbol{x}_\theta(t)$, the simultaneous estimation of the state and parameter can be achieved. When $\boldsymbol{x}_\theta(t)$ is observable, various existing state estimation methods can be used to obtain the results. However, in unobservable situations, it will be challenging to obtain accurate estimation results. Therefore, the first step is to check whether the entire augmented state vector $\boldsymbol{x}_\theta(t)$ is observable based on given output measurements.

3.2 Observability of the augmented system

Consider N sampling points from $t-N+1$ to t along a trajectory of the system in (3) and (4). The linearization of the system at each sampling point can be described using the following form with the assumption of zero process and measurement noise (without the loss of generality):

$$\boldsymbol{x}_\theta(t+1) = \boldsymbol{A}_\theta(t)\boldsymbol{x}_\theta(t) + \boldsymbol{B}_\theta(t)\boldsymbol{u}(t), \tag{5}$$

$$\boldsymbol{y}_\theta(t) = \boldsymbol{C}_\theta(t)\boldsymbol{x}_\theta(t), \tag{6}$$

where matrices $\boldsymbol{A}_\theta(t)$, $\boldsymbol{B}_\theta(t)$ and $\boldsymbol{C}_\theta(t)$ are time variant. The observability matrix of Equations (5) and (6) at each sampling point is as follows:

$$\boldsymbol{O}(t) = \begin{bmatrix} \boldsymbol{C}_\theta(t-N+1) \\ \boldsymbol{C}_\theta(t-N+2)\boldsymbol{A}_\theta(t-N+1) \\ \vdots \\ \boldsymbol{C}_\theta(t)\boldsymbol{A}_\theta(t-1)\cdots\boldsymbol{A}_\theta(t-N+1) \end{bmatrix}.$$

If $\boldsymbol{O}(t)$ is full rank, the nonlinear system is locally observable along the trajectory (Proletarsky et al. (2017)).

3.3 Sensitivity matrix

To quantitatively investigate parameter effects, we will illustrate how the sensitivity analysis can be performed. The sensitivity of the output $\boldsymbol{y}_\theta(t)$ with respect to the parameter $\boldsymbol{\theta}$ is defined as $\boldsymbol{S}_{y,\theta}(t) := \frac{\partial \boldsymbol{y}_\theta(t)}{\partial \boldsymbol{\theta}}$. Similarly, the sensitivity of $\boldsymbol{y}_\theta(t)$ with respect to the initial state $\boldsymbol{x}_\theta(0)$ is represented as $\boldsymbol{S}_{y,x(0)}(t) := \frac{\partial \boldsymbol{y}_\theta(t)}{\partial \boldsymbol{x}_\theta(0)}$. In order to calculate the two sensitivities, we first define the sensitivity of the state to the parameter as $\boldsymbol{S}_{x,\theta}(t) := \frac{\partial \boldsymbol{x}_\theta(t)}{\partial \boldsymbol{\theta}}$. According to the nonlinear model shown in (3)–(4), $\boldsymbol{S}_{y,\theta}(t)$ can be computed by solving the following two equations

$$\boldsymbol{S}_{x,\theta}(t+1) = \frac{\partial \boldsymbol{f}_\theta}{\partial \boldsymbol{x}_\theta}\boldsymbol{S}_{x,\theta}(t) + \frac{\partial \boldsymbol{f}_\theta}{\partial \boldsymbol{\theta}}, \tag{7}$$

$$\boldsymbol{S}_{y,\theta}(t) = \frac{\partial \boldsymbol{h}_\theta}{\partial \boldsymbol{x}_\theta}\boldsymbol{S}_{x,\theta}(t) + \frac{\partial \boldsymbol{h}_\theta}{\partial \boldsymbol{\theta}}, \tag{8}$$

with the initial condition $\boldsymbol{S}_{x,\theta}(0) = \boldsymbol{0}$, where

$$\frac{\partial \boldsymbol{f}_\theta}{\partial \boldsymbol{x}_\theta} = \begin{bmatrix} \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}} \\ \boldsymbol{0} & \boldsymbol{I}_{n_p\times n_p} \end{bmatrix}, \quad \frac{\partial \boldsymbol{h}_\theta}{\partial \boldsymbol{x}_\theta} = \begin{bmatrix} \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} & \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}} \end{bmatrix}.$$

To obtain another sensitivity $\boldsymbol{S}_{y,x(0)}(t)$, the initial state $\boldsymbol{x}_\theta(0)$ can be considered as a virtual parameter of the system. Define the sensitivity of the state to the initial condition as $\boldsymbol{S}_{x,x(0)}(t) := \frac{\partial \boldsymbol{x}_\theta(t)}{\partial \boldsymbol{x}_\theta(0)}$. The sensitivity $\boldsymbol{S}_{y,x(0)}(t)$ can be calculated by solving the following equations:

$$\boldsymbol{S}_{x,x(0)}(t+1) = \frac{\partial \boldsymbol{f}_\theta}{\partial \boldsymbol{x}_\theta}\boldsymbol{S}_{x,x(0)}(t), \tag{9}$$

$$\boldsymbol{S}_{y,x(0)}(t) = \frac{\partial \boldsymbol{h}_\theta}{\partial \boldsymbol{x}_\theta}\boldsymbol{S}_{x,x(0)}(t), \tag{10}$$

with the initial value $\boldsymbol{S}_{x,x(0)}(0) = \boldsymbol{I}$.

Based on Equations (9) and (10), the sensitivity $\boldsymbol{S}_{y,x_\theta(0)}(t)$ of the linearized system in (5) and (6) can be rewritten in the following at the sampling point t:

$$\boldsymbol{S}_{y,x_\theta(0)}(t) = \boldsymbol{C}_\theta(t)\boldsymbol{A}_\theta(t-1)\boldsymbol{A}_\theta(t-2)\cdots\boldsymbol{A}_\theta(0).$$

For each sampling time t, we can collect the most recent N sensitivities $\boldsymbol{S}_{y,x_\theta(t-N+1)}(i)$, $i = t, t-1, \ldots, t-N+1$, to form a sensitivity matrix $\boldsymbol{S}_\theta(t)$ as follows, and test the rank of the following sensitivity matrix along a typical trajectory in a data window of the augmented system:

$$\boldsymbol{S}_\theta(t) = \begin{bmatrix} \boldsymbol{S}_{y,x_\theta(t-N+1)}(t-N+1) \\ \boldsymbol{S}_{y,x_\theta(t-N+1)}(t-N+2) \\ \vdots \\ \boldsymbol{S}_{y,x_\theta(t-N+1)}(t) \end{bmatrix}, \tag{11}$$

where N is the data window length.

Remark 1. According to the derivations of $\boldsymbol{O}(t)$ and $\boldsymbol{S}_\theta(t)$, we can find that $\boldsymbol{O}(t)$ and $\boldsymbol{S}_\theta(t)$ include the same information from $t-N+1$ to t. Therefore, $\boldsymbol{S}_\theta(t)$ has been used as an indication of the observability of nonlinear systems because it can capture the dynamics of the system. By checking the rank of the sensitivity matrix $\boldsymbol{S}_\theta(t)$ at each sampling time, we can conclude whether the entire augmented state vector $\boldsymbol{x}_\theta$ can be estimated locally using the input and output information. Especially, the sensitivity matrix has been used in the parameter selection (Stigter and Molenaar (2015); Kim and Lee (2019)).

3.4 Parameter selection based on the orthogonalization method

If the sensitivity matrix $\boldsymbol{S}_\theta(t)$ is full rank along all of the sampling points and is well-conditioned, the augmented state vector $\boldsymbol{x}_\theta(t)$ can be estimated locally using the input and output data. However, when $\boldsymbol{S}_\theta(t)$ is not full rank along all of the sampling points, this means that not all of the variables in the augmented state vector $\boldsymbol{x}_\theta(t)$ can be estimated. In view of this situation, we aim to resort to the information contained in the sensitivity matrix $\boldsymbol{S}_\theta(t)$ to select those variables that are important for the output prediction to estimate. The larger the sensitivity of one parameter, the more sensitive the system response with respect to small perturbations of this parameter.

Here, we apply the orthogonalization method to select the variables. The procedure for sequentially selecting the most important parameters for simultaneous state and parameter estimation using the orthogonalization method is as follows:

(1) At time t, calculate the norm of each column of the normalized $\boldsymbol{S}_\theta(t)$.
(2) Initialize $l = 1$, select the parameter whose column in $\boldsymbol{S}_\theta(t)$ has the largest norm, and denote the corresponding column as $\boldsymbol{X}_l$.
(3) Project the sensitivity vectors of the unselected parameters onto the space orthogonal to the space formed by the sensitivity vectors of the previously selected parameters, and compute the orthogonal projection matrix that cannot be expressed by $\boldsymbol{X}_l$: $\boldsymbol{P}_l = \boldsymbol{I} - \boldsymbol{X}_l(\boldsymbol{X}_l^{\mathrm{T}}\boldsymbol{X}_l)^{-1}\boldsymbol{X}_l^{\mathrm{T}}$.
(4) Calculate the norm of each column of the residual matrix: $\boldsymbol{R}_l = \boldsymbol{P}_l\boldsymbol{S}_\theta(t)$, select the column from $\boldsymbol{S}_\theta(t)$ that corresponds to the column with the largest norm in $\boldsymbol{R}_l$, and add it to $\boldsymbol{X}_l$ as a new column to form $\boldsymbol{X}_{l+1}$.
(5) If the largest norm of the columns of $\boldsymbol{R}_l$ is larger than a prescribed cutoff value, go to Step 3 with $l := l+1$; otherwise, obtain the selected parameters, terminate this process.

Remark 2. The projection aims to remove the parameter's effect on the output covered by the previously selected parameters. In each step, we select the parameter which has the largest not yet covered effect.

3.5 DMHE based on the sensitivity analysis

Following the algorithm in the previous subsections, we know the subset of states and parameters that can be estimated. Taking that into account, we should proceed to decompose the entire system into subsystems for distributed estimation. The decomposition may be performed based on, for example, Yin and Liu (2019). Due to the page limit, this is omitted in this paper. It is noted that in the decomposition, special attention should be given to the calculation of the adjacent matrix.

When the system is decomposed, the distributed simultaneous state and parameter estimation is developed in the framework of the distributed MHE. Specifically, at time t, after selecting variables according to the sensitivity matrix $\boldsymbol{S}_\theta(t)$, it determines that the elements of $\boldsymbol{x}_\theta(t)$ can be estimated based on the input-output data from $t-N$ to t. In the proposed MHE design for each subsystem, the estimation window used in each MHE is considered to be the same as the data window length N used in $\boldsymbol{S}_\theta(t)$. The design of the proposed distributed MHE for each subsystem i, $i = 1, 2, \ldots$, at time t based on the augmented system (3)–(4) is described as follows:

$$\min_{\tilde{x}_\theta^i(t-N),\ldots,\tilde{x}_\theta^i(t)} \Bigg\{ \sum_{l=t-N}^{t-1} \|\hat{\boldsymbol{w}}_\theta^i(l)\|^2_{Q_i^{-1}} + \sum_{j=t-N}^{t} \|\hat{\boldsymbol{v}}^i(j)\|^2_{R_i^{-1}} \\ \Gamma^i(\tilde{\boldsymbol{x}}_\theta^i(t-N)) \Bigg\}, \tag{12}$$

$$\text{s.t.}\,\tilde{\boldsymbol{x}}_\theta^i(l+1) = f_\theta^i(\tilde{\boldsymbol{x}}_\theta^i(l), \boldsymbol{u}^i(l), \boldsymbol{w}_\theta^i(l)) + \tilde{f}^i(\hat{\boldsymbol{\chi}}_\theta^i(t)), \tag{13}$$

$$y^i(l) = h_\theta^i(\hat{\boldsymbol{x}}_\theta^i(l)) + \boldsymbol{v}^i(l), \tag{14}$$

$$\tilde{\boldsymbol{x}}_\theta^i(l) \in \mathbb{X}_\theta^i, \quad \boldsymbol{v}^i(l) \in \mathbb{V}^i, \quad l = t-N, \ldots, t, \tag{15}$$

$$\boldsymbol{w}_\theta^i(l) \in \mathbb{W}_\theta^i, \quad l = t-N, \ldots, t-1, \tag{16}$$

$$\tilde{\boldsymbol{x}}_{\theta,u}^i(t-N) = \tilde{\boldsymbol{x}}_{\theta,u}^i(t-N|t-1), \ u \in U(t), \tag{17}$$

$$\boldsymbol{w}_{\theta,u}^i(l) = 0, \ u \in U(t), \ l = t-N, \ldots, t-1, \tag{18}$$

where $\tilde{\boldsymbol{x}}_\theta^i$ is the prediction of $\boldsymbol{x}_\theta^i$ within the optimization problem, N is the estimation horizon, $\hat{\boldsymbol{x}}_\theta^i$, $\boldsymbol{Q}_i$ and $\boldsymbol{R}_i$ denote the covariance matrices of $\boldsymbol{w}^i$ and $\boldsymbol{v}^i$, respectively, $\hat{\boldsymbol{\chi}}_\theta$ represents the estimate obtained by other MHEs. Note that the subsystems communicate with each other and the communicated information is used in interactive compensation.

Once the optimization problem (12)–(18) is solved, a series of solution is determined as $\{\tilde{x}_\theta^{i*}(t-N), \ldots, \tilde{x}_\theta^{i*}(t)\}$, of which the last element $\tilde{x}_\theta^{i*}(t)$ is adopted as the current optimal estimate and denoted as $\hat{x}_\theta^i(t)$.

In the optimization problem (12)–(18), Equation (12) is the cost function for each MHE estimator, and the last term $\Gamma^i(\tilde{\boldsymbol{x}}_\theta^i(t-N))$, called the arrival cost in the MHE, summarizes the previous information of the measurements before the current window. However, an algebraic expression for the arrival cost only exists sometimes, such as the linear unconstrained case. Hence, we choose a quadratic arrival cost with a constant weighting matrix as an approximated arrival cost:

$$\Gamma(\hat{\boldsymbol{x}}_\theta^i(t-N)) = \|\hat{\boldsymbol{x}}_\theta^i(t-N) - \hat{\boldsymbol{x}}_\theta^i(t-N|t-1)\|^2_{p_i^{-1}},$$

where $\boldsymbol{P}$ is a constant weighting matrix and $\tilde{\boldsymbol{x}}_\theta^i(t-N|t-1)$ is the estimate of $\tilde{\boldsymbol{x}}_\theta^i(t-N)$ obtained at the previous time instant $t-1$. It has been shown that, under quite general assumptions, such a simple arrival cost is sufficient to ensure the convergence of the algorithm provided that $\boldsymbol{P}$ is adequately chosen to avoid an overconfidence on the available estimates (Alessandri et al. (2008)). Equations (13) and (14) are the subsystem model constraints, while Equations (15) and (16) consider the constraints on system states, measurement noise, and system disturbances. Compared with other distributed MHE, the difference of the algorithm proposed in this paper is that Equations (17) and (18) take into account the constraints of the variable selection results for each subsystem. Here, let $U(t)$ denote the set containing the indices of the unselected elements of $\boldsymbol{x}_\theta(t)$ based on the variable selection presented in subsection 3.4. If $x_{\theta,3}$, $x_{\theta,7}$ and $x_{\theta,8}$ are not selected, then $U(t) = \{3, 7, 8\}$. This means that the elements in subsystem disturbance vector corresponding to the unselected variable are zero, that is, $\boldsymbol{w}_{\theta,u}^i = 0$ $(u \in U(t))$. In this case, the unselected variables $\tilde{\boldsymbol{x}}_{\theta,u}^i$ will only evolve in an open-loop manner according to the subsystem model, and the initial condition $\tilde{\boldsymbol{x}}_{\theta,u}^i(t-N)$ is designated as the value obtained at the previous instant.

For the proposed distributed framework, the each MHE estimator is required to exchange information with each other. Therefore, the interaction among the configured subsystems needs to be considered. At each sampling time, each local MHE is performed to provide the state estimates of the subsystem based on the information gathered from its associated subsystem as well as its interactive subsystems.

Remark 3. The method proposed in this work does not require accurate nominal parameters. Rough initial guess $\hat{\boldsymbol{x}}_\theta^i(0)$ would suffice for an SA based DMHE algorithm. The variable selection based on the sensitivity analysis and distributed state estimation run online simultaneously. Parameters and states are updated through the distributed MHE, and they are used for the next sensitivity calculations.

4. ILLUSTRATIVE EXAMPLE

In this section, we apply the proposed method to a chemical process consisting of four connected continuous-stirred tank reactors (CSTRs) with different volumes (Liu et al. (2022)). Based on mass and energy balances, a model that comprises eight differential equations describes the process dynamics:

$$\begin{aligned}\frac{dC_{A1}}{dt} =& \frac{F_{01}}{V_1}(C_{A01} - C_{A1}) + \frac{F_{r1}}{V_1}(C_{A2} - C_{A1}) \\ &+ \frac{F_{r2}}{V_1}(C_{A4} - C_{A1}) - \sum_{i=1}^{3} k_i e^{\frac{-E_i}{RT_1}} C_{A1},\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{dT_1}{dt} =& \frac{F_{01}}{V_1}(T_{01} - T_1) + \frac{F_{r1}}{V_1}(T_2 - T_1) + \frac{F_{r2}}{V_1}(T_4 - T_1) \\ &- \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_i e^{\frac{-E_i}{RT_1}} C_{A1} + \frac{Q_1}{\rho c_p V_1},\end{aligned} \tag{20}$$

$$\begin{aligned}\frac{dC_{A2}}{dt} =& \frac{F_1}{V_2}(C_{A1} - C_{A2}) + \frac{F_{02}}{V_2}(C_{A02} - C_{A2}) \\ &- \sum_{i=1}^{3} k_i e^{\frac{-E_i}{RT_2}} C_{A2},\end{aligned} \tag{21}$$

$$\begin{aligned}\frac{dT_2}{dt} =& \frac{F_1}{V_2}(T_1 - T_2) + \frac{F_{02}}{V_2}(T_{02} - T_2) \\ &- \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_i e^{\frac{-E_i}{RT_2}} C_{A2} + \frac{Q_2}{\rho c_p V_2},\end{aligned} \tag{22}$$

$$\begin{aligned}\frac{dC_{A3}}{dt} =& \frac{F_2 - F_{r1}}{V_3}(C_{A2} - C_{A3}) + \frac{F_{03}}{V_3}(C_{A03} - C_{A3}) \\ &- \sum_{i=1}^{3} k_i e^{\frac{-E_i}{RT_3}} C_{A3},\end{aligned} \tag{23}$$

$$\begin{aligned}\frac{dT_3}{dt} =& \frac{F_2 - F_{r1}}{V_3}(T_2 - T_3) + \frac{F_{03}}{V_3}(T_{03} - T_3) \\ &- \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_i e^{\frac{-E_i}{RT_3}} C_{A3} + \frac{Q_3}{\rho c_p V_3},\end{aligned} \tag{24}$$

$$\begin{aligned}\frac{dC_{A4}}{dt} =& \frac{F_3}{V_4}(C_{A3} - C_{A4}) + \frac{F_{04}}{V_4}(C_{A04} - C_{A4}) \\ &- \sum_{i=1}^{3} k_i e^{\frac{-E_i}{RT_4}} C_{A4},\end{aligned} \tag{25}$$

$$\begin{aligned}\frac{dT_4}{dt} =& \frac{F_3}{V_4}(T_3 - T_4) + \frac{F_{04}}{V_4}(T_{04} - T_4) \\ &- \sum_{i=1}^{3} \frac{\Delta H_i}{\rho c_p} k_i e^{\frac{-E_i}{RT_4}} C_{A4} + \frac{Q_4}{\rho c_p V_4},\end{aligned} \tag{26}$$

For this process, the heating inputs Q_i to the four vessels are usually chosen to be the input variable, C_{Ai} and T_i are the system states, where T_i is online measured and is used as the output measurements. Each measurement is associated with a subsystem and a local estimator is developed based on the measurement. The values of the model parameters are shown in Table 1. The continuous model is discretized using the fourth-order Runge-Kutta method with a sample time $\Delta t = \frac{1}{120}$h.

Since the unknown parameters in Table 1 cannot be known exactly, we need to estimate the state variables

Table 1. Parameter values of the four-CSTR

Known parameters	
$T_{01} = 300$ K	$\Delta H_1 = -5.0 \times 10^4$ kJ/kmol
$T_{03} = 300$ K	$\Delta H_2 = -5.2 \times 10^4$ kJ/kmol
$T_{02} = 300$ K	$\Delta H_3 = -5.0 \times 10^4$ kJ/kmol
$T_{04} = 300$ K	$k_1 = 3.0 \times 10^6$ h^{-1}
$c_p = 0.231$ kJ/(kg·K)	$k_2 = 3.0 \times 10^5$ h^{-1}
$\rho = 1000$ kg/m^3	$k_3 = 3.0 \times 10^5$ h^{-1}
Unknown parameters to be estimated	
$F_{01} = 5$ m^3/h	$C_{01} = 4.0$ kmol/m^3
$F_{02} = 10$ m^3/h	$C_{02} = 2.0$ kmol/m^3
$F_{03} = 8$ m^3/h	$C_{03} = 3.0$ kmol/m^3
$F_{04} = 12$ m^3/h	$C_{04} = 3.5$ kmol/m^3
$V_1 = 1$ m^3	$F_1 = 35$ m^3/h
$V_2 = 3$ m^3	$F_2 = 45$ m^3/h
$V_3 = 4$ m^3	$F_3 = 33$ m^3/h
$V_4 = 6$ m^3	$F_{r1} = 20$ m^3/h
$E_1 = 5.0 \times 10^4$ kJ/kmol	$F_{r2} = 10$ m^3/h
$E_2 = 7.5 \times 10^4$ kJ/kmol	$R = 8.314$ kJ/(kmol·K)
$E_3 = 7.53 \times 10^4$ kJ/kmol	

and parameter variables of the four-CSTR process by extracting as much information as possible from the four output measurements. The first step is to construct the augmented system, and 21 unknown parameters are considered in the augmented system. The state vector $\boldsymbol{x} = [C_{A1}, T_1, C_{A2}, T_2, C_{A3}, T_3, C_{A4}, T_4]^{\mathrm{T}}$ and parameter vector $\boldsymbol{\theta} = [F_{0i}, V_i, C_{0i}, E_1, E_2, E_3, F_1, F_2, F_3, F_{r1}, F_{r2}, R]^{\mathrm{T}}$ $(i = 1, 2, 3, 4)$ constructs the augmented state as follows:

$$\boldsymbol{x}_\theta := [\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}}]^{\mathrm{T}}.$$

The four-CSTR process has a steady state

$$\begin{aligned}\boldsymbol{x}_s =& [2.78\mathrm{kmol/m^3}, 363\mathrm{K}, 2.58\mathrm{kmol/m^3}, 356\mathrm{K},\\ & 2.6\mathrm{kmol/m^3}, 355\mathrm{K}, 2.6\mathrm{kmol/m^3}, 392\mathrm{K}]^{\mathrm{T}}.\end{aligned}$$

Next, we normalize the augmented model around the augmented steady state $\boldsymbol{x}_{\theta,s}$ formed by the above steady state $\boldsymbol{x}_s$ and the nominal parameter values in Table 1 to avoid the potential impact of the parameter value in rank calculation.

In the simulations, the constant heat inputs to the four vessels are selected as: $Q_1 = 1.0 \times 10^4$kJ/h, $Q_2 = 2.0 \times 10^4$kJ/h, $Q_3 = 2.5 \times 10^4$kJ/h, and $Q_4 = 1.0 \times 10^4$kJ/h. In the variable selection, a predetermined cut-off value is required to terminate the variable selection process. Here, we propose to use the cut-off value $\alpha = 3\sqrt{\bar{\boldsymbol{\omega}}^2 + \bar{\boldsymbol{v}}^2}$ (Liu et al. (2021)), where $\bar{\boldsymbol{\omega}} = 10^{-3}$ and $\bar{\boldsymbol{v}} = 10^{-3}$ denote the standard deviation of the normalized process noise and measurement noise, respectively. In this way, the influence of noises on different elements in the normalized model is basically at the same level.

In the variable selection, the eight original states are considered important and must be estimated at each sampling time. In this case, the variable selection is only performed among the parameters. At each sampling time, the information represented by the eight original states needs to be removed from the obtained sensitivity matrix $\boldsymbol{S}_\theta(k)$, and then some additional estimable parameters are selected in turn using the residual matrix. Table 2 shows the number of sampling times of each state and parameter when the total sampling times are 500. It can be seen that nine parameters are selected based on the sensitivity information and will be estimated together with the states.

The following is to estimate the system states and parameters based on the four outputs using the DMHE based on four subsystems. The covariance parameters of each MHE are tuned as $\boldsymbol{Q}_i = \mathrm{diag}\{0.05^2, 0.05^2\}$, $i = 1, 2, 3, 4$, and $R_i = 0.05^2$, and the matrix $\boldsymbol{P}_1 = \mathrm{diag}\{0.1^2\boldsymbol{I}_2, 0.07^2\boldsymbol{I}_3\}$, $\boldsymbol{P}_2 = \boldsymbol{P}_3 = \boldsymbol{P}_4 = \mathrm{diag}\{0.1^2\boldsymbol{I}_2, 0.07^2\boldsymbol{I}_2\}$.

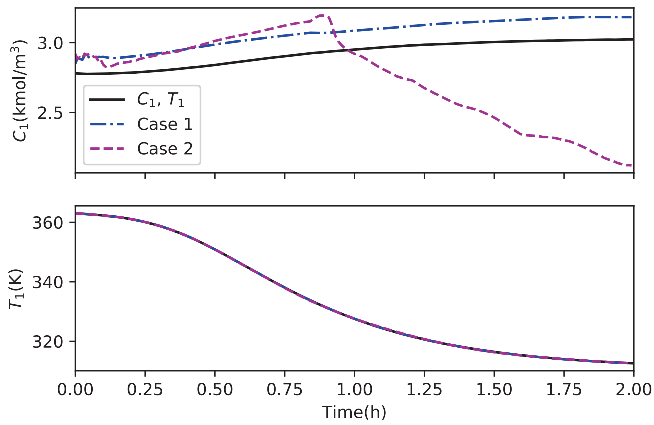

Fig. 1. Trajectories of the actual states C_1, T_1, and state estimates in cases 1 and 2

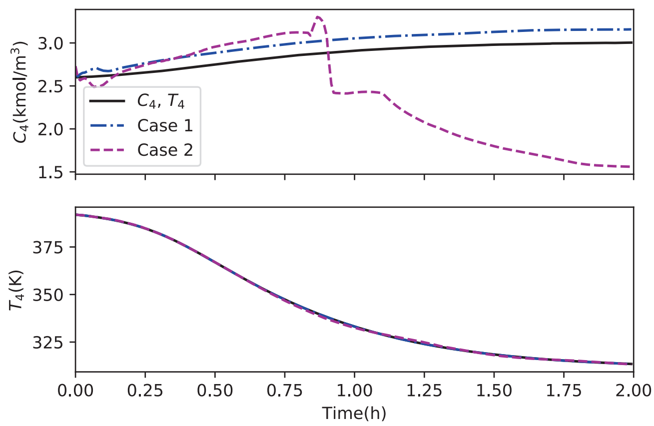

Fig. 2. Trajectories of the actual states C_4, T_4, and state estimates in cases 1 and 2

To verify the effectiveness of the proposed method, we conduct the two different cases. In case 1, the distributed MHE is considered to estimate the most appropriate 17 variables in Table 3 selected by the variable selection algorithm based on the sensitivity analysis. In case 2, all of the 29 variables in augmented state are estimated simultaneously without the sensitivity analysis. That means the insensitive variables are randomly assigned to the four subsystems in case 2.

In both cases, a 5% mismatch in the initial state of each of the eight original states is considered. It is also assumed that the parameters are not known exactly and there is a 5% mismatch in each of the parameters. The two cases use the same input and output data. The simulation results of two cases are shown in Figures 1-3 (Due to too many system variables, we show the estimation results of 4 states and 4 parameters). From Figures 1-3, it can be

Table 2. Number of sampling times for each variable

	C_{A1}	T_1	C_{A2}	T_2	C_{A3}	T_3	C_{A4}	T_4	F_1	F_2	F_3	V_1	V_2	V_3	V_4
Count	499	500	499	500	499	500	499	500	0	0	0	498	498	498	498
	F_{r1}	F_{r2}	E_1	E_2	E_3	R	F_{01}	F_{02}	F_{03}	F_{04}	C_{01}	C_{02}	C_{03}	C_{04}	
Count	0	497	0	0	0	0	497	497	497	496	0	0	0	0	

Table 3. Subsystem description for cases 1, 2

Case 1	States	Inputs
sub 1	C_1, $\mathbf{T_1}$, F_{01}, V_1, F_{r2}	Q_1
sub 2	C_2, $\mathbf{T_2}$, F_{02}, V_2	Q_2
sub 3	C_3, $\mathbf{T_3}$, F_{03}, V_3	Q_3
sub 4	C_4, $\mathbf{T_4}$, F_{04}, V_4	Q_4
Case 2	States	Inputs
sub 1	C_1, $\mathbf{T_1}$, F_{01}, V_1, C_{01}, F_{r1}, F_{r2}, R	Q_1
sub 2	C_2, $\mathbf{T_2}$, F_{02}, V_2, C_{02}, F_1, E_1	Q_2
sub 3	C_3, $\mathbf{T_3}$, F_{03}, V_3, C_{03}, F_2, E_2	Q_3
sub 4	C_4, $\mathbf{T_4}$, F_{04}, V_4, C_{04}, F_3, E_3	Q_4

seen that the state estimation performance of case 2 is much poorer compared with case 1. The main reason is that the rank of the sensitivity matrix of the augmented system is 17, which is smaller than 29. This means that ignoring the observability when estimating the augmented state vector $\boldsymbol{x}_\theta$ will result in bad estimation results. In case 1, the observability is considered, and a subset of variables selected based on the sensitivity matrix are estimated.

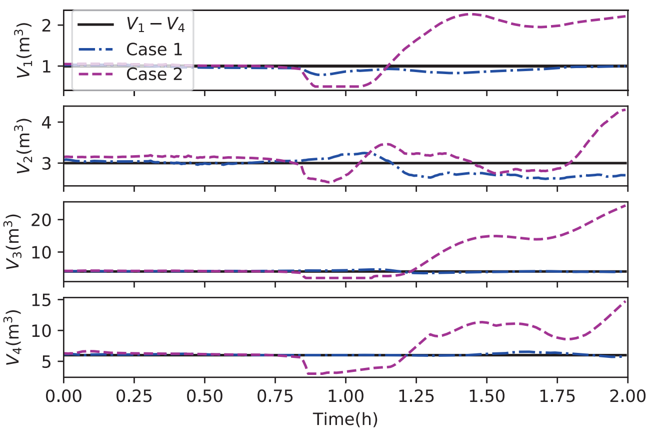

Fig. 3. Trajectories of the actual parameters V_1, V_2, V_3, V_4, and parameter estimates in cases 1 and 2

REFERENCES

Alessandri, A., Baglietto, M., & Battistelli, G. (2008). Moving-horizon state estimation for nonlinear discrete-time systems: New stability results and approximation schemes. *Automatica*, 44, 1753-1765.

Battistelli, G. (2019). Distributed moving-horizon estimation with arrival-cost consensus. *IEEE Transactions on Automatic Control*, 64(8), 3316-3323.

Fysikopoulos, D., Benyahia, B., Borsos, A., Nagy, Z.K., & Rielly, C.D. (2019). A framework for model reliability and estimability analysis of crystallization processes with multi-impurity multi-dimensional population balance models. *Computers & Chemical Engineering*, 122, 275-292.

Grubben, N.L.M., & Keesman, K.J. (2018). Controllability and observability of 2D thermal flow in bulk storage facilities using sensitivity fields. *International Journal of Control*, 97(7), 1554-1566.

Kim, B., & Lee, J.H. (2019). Parameter subset selection and biased estimation for a class of ill-conditioned estimation problems. *Journal of Process Control*, 81, 65-75.

Liu, J.B., Gnanasekar, A., Zhang, Y., et al. (2021). Simultaneous state and parameter estimation: the role of sensitivity analysis. *Industrial & Engineering Chemistry Research*, 60(7), 2971-2982.

Liu, S.Y., Ding, F., & Hayat, T. (2020). Moving data window gradient-based iterative algorithm of combined parameter and state estimation for bilinear systems. *International Journal of Robust and Nonlinear Control*, 30(6), 2413-2429.

Liu, S.Y., Yin, X., Liu, J.B., Liu, J., & Ding, F. (2022). Distributed simultaneous state and parameter estimation of nonlinear systems. *Chemical Engineering Research and Design*, 118, 74-86.

Mahmoud, M.S. (2016). Distributed estimation based on information-based covariance intersection algorithms. *International Journal of Adaptive Control and Signal Processing*, 30(5), 750-778.

Nahar, J., Liu, J., & Shah, S.L. (2019). Parameter and state estimation of an agro-hydrological system based on system observability analysis. *Computers & Chemical Engineering*, 121, 450-464.

Proletarsky, A.V., Neusypin, K.A., Shen, K., Selezneva, M.S., & Grout, V. (2017). Development and analysis of the numerical criterion for the degree of observability of state variables in nonlinear systems. *2017 Internet Technologies and Applications*, 150-154.

Stigter, J.D., & Molenaar, J. (2015). A fast algorithm to assess local structural identifiability. *Automatica*, 58, 118-124.

Stojanovic, V., He, S.P., & Zhang, B.Y. (2020). State and parameter joint estimation of linear stochastic systems in presence of faults and non-Gaussian noises. *International Journal of Robust and Nonlinear Control*, 30(16), 6683-6700.

Yin, X., & Liu, J. (2017). Distributed moving horizon state estimation of two-time-scale nonlinear systems. *Automatica*, 79, 152-161.

Yin, X., & Liu, J. (2019). Subsystem decomposition of process networks for simultaneous distributed state estimation and control. *AIChE Journal*, 65, 904-914.

Yin, X., Zeng, J., & Liu, J. (2019). Forming distributed state estimation network from decentralized estimators. *IEEE Transactions on Control Systems Technology*, 27(6), 2430-2443.

Zhang, J., & Liu, J. (2013). Distributed moving horizon state estimation for nonlinear systems with bounded uncertainties. *Journal of Process Control*, 23(9), 1281-1295.