

ORIGINAL RESEARCH PAPER

State filtering and parameter estimation for two-input two-output systems with time delay

Ya Gu¹  | Quanmin Zhu² | Chuanjiang Li³ | Peiyi Zhu¹ | Hassan Nouri²

¹ School of Electrical Engineering and Automation, Changshu Institute of Technology, Changshu, Jiangsu 215500, P. R. China

² Department of Engineering Design and Mathematics, University of the West of England, Bristol BS16 1QY, UK

³ College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, P. R. China

Correspondence

Ya Gu, School of Electrical Engineering and Automation, Changshu Institute of Technology, Changshu, Jiangsu 215500, P. R. China.
Email: guya927819@163.com

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 61903050, 61803049; Natural Science Foundation of Jiangsu Province, Grant/Award Number: BK20181033; Natural Science Fundamental Research Project of Colleges and Universities in Jiangsu province, Grant/Award Numbers: 19KJB470009, 17KJB510002

Abstract

This paper focuses on presenting a new identification algorithm to estimate the parameters and state variables for two-input two-output dynamic systems with time delay based on canonical state space models. First, the related input-output equation is determined and transformed into an identification oriented model, which does not involve in the unmeasurable states, and then a residual based least squares identification algorithm is presented for the estimations. After the parameters being estimated, the system states are subsequently estimated by using the estimated parameters. Through theoretical analysis, the convergence of the algorithm is derived to provide assurance for applicability. Finally, a selected simulation example is given for a meaningful case study to show the effectiveness of the proposed algorithm.

1 | INTRODUCTION

Parameter estimation and state identification have been the core issues of data driven modelling, signal filtering and controller design [1–4]. The mathematical model is the foundation to quantitatively represent the system, and the system identification makes use of statistical algorithms to develop the mathematical model of the dynamic systems from the known information [5–8]. The iterative/recursive algorithms are the typical parameter identification algorithms [9], which have a wide range of applications in seeking the roots of the equation and developing parameter estimation methods [10–14]. Ansari and Bernstein proposed the deadbeat unknown-input state estimation and input reconstruction for linear discrete-time systems [15]; Xu et al. presented a hierarchical Newton and least squares iterative estimation algorithm for dynamic systems based on the impulse responses [16]. Recently, Xia et al presented an improved least-squares identification for multiple-output non-

linear stochastic systems [17]. Telmoudi et al discussed the modelling and state of health estimation of nickel-metal hydride battery using an EPSO-based fuzzy c-regression model [18] and the parameter estimation of nonlinear systems using a robust possibilistic c-regression model algorithm [19]. In this work, the differential equations are expanded to the two-input two-output model with time delay, which is difficult from the point of view of identification and the time delay system is generally ubiquitous in industry.

Regarding the model structures, state space models have been predominantly adopted in system identification and control system design [20–24], and thus have received much research attention in parameter and state estimation over decades [25, 26] and witnessed the applications [27]. In engineering applications, it should be noted that the states of some systems are not completely known due to various reasons (e.g. no available sensors, heavy cost in measurements, etc.) Therefore, the system state estimation has played an important role in control design and

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Control Theory & Applications* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

system identification. There are many state and parameter estimation algorithms: Meurer et al. proposed the nonlinear state estimation for the Czochralski process based on the weighing signal using an extended Kalman filter [28]; Alessandri and Gaggero discussed the fast moving horizon state estimation for discrete-time systems using single and multi iteration descent methods [29].

The complexity and uncertainty of the analysis or sampling often accompanied with the outputs subject to uncertain delays [30, 31]. The existence of time delay makes it difficult for the control system to respond to timely changes in input [32, 33]. Besides, the time delay can lead to instability and poor performance of controlled processes. The parameter and delay estimation of such systems is a challenging problem and meaningful in academic research and applications which has attracted a lot of attention, especially in the case of measurement noise. Some useful techniques have been introduced in this aspect. Stojanovic discussed the robust finite-time stability of discrete time systems with interval time-varying delay and nonlinear perturbations [34]. For multivariate delayed state space models, more research has paid attention to parameter estimation and status estimation, ignored the computation demanding of the algorithms [35]. The expectation maximization (EM) algorithm has been a widely used for computing maximum likelihood estimates of unknown parameters in probabilistic models involving latent variables. The EM algorithm takes up an iterative process that alternates between computing a conditional expectation and solving a maximization problem. However EM cannot be directly used to estimate state variables.

This article investigates a residual-based least squares identification algorithm to simultaneously estimate states and parameters of a class of two-input two-output systems. Based on the thought of decomposition in the identification model, the two-input two-output system is decomposed into two less dimension and variables two-input single-output subsystems, again to identify each subsystem. To overcome the difficulty of the information matrix including unmeasurable noise terms, the unknown noise terms are replaced with their estimated residuals, which are computed through the preceding parameter estimates. The simulation results are provided to show the computational experimental validity tests.

The contributions of the study, in terms of reducing computational complexity/demanding in parameter and state estimation, lie in three aspects:

- The two-input two-output model with time delay is decomposed into two two-input single-output models with few dimensions and few variables based on the idea of identification model decomposition.
- The presented algorithm can make full use of all data to generate highly accurate parameter estimates.
- The deducing process of the identification model is simplified to reduce the computational load of multivariable system identification.

Regarding the related research, it should be noted that the gradient iterative identification algorithm has a small amount of

calculation, but low estimation accuracy, and slow convergence speed [36]. The least squares algorithm has high accuracy, but it has heavy computational demand [37]. This study presents a hierarchical identification algorithm to decompose the identification system into two subsystems, reducing the dimensionality of the covariance matrix, reducing the computational load, and improving the estimation accuracy by filtering the input and output data from noise.

The communicate is organized as follows. Section 2 gives the identification model for two-input two-output state space system with time delay. Section 3 derives a parameter identification algorithm for canonical state space systems with time delay. Section 4 presents the state estimation identification algorithm. Section 5 provides an illustrative example for the results in this study. Finally, we offer some concluding remarks in Section 6.

2 | THE PROBLEM FORMULATION

Consider the following model describing two-input two-output state space system with time delay,

$$\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{x}(t-d) + \mathbf{F}\mathbf{u}(t), \quad (1)$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{v}(t). \quad (2)$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is the unmeasurable state vector, $\mathbf{u}(t) = [u_1(t), u_2(t)]^T \in \mathbb{R}^2$ is the input vector, $\mathbf{y}(t) = [y_1(t), y_2(t)]^T \in \mathbb{R}^2$ is the output vector, and $\mathbf{v}(t) = [v_1(t), v_2(t)]^T \in \mathbb{R}^2$ is the uncorrelated stochastic noise with zero mean. Assume that n is known, and $\mathbf{u}(t) = \mathbf{0}$ for $t \leq 0$. Since it is a multivariable system with coupling, the matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{F} \in \mathbb{R}^{n \times 2}$ and $\mathbf{c} \in \mathbb{R}^{2 \times n}$ are the system parameter matrices to be identified.

$$\mathbf{A} := \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_2 \end{bmatrix} \in \mathbb{R}^{n \times n},$$

$$\mathbf{A}_i := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ a_i(1) & a_i(2) & a_i(3) & \cdots & a_i(n_i) \end{bmatrix} \in \mathbb{R}^{n_i \times n_i},$$

$$i, j = 1, 2, \quad i \neq j,$$

$$\mathbf{A}_{ij} := \begin{bmatrix} 0 & & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & & \cdots & 0 \\ a_{ij}(1) & \cdots & a_{ij}(n_{ij}) & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times n_j},$$

$$n_{ij} \leq \begin{cases} n_i + 1, & i > j, \\ n_i, & i \leq j, \end{cases}$$

$$\mathbf{B} := [\mathbf{B}_1^T, \mathbf{B}_2^T]^T \in \mathbb{R}^{n \times n},$$

$$\mathbf{B}_i := [\mathbf{b}_{i1}^T, \mathbf{b}_{i2}^T, \dots, \mathbf{b}_{in_i}^T]^T \in \mathbb{R}^{n_i \times n}, \quad \mathbf{b}_{ik} \in \mathbb{R}^{1 \times n},$$

$$F := [F_1^T, F_2^T]^T \in \mathbb{R}^{n \times 2},$$

$$F_i := [f_{i1}^T, f_{i2}^T, \dots, f_{in_i}^T]^T \in \mathbb{R}^{n_i \times 2}, f_{ik} \in \mathbb{R}^{1 \times 2},$$

$$C := \begin{bmatrix} e_1^T & 0 \\ 0 & e_2^T \end{bmatrix} \in \mathbb{R}^{2 \times n}, e_i = [1, 0, \dots, 0]^T \in \mathbb{R}^{n_i}.$$

Here, $n_i \geq 1$ are the observability indices, satisfying $n_1 + n_2 = n$. Let $\mathbf{x}(t) = [\mathbf{x}_1^T(t), \mathbf{x}_2^T(t)]^T \in \mathbb{R}^n$, $\mathbf{x}_i(t) \in \mathbb{R}^{n_i}$. Because the model (1)–(2) contains the unknown parameter vectors/matrices of the system and unmeasurable state vectors, which is the difficulty of identification, the idea of this paper is to replace the state vector with a measurable input and output. First analyze the first subsystem, we have

$$\begin{aligned} y_1(t+i) &= e_1^T A_1^i \mathbf{x}_1(t) + e_1^T A_1^{i-1} A_{12} \mathbf{x}_2(t) \\ &+ e_1^T A_1^{i-2} A_{12} \mathbf{x}_2(t+1) \\ &+ \dots + e_1^T A_{12} \mathbf{x}_2(t+i-1) \\ &+ e_1^T A_1^{i-1} B_1 \mathbf{x}(t-d) + e_1^T A_1^{i-2} B_1 \mathbf{x}(t-d+1) \\ &+ \dots + e_1^T B_1 \mathbf{x}(t-d+i-1) \\ &+ e_1^T A_1^{i-1} F_1 \mathbf{u}(t) + \dots + e_1^T F_1 \mathbf{u}(t+i-1) \\ &+ v_1(t+i), i = 0, 1, \dots, n_1-1, \end{aligned} \quad (3)$$

$$\begin{aligned} y_1(t+n_1) &= e_1^T A_1^{n_1} \mathbf{x}_1(t) + e_1^T A_1^{n_1-1} A_{12} \mathbf{x}_2(t) \\ &+ e_1^T A_1^{n_1-2} A_{12} \mathbf{x}_2(t+1) + \dots \\ &+ e_1^T A_{12} \mathbf{x}_2(t+n_1-1) + e_1^T A_1^{n_1-1} B_1 \mathbf{x}(t-d) \\ &+ e_1^T A_1^{n_1-2} B_1 \mathbf{x}(t-d+1) + \dots \\ &+ e_1^T B_1 \mathbf{x}(t-d+n_1-1) + e_1^T A_1^{n_1-1} F_1 \mathbf{u}(t) \\ &+ e_1^T A_1^{n_1-2} F_1 \mathbf{u}(t+1) + \dots \\ &+ e_1^T F_1 \mathbf{u}(t+n_1-1) + v_1(t+n_1). \end{aligned} \quad (4)$$

Since the model (1)–(2) is in the observable canonical form, the decomposed subsystem is still the observable canonical model. According to the special structure of the matrix A_1 and e_1^T , it is known that the observability matrix T is an identity matrix:

$$T = \begin{bmatrix} e_1^T \\ e_1^T A_1 \\ \vdots \\ e_1^T A_1^{n_1-1} \end{bmatrix} = I_{n_1}. \quad (5)$$

Post-multiplying both sides of the above equation by the matrix A_{12} yields:

$$\begin{bmatrix} e_1^T A_{12} \\ e_1^T A_1 A_{12} \\ \vdots \\ e_1^T A_1^{n_1-1} A_{12} \end{bmatrix} = A_{12}. \quad (6)$$

Observing the structure of matrix A_{12} , we get

$$e_1^T A_1^{k-1} A_{12} = 0, k = 1, 2, \dots, n_1-1.$$

Observing Equation (6), $e_1^T A_1^{k-1} A_{12}$ is the last row of matrix A_{12} , which is

$$e_1^T A_1^{n_1-1} A_{12} = [a_{12}(1), a_{12}(2), \dots, a_{12}(n_{12}), 0, \dots, 0].$$

Define

$$Y_i(t+n_1) := [y_i(t), y_i(t+1), \dots, y_i(t+n_i-1)]^T \in \mathbb{R}^{n_i},$$

$$U_i(t+n_1) := [u^T(t), u^T(t+1), \dots, u^T(t+n_i-1)]^T \in \mathbb{R}^{2n_i},$$

$$X(t-d+n_i) := [\mathbf{x}^T(t-d), \mathbf{x}^T(t-d+1), \dots,$$

$$\mathbf{x}^T(t-d+n_i-1)]^T \in \mathbb{R}^{m_i},$$

$$V_i(t+n_i) := [v_i(t), v_i(t+1), \dots, v_i(t+n_i-1)]^T \in \mathbb{R}^{n_i},$$

$$M_i := \begin{bmatrix} 0 & \dots & 0 & 0 \\ e_i^T B_i & \dots & 0 & 0 \\ e_i^T A_i B_i & & \vdots & \vdots \\ \vdots & & 0 & 0 \\ e_i^T A_i^{n_i-2} B_i & \dots & e_i^T B_i & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times (m_i)},$$

$$Q_i := \begin{bmatrix} 0 & \dots & 0 & 0 \\ e_i^T F_i & \dots & 0 & 0 \\ e_i^T A_i F_i & & \vdots & \vdots \\ \vdots & & 0 & 0 \\ e_i^T A_i^{n_i-2} F_i & \dots & e_i^T F_i & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times (2n_i)}.$$

From Equations (3) and (4), we have

$$\begin{aligned} Y_1(t+n_1) &= T \mathbf{x}_1(t) + M_1 X(t-d+n_1) + Q_1 U_1(t+n_1) \\ &+ V_1(t+n_1). \end{aligned}$$

Combining the observable matrix in (5) and the above equation gives

$$\begin{aligned} \mathbf{x}_1(t) &= Y_1(t+n_1) - M_1 X(t-d+n_1) - Q_1 U_1(t+n_1) \\ &- V_1(t+n_1). \end{aligned} \quad (7)$$

For the second subsystem, we have

$$\begin{aligned} \mathbf{x}_2(t) = & \mathbf{Y}_2(t + n_2) - \mathbf{M}_2\mathbf{X}(t - d + n_2) - \mathbf{Q}_2\mathbf{U}_2(t + n_2) \\ & - \mathbf{V}_2(t + n_2). \end{aligned} \quad (8)$$

For $n_{12} \leq n_1$ and $n_{12} \leq \min\{n_1, n_2\}$, the number of non-zero elements of $\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{Q}_2 \in \mathbb{R}^{1 \times (2n_2)}$ is $2n_{12}$, now construct a vector associated with it:

$$\mathbf{l}_1 := \begin{cases} [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{Q}_2, 0, \dots, 0] \in \mathbb{R}^{1 \times (2n_1)}, n_1 \geq n_2 \\ [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{Q}_2] \begin{bmatrix} \mathbf{I}_{2n_1} \\ 0 \end{bmatrix} \in \mathbb{R}^{1 \times (2n_1)}, n_1 \geq n_2, \end{cases}$$

similarly,

$$\mathbf{h}_1 := \begin{cases} [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{M}_2, 0, \dots, 0] \in \mathbb{R}^{1 \times (2n_1)}, n_1 \geq n_2 \\ [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{M}_2] \begin{bmatrix} \mathbf{I}_{2n_1} \\ 0 \end{bmatrix} \in \mathbb{R}^{1 \times (2n_1)}, n_1 \geq n_2. \end{cases}$$

In order to get the parameter estimates, define the information set at time t by $\boldsymbol{\varphi}_1(t)$ and the parameter vector $\boldsymbol{\theta}_1$ as

$$\begin{aligned} \boldsymbol{\varphi}_1(t + n_1) &:= [\boldsymbol{\varphi}_{11}^T(t + n_1), \mathbf{U}_1^T(t + n_1), \mathbf{X}^T(t - d + n_1)]^T \\ &\in \mathbb{R}^{n+2n_1+nn_1}, \\ \boldsymbol{\varphi}_{11}(t + n_1) &:= \begin{bmatrix} \mathbf{Y}_1(t + n_1) - \mathbf{V}_1(t + n_1) \\ \mathbf{Y}_2(t + n_2) - \mathbf{V}_2(t + n_2) \end{bmatrix} \in \mathbb{R}^n, \\ \boldsymbol{\theta}_1 &:= [\boldsymbol{\theta}_{11}^T, \boldsymbol{\theta}_{12}^T, \boldsymbol{\theta}_{13}^T]^T \in \mathbb{R}^{n+2n_1+nn_1}, \\ \boldsymbol{\theta}_{11} &:= [\mathbf{e}_1^T \mathbf{A}_1^{n_1}, \mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12}]^T \in \mathbb{R}^n, \end{aligned} \quad (9)$$

$$\begin{aligned} \boldsymbol{\theta}_{12} &:= [-\mathbf{e}_1^T \mathbf{A}_1^{n_1} \mathbf{Q}_1 - \mathbf{l}_1 + [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{F}_1, \\ &\quad \mathbf{e}_1^T \mathbf{A}_1^{n_1-2} \mathbf{F}_1, \dots, \mathbf{e}_1^T \mathbf{F}_1]]^T \in \mathbb{R}^{2n_1}, \end{aligned} \quad (10)$$

$$\begin{aligned} \boldsymbol{\theta}_{13} &:= [-\mathbf{e}_1^T \mathbf{A}_1^{n_1} \mathbf{M}_1 - \mathbf{h}_1 + [\mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{B}_1, \\ &\quad \mathbf{e}_1^T \mathbf{A}_1^{n_1-2} \mathbf{B}_1, \dots, \mathbf{e}_1^T \mathbf{B}_1]]^T \in \mathbb{R}^{nn_1}. \end{aligned} \quad (11)$$

Substituting the state variable in (7)–(8) into (4) gives

$$\begin{aligned} y_1(t + n_1) = & \mathbf{e}_1^T \mathbf{A}_1^{n_1} \mathbf{x}_1(t) + \mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{A}_{12} \mathbf{x}_2(t) \\ & + \mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{B}_1 \mathbf{x}(t - d) + \dots \\ & + \mathbf{e}_1^T \mathbf{B}_1 \mathbf{x}(t - d + n_1 - 1) + \mathbf{e}_1^T \mathbf{A}_1^{n_1-1} \mathbf{F}_1 \mathbf{u}(t) \end{aligned}$$

$$\begin{aligned} & + \mathbf{e}_1^T \mathbf{A}_1^{n_1-2} \mathbf{F}_1 \mathbf{u}(t + 1) + \dots \\ & + \mathbf{e}_1^T \mathbf{F}_1 \mathbf{u}(t + n_1 - 1) + v_1(t + n_1) \\ = & \boldsymbol{\varphi}_{11}^T(t + n_1) \boldsymbol{\theta}_{11} + \mathbf{U}_1^T(t + n_1) \boldsymbol{\theta}_{12} \\ & + \mathbf{X}^T(t - d + n_1) \boldsymbol{\theta}_{13} + v_1(t + n_1) \\ = & [\boldsymbol{\varphi}_{11}^T(t + n_1), \mathbf{U}_1^T(t + n_1), \mathbf{X}^T(t - d + n_1)] \\ & \times \begin{bmatrix} \boldsymbol{\theta}_{11} \\ \boldsymbol{\theta}_{12} \\ \boldsymbol{\theta}_{13} \end{bmatrix} + v_1(t + n_1) \\ = & \boldsymbol{\varphi}_1^T(t + n_1) \boldsymbol{\theta}_1 + v_1(t + n_1). \end{aligned} \quad (12)$$

Replacing t in (12) with $t - n_1$ can be simplified as the following regression model,

$$y_1(t) = \boldsymbol{\varphi}_1^T(t) \boldsymbol{\theta}_1 + v_1(t). \quad (13)$$

The proposed parameter estimation algorithms are based on this identification model. Many identification methods are derived based on the identification models of the systems [38–43] and can be used to estimate the parameters of other linear systems and nonlinear systems [44–50] and can be applied to other fields [51–57] such as chemical process control systems.

Remark 1. The above equation is the identification model of the two-input two-output state space system with time delay. For research convenience, assume t is the current moment, $\{\mathbf{u}(t), \mathbf{y}(t) : t = 0, 1, 2, \dots\}$ is the measurable input-output information, $\mathbf{y}(t)$ and $\boldsymbol{\varphi}(t)$ are the current information, $\{\mathbf{y}(t - i), \boldsymbol{\varphi}(t - i) : i = 1, 2, \dots, p - 1\}$ are the past information.

3 | THE PARAMETER ESTIMATION ALGORITHM

According to the least squares theory for Equation (13), minimizing the criterion function,

$$J_1(\boldsymbol{\theta}_1) := \sum_{k=1}^t [y_1(k) - \boldsymbol{\varphi}_1^T(k) \boldsymbol{\theta}_1]^2.$$

We get the following recursive least square (RLS) algorithm to estimate the parameter vector $\boldsymbol{\theta}_1$:

$$\hat{\boldsymbol{\theta}}_1(t) = \hat{\boldsymbol{\theta}}_1(t - 1) + \mathbf{L}_1(t)[y_1(t) - \boldsymbol{\varphi}_1^T(t) \hat{\boldsymbol{\theta}}_1(t - 1)], \quad (14)$$

$$\mathbf{L}_1(t) = \mathbf{P}_1(t) \boldsymbol{\varphi}_1(t) = \frac{\mathbf{P}_1(t - 1) \boldsymbol{\varphi}_1(t)}{1 + \boldsymbol{\varphi}_1^T(t) \mathbf{P}_1(t - 1) \boldsymbol{\varphi}_1(t)}, \quad (15)$$

$$\mathbf{P}_1(t) = \mathbf{P}_1(t-1) - \frac{\mathbf{P}_1(t-1)\boldsymbol{\varphi}_1(t)\boldsymbol{\varphi}_1^T(t)\mathbf{P}_1(t-1)}{1 + \boldsymbol{\varphi}_1^T(t)\mathbf{P}_1(t-1)\boldsymbol{\varphi}_1(t)}, \quad (16)$$

where $\mathbf{L}_1(t) \in \mathbb{R}^{n+2n_1+nm_1}$ is the gain vector,

$\mathbf{P}_1(t) \in \mathbb{R}^{(n+2n_1+nm_1) \times (n+2n_1+nm_1)}$ is the covariance matrix.

Remark 2. For the information vector $\boldsymbol{\varphi}_1(t)$ in (14)–(16) contains the unknown noise item $v_1(t-i)$ and the state vector $\mathbf{x}(t-d-i)$, the above algorithm cannot be realized, which is the difficulty in identification. This section adopts the basic idea of replacing the unknown noise item $v_1(t-i)$ and the state vector $\mathbf{x}(t-d-i)$ in $\boldsymbol{\varphi}_1(t)$ with the estimated residual $\hat{v}_1(t-i)$ and the estimated state vector $\hat{\mathbf{x}}(t-d-i)$.

Use the estimates $\hat{v}_i(t)$ and $\hat{\mathbf{x}}(t)$ of $v_i(t)$ and $\mathbf{x}(t)$ to construct the estimates $\hat{\mathbf{V}}_i(t)$ and $\hat{\mathbf{X}}(t)$ of $\mathbf{V}_i(t)$ and $\mathbf{X}(t)$:

$$\hat{\boldsymbol{\varphi}}_1(t) := [\hat{\boldsymbol{\varphi}}_{11}^T(t), \mathbf{U}_1^T(t), \hat{\mathbf{X}}^T(t-d)]^T \in \mathbb{R}^{n+2n_1+nm_1},$$

$$\hat{\boldsymbol{\varphi}}_{11}(t) := \begin{bmatrix} \mathbf{Y}_1(t) - \hat{\mathbf{V}}_1(t) \\ \mathbf{Y}_2(t) - \hat{\mathbf{V}}_2(t) \end{bmatrix} \in \mathbb{R}^n,$$

$$\hat{\mathbf{X}}(t-d) := [\hat{\mathbf{x}}^T(t-n_1-d), \dots, \hat{\mathbf{x}}^T(t-d-1)]^T \in \mathbb{R}^{nm_1},$$

$$\hat{\mathbf{V}}_1(t) := [\hat{v}_1^T(t-n_1), \hat{v}_1^T(t-n_1+1), \dots, \hat{v}_1^T(t-1)]^T \in \mathbb{R}^{n_1},$$

$$\hat{\mathbf{V}}_2(t) := [\hat{v}_2^T(t-n_2), \hat{v}_2^T(t-n_2+1), \dots, \hat{v}_2^T(t-1)]^T \in \mathbb{R}^{n_2}.$$

Define $\hat{\boldsymbol{\theta}}_1(t) := [\hat{\boldsymbol{\theta}}_{11}^T(t), \hat{\boldsymbol{\theta}}_{12}^T(t), \hat{\boldsymbol{\theta}}_{13}^T(t)]^T$ is the estimate of $\boldsymbol{\theta}_1 = [\boldsymbol{\theta}_{11}^T, \boldsymbol{\theta}_{12}^T, \boldsymbol{\theta}_{13}^T]^T$ at time t . According to Equation (13), the estimate of $v_1(t)$ is calculated as $\hat{v}_1(t) = y_1(t) - \hat{\boldsymbol{\varphi}}_1^T(t)\hat{\boldsymbol{\theta}}_1(t)$. Thus, replacing the unknown variable $\boldsymbol{\varphi}_1(t)$ on the right-hand sides of algorithm (14)–(16) with its corresponding estimate $\hat{\boldsymbol{\varphi}}_1(t)$, replacing the unknown $\boldsymbol{\theta}_1$ with its estimate $\hat{\boldsymbol{\theta}}_1(t-1)$ at the previous time $t-1$, we obtain the following parameter estimation based recursive least squares algorithm to calculate $\boldsymbol{\theta}_1$:

$$\hat{\boldsymbol{\theta}}_1(t) = \hat{\boldsymbol{\theta}}_1(t-1) + \mathbf{L}_1(t)[y_1(t) - \hat{\boldsymbol{\varphi}}_1^T(t)\hat{\boldsymbol{\theta}}_1(t-1)], \quad (17)$$

$$\mathbf{L}_1(t) = \mathbf{P}_1(t)\hat{\boldsymbol{\varphi}}_1(t) = \frac{\mathbf{P}_1(t-1)\hat{\boldsymbol{\varphi}}_1(t)}{1 + \hat{\boldsymbol{\varphi}}_1^T(t)\mathbf{P}_1(t-1)\hat{\boldsymbol{\varphi}}_1(t)}, \quad (18)$$

$$\mathbf{P}_1(t) = [\mathbf{I} - \mathbf{L}_1(t)\hat{\boldsymbol{\varphi}}_1^T(t)]\mathbf{P}_1(t-1), \quad \mathbf{P}_1(0) = \mathbf{P}_0\mathbf{I}, \quad (19)$$

$$\hat{v}_1(t) = y_1(t) - \hat{\boldsymbol{\varphi}}_1^T(t)\hat{\boldsymbol{\theta}}_1(t), \quad (20)$$

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_1(t) = & [y_1(t-n_1) - \hat{v}_1(t-n_1), y_1(t-n_1+1) \\ & - \hat{v}_1(t-n_1+1), \dots, y_1(t-1) - \hat{v}_1(t-1), \end{aligned}$$

$$\begin{aligned} & y_2(t-n_2) - \hat{v}_2(t-n_2), y_2(t-n_2+1) \\ & - \hat{v}_2(t-n_2+1), \dots, y_2(t-1) - \hat{v}_2(t-1), \\ & \hat{\mathbf{x}}(t-n_1-d), \hat{\mathbf{x}}(t-n_1-d+1), \dots, \hat{\mathbf{x}}(t-d-1), \\ & \mathbf{u}^T(t-n_1), \dots, \mathbf{u}^T(t-1)]^T. \end{aligned} \quad (21)$$

Similar to the derivation process of the parameter vector $\boldsymbol{\theta}_1$, the second subsystem is obtained as follows.

Define $\mathbf{L}_2$ and $\mathbf{h}_2$ as

$$\begin{aligned} \mathbf{L}_2 &:= \begin{cases} [e_2^T \mathbf{A}_2^{n_2-1} \mathbf{A}_{21} \mathbf{Q}_1, 0, \dots, 0] \in \mathbb{R}^{1 \times (2n_2)}, n_2 \geq n_1 \\ [e_2^T \mathbf{A}_2^{n_2-1} \mathbf{A}_{21} \mathbf{Q}_1] \begin{bmatrix} \mathbf{I}_{2n_1} \\ 0 \end{bmatrix} \in \mathbb{R}^{1 \times (2n_2)}, n_2 < n_1, \end{cases} \\ \mathbf{h}_2 &:= \begin{cases} [e_2^T \mathbf{A}_2^{n_1-1} \mathbf{A}_{21} \mathbf{M}_1, 0, \dots, 0] \in \mathbb{R}^{1 \times (nm_2)}, n_2 \geq n_1 \\ [e_2^T \mathbf{A}_2^{n_1-1} \mathbf{A}_{21} \mathbf{M}_1] \begin{bmatrix} \mathbf{I}_{2n_1} \\ 0 \end{bmatrix} \in \mathbb{R}^{1 \times (nm_2)}, n_2 < n_1. \end{cases} \end{aligned}$$

When computing one parameter vector, others are replaced with their estimates, then we get the parameter estimation based recursive least squares algorithm as follows:

$$\hat{\boldsymbol{\theta}}_2(t) = \hat{\boldsymbol{\theta}}_2(t-1) + \mathbf{L}_2(t)[y_2(t) - \hat{\boldsymbol{\varphi}}_2^T(t)\hat{\boldsymbol{\theta}}_2(t-1)], \quad (22)$$

$$\mathbf{L}_2(t) = \mathbf{P}_2(t)\hat{\boldsymbol{\varphi}}_2(t) = \frac{\mathbf{P}_2(t-1)\hat{\boldsymbol{\varphi}}_2(t)}{1 + \hat{\boldsymbol{\varphi}}_2^T(t)\mathbf{P}_2(t-1)\hat{\boldsymbol{\varphi}}_2(t)}, \quad (23)$$

$$\mathbf{P}_2(t) = [\mathbf{I} - \mathbf{L}_2(t)\hat{\boldsymbol{\varphi}}_2^T(t)]\mathbf{P}_2(t-1), \quad \mathbf{P}_2(0) = \mathbf{P}_0\mathbf{I}, \quad (24)$$

$$\hat{v}_2(t) = y_2(t) - \hat{\boldsymbol{\varphi}}_2^T(t)\hat{\boldsymbol{\theta}}_2(t), \quad (25)$$

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_2(t) = & [y_1(t-n_1) - \hat{v}_1(t-n_1), y_1(t-n_1+1) \\ & - \hat{v}_1(t-n_1+1), \dots, y_1(t-1) - \hat{v}_1(t-1), \\ & y_2(t-n_2) - \hat{v}_2(t-n_2), y_2(t-n_2+1) \\ & - \hat{v}_2(t-n_2+1), \dots, y_2(t-1) - \hat{v}_2(t-1), \\ & \hat{\mathbf{x}}(t-n_2-d), \hat{\mathbf{x}}(t-n_2-d+1), \dots, \\ & \hat{\mathbf{x}}(t-d-1), \mathbf{u}^T(t-n_2), \dots, \mathbf{u}^T(t-1)]^T, \end{aligned} \quad (26)$$

where $\hat{\boldsymbol{\theta}}_2(t) := [\hat{\boldsymbol{\theta}}_{21}^T(t), \hat{\boldsymbol{\theta}}_{22}^T(t), \hat{\boldsymbol{\theta}}_{23}^T(t)]^T$ is the estimate of $\boldsymbol{\theta}_2 = [\boldsymbol{\theta}_{21}^T, \boldsymbol{\theta}_{22}^T, \boldsymbol{\theta}_{23}^T]^T$ at time t , the gain matrix $\mathbf{L}_2(t) \in \mathbb{R}^{n+2n_2+nm_2}$, the covariance matrix $\mathbf{P}_2(t) \in \mathbb{R}^{(n+2n_2+nm_2) \times (n+2n_2+nm_2)}$.

Equations (17)–(21) and (22)–(26) form the recursive least squares algorithm whose initial values $\hat{\theta}_1(0)$ and $\hat{\theta}_2(0)$ are taken as zero vectors of appropriate sizes, $\hat{v}_i(j)$, $\hat{x}(j)$, $\hat{\phi}_1(j)$, $\hat{\phi}_2(j)$, $u(j)$ and $y(j)$ as zero vectors or zero matrices of appropriate sizes for $j \leq 0$, and $P_1(0) = p_0 I$, $P_2(0) = p_0 I$, $p_0 = 10^6$, I is an identity matrix of appropriate dimensions.

Remark 3. Since we consider a multivariable system, the coupling of the system needs to be analyzed during the decomposition, which is to realize the decoupling of the system: the multivariable system that makes the input and output are correlated to each other realizes that each output is only controlled by the corresponding input.

Theorem 1. For the system in (1) and (2), the identification model in (13) and the least squares algorithm in (17)–(26), suppose that $\{v(t)\}$ is a white noise sequence with zero mean and variance σ^2 , that is, $\mathbb{E}[v(t)] = 0$, $\mathbb{E}[v^2(t)] = \sigma^2$, $\mathbb{E}[v(t)v(s)] = 0$, $s \neq t$, the input $u(t)$ is deterministic, the following least squares parameter estimate $\hat{\theta}_i$ is an unbiased estimate of θ_i ,

$$\begin{aligned} \mathbb{E}[\hat{\theta}_i] &= \theta_i + \mathbb{E} \left[\left[\sum_{k=1}^t \varphi_i(k) \varphi_i^T(k) \right]^{-1} \sum_{k=1}^t \varphi_i(k) v(k) \right] \\ &= \theta_i + \left[\sum_{k=1}^t \varphi_i(k) \varphi_i^T(k) \right]^{-1} \sum_{k=1}^t \varphi_i(k) \mathbb{E}[v(k)] \\ &= \theta_i. \end{aligned}$$

4 | THE STATE ESTIMATION ALGORITHM

The relationship between the parameter vector θ_1 and the matrices/vector A_1 , A_{12} , B_1 , e_1^T has been established; post-multiplying the observable matrix in (5) on both sides by B_1 gives

$$\begin{bmatrix} e_1^T B_1 \\ e_1^T A_1 B_1 \\ \vdots \\ e_1^T A_1^{n_1-1} B_1 \end{bmatrix} = B_1. \quad (27)$$

From Equation (27) and the definition of B_1 , we have

$$e_1^T A_1^{k-1} B_1 = b_{1k}, \quad k = 1, 2, \dots, n_1.$$

Using the above equation, the matrix M_1 and the similar matrix M_2 are simplified as

$$M_i = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ b_{i1} & 0 & \cdots & 0 & 0 \\ b_{i2} & b_{i1} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ b_{i,n_i-1} & b_{i,n_i-2} & \cdots & b_{i1} & 0 \end{bmatrix}, \quad i = 1, 2. \quad (28)$$

Observing the structure of A_{12} , from the last line of the formula (6) gives

$$e_1^T A_1^{n_1-1} A_{12} = [a_{12}(1), a_{12}(2), \dots, a_{12}(n_{12}), 0, \dots, 0], \quad (29)$$

$$e_1^T A_1^{n_1} = [a_1(1), a_1(2), \dots, a_1(n_1), 0, \dots, 0]. \quad (30)$$

Post-multiplying Equation (29) on both sides with the matrix M_2 , and post-multiplying Equation (30) on both sides with the matrix M_1 , we have

$$\begin{aligned} e_1^T A_1^{n_1-1} A_{12} M_2 &= [a_{12}(2)b_{21} + a_{12}(3)b_{22} + \cdots \\ &\quad + a_{12}(n_{12})b_{2,n_{12}-1}, \\ &\quad a_{12}(3)b_{21} + a_{12}(4)b_{22} + \cdots \\ &\quad + a_{12}(n_{12})b_{2,n_{12}-2}, \dots, 0], \end{aligned} \quad (31)$$

$$\begin{aligned} e_1^T A_1^{n_1} M_1 &= [a_1(2)b_{11} + a_1(3)b_{12} + \cdots \\ &\quad + a_1(n_1)b_{1,n_1-1}, \\ &\quad a_1(3)b_{11} + a_1(4)b_{12} + \cdots \\ &\quad + a_1(n_1)b_{1,n_1-2}, \dots, a_1(n_1)b_{11}, 0]. \end{aligned} \quad (32)$$

From Equation (31), we get that the number of non-zero elements of $e_1^T A_1^{n_1-1} A_{12}$ is $2n_{12}$, according to the above equations, the parameters of θ_{11} in (9), θ_{12} in (10) and θ_{13} in (11) are expressed as

$$\begin{aligned} \theta_{11} &= [e_1^T A_1^{n_1}, e_1^T A_1^{n_1-1} A_{12}]^T \\ &= [a_1(1), a_1(2), \dots, a_1(n_1), a_{12}(1), a_{12}(2), \dots, \\ &\quad a_{12}(n_{12}), 0, \dots, 0]^T, \end{aligned} \quad (33)$$

$$\begin{aligned} \theta_{12} &= [-e_1^T A_1^{n_1} Q_1 - l_1 + [e_1^T A_1^{n_1-1} F_1, e_1^T A_1^{n_1-2} F_1, \dots, \\ &\quad e_1^T F_1]] \\ &= [-a_1(2)f_{11} - a_1(3)f_{12} - \cdots - a_1(n_1)f_{1,n_1-1} + f_{1,n_1} \\ &\quad - a_{12}(2)f_{21} - a_{12}(3)f_{22} - \cdots - a_{12}(n_{12})f_{2,n_{12}-1}, \\ &\quad - a_1(3)f_{11} - a_1(4)f_{12} - \cdots - a_1(n_1)f_{1,n_1-2} \end{aligned}$$

$$\begin{aligned}
& +f_{1,(n_1-1)} - a_{12}(3)f_{21} - a_{12}(4)f_{22} - \cdots \\
& -a_{12}(n_{12})f_{2,(n_{12}-2)}, \dots, -a_1(n_1)f_{11} + f_{12}, f_{11}]^T, \quad (34) \\
\theta_{13} = & [-e_1^T A_1^{n_1} M_1 - h_1 + [e_1^T A_1^{n_1-1} B_1, e_1^T A_1^{n_1-2} B_1, \dots, \\
& e_1^T B_1]]^T \\
= & [-a_1(2)b_{11} - a_1(3)b_{12} - \cdots - a_1(n_1)b_{1,(n_1-1)} + b_{1,n_1} \\
& -a_{12}(2)b_{21} - a_{12}(3)b_{22} - \cdots - a_{12}(n_{12})b_{2,(n_{12}-1)}, \\
& -a_1(3)b_{11} - a_1(4)b_{12} - \cdots - a_1(n_1)b_{1,(n_1-2)} \\
& +b_{1,(n_1-1)} - a_{12}(3)b_{21} - a_{12}(4)b_{22} - \cdots \\
& -a_{12}(n_{12})b_{2,(n_{12}-2)}, \dots, -a_1(n_1)b_{11} + b_{12}, b_{11}]^T. \quad (35)
\end{aligned}$$

Let $\theta_{12}^T =: [g_1, g_2, \dots, g_{n_1}] \in \mathbb{R}^{1 \times (2n_1)}$. Equation (34) is converted into the following form:

$$\begin{aligned}
& \left[\begin{array}{ccc|ccc} -a_1(2) & \cdots & 1 & -a_{12}(2) & \cdots & -a_{12}(n_{12}) & 0 \\ -a_1(3) & \cdots & & -a_{12}(3) & \cdots & 0 & \\ \vdots & \ddots & & \vdots & & & \\ \vdots & \ddots & & -a_{12}(n_{12}) & 0 & & \\ -a_1(n_1) & & & 0 & & & \\ 1 & & & 0 & & & \end{array} \right] F \\
& = \left[\begin{array}{c} g_1 \\ g_2 \\ \vdots \\ g_{n_1-1} \\ g_{n_1} \end{array} \right]. \quad (36)
\end{aligned}$$

Define $\theta_{13}^T =: [r_1, r_2, \dots, r_{n_1}] \in \mathbb{R}^{1 \times (n_{n_1})}$. By arranging Equation (35), the following matrix equation is obtained.

$$\begin{aligned}
& \left[\begin{array}{ccc|ccc} -a_1(2) & \cdots & 1 & -a_{12}(2) & \cdots & -a_{12}(n_{12}) & 0 \\ -a_1(3) & \cdots & & -a_{12}(3) & \cdots & 0 & \\ \vdots & \ddots & & \vdots & & & \\ \vdots & \ddots & & -a_{12}(n_{12}) & 0 & & \\ -a_1(n_1) & & & 0 & & & \\ 1 & & & 0 & & & \end{array} \right] B \\
& = \left[\begin{array}{c} r_1 \\ r_2 \\ \vdots \\ r_{n_1-1} \\ r_{n_1} \end{array} \right]. \quad (37)
\end{aligned}$$

Similarly, the parameters θ_{21} , θ_{22} and θ_{23} are expressed as

$$\begin{aligned}
\theta_{21} &= [e_2^T A_2^{n_2-1} A_{21}, e_2^T A_2^{n_2}]^T \\
&= [a_{21}(1), a_{21}(2), \dots, a_{21}(n_{21}), 0, \dots, 0, \\
& \quad a_2(1), a_2(2), \dots, a_2(n_2)]^T, \\
\theta_{22} &= [-e_2^T A_2^{n_2} Q_2 - l_2 + [e_2^T A_2^{n_2-1} F_2, e_2^T A_2^{n_2-2} F_2, \dots, \\
& \quad e_2^T F_2]]^T \\
&= [-a_2(2)f_{21} - a_2(3)f_{22} - \cdots - a_2(n_2)f_{2,(n_2-1)} + f_{2,n_2} \\
& \quad -a_{22}(2)f_{11} - a_{21}(3)f_{12} - \cdots - a_{21}(n_{21})f_{1,(n_{21}-1)}, \\
& \quad -a_2(3)f_{21} - a_2(4)f_{22} - \cdots - a_2(n_2)f_{2,(n_2-2)} \\
& \quad + f_{2,(n_2-1)} - a_{21}(3)f_{11} - a_{21}(4)f_{12} - \cdots \\
& \quad -a_{21}(n_{21})f_{1,(n_{21}-2)}, \dots, -a_2(n_2)f_{21} + f_{22}, f_{21}]^T, \quad (38)
\end{aligned}$$

$$\begin{aligned}
\theta_{23} &= [-e_2^T A_2^{n_2} M_2 - h_2 + [e_2^T A_2^{n_2-1} B_2, e_2^T A_2^{n_2-2} B_2, \dots, \\
& \quad e_2^T B_2]]^T \\
&= [-a_2(2)b_{21} - a_2(3)b_{22} - \cdots - a_2(n_2)b_{2,(n_2-1)} + b_{2,n_2} \\
& \quad -a_{21}(2)b_{11} - a_{21}(3)b_{12} - \cdots - a_{21}(n_{21})b_{1,(n_{21}-1)}, \\
& \quad -a_2(3)b_{21} - a_2(4)b_{22} - \cdots - a_2(n_2)b_{2,(n_2-2)} \\
& \quad +b_{2,(n_2-1)} - a_{21}(3)b_{11} - a_{21}(4)b_{12} - \cdots \\
& \quad -a_{21}(n_{21})b_{1,(n_{21}-2)}, \dots, -a_2(n_2)b_{21} + b_{22}, b_{21}]^T. \quad (39)
\end{aligned}$$

Let $\theta_{22}^T =: [h_1, h_2, \dots, h_{n_2}] \in \mathbb{R}^{1 \times (2n_2)}$. Using Equation (38), the following equation is established.

$$\begin{aligned}
& \left[\begin{array}{ccc|ccc} -a_{21}(2) & \cdots & -a_{21}(n_{21}) & 0 & -a_2(2) & \cdots & 1 \\ -a_{21}(3) & \cdots & 0 & & -a_2(3) & \cdots & \\ \vdots & & & & \vdots & & \\ -a_{21}(n_{21}) & 0 & & & \cdot & & \\ 0 & & & & -a_2(n_2) & & \\ 0 & & & & 1 & & \end{array} \right] F \\
& = \left[\begin{array}{c} h_1 \\ h_2 \\ \vdots \\ h_{n_2-1} \\ h_{n_2} \end{array} \right]. \quad (40)
\end{aligned}$$

Let $\theta_{23}^T = [s_1, s_2, \dots, s_{n_2}] \in \mathbb{R}^{1 \times (n_{n_2})}$. Combing with Equation (39), we have

$$\begin{bmatrix} -a_{21}(2) & \cdots & -a_{21}(n_{21}) & 0 & -a_2(2) & \cdots & 1 \\ -a_{21}(3) & \cdots & 0 & & -a_2(3) & \cdots & \\ \vdots & & & & \vdots & & \\ -a_{21}(n_{21}) & 0 & & & \cdot & & \\ 0 & & & & -a_2(n_2) & & \\ 0 & & & & 1 & & \end{bmatrix} B = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{n_2-1} \\ s_{n_2} \end{bmatrix}. \quad (41)$$

Combing Equations (36) and (40) gives

$$\begin{bmatrix} -a_1(2) & \cdots & 1 & -a_{12}(2) & \cdots & 0 \\ -a_1(3) & \cdots & & -a_{12}(3) & \cdots & \\ \cdot & & & \vdots & & \\ \cdot & & & -a_{12}(n_{12}) & 0 & \\ -a_1(n_1) & & & 0 & & \\ 1 & & & 0 & & \\ -a_{21}(2) & \cdots & 0 & -a_2(2) & \cdots & 1 \\ -a_{21}(3) & \cdots & & -a_2(3) & \cdots & \\ \vdots & & & \cdot & & \\ -a_{21}(n_{21}) & 0 & & \cdot & & \\ 0 & & & -a_2(n_2) & & \\ 0 & & & 1 & & \end{bmatrix} F = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{n_1} \\ h_1 \\ h_2 \\ \vdots \\ h_{n_2} \end{bmatrix}. \quad (42)$$

Similarly, combining Equations (37) and (41), we have

$$\begin{bmatrix} -a_1(2) & \cdots & 1 & -a_{12}(2) & \cdots & 0 \\ -a_1(3) & \cdots & & -a_{12}(3) & \cdots & \\ \cdot & & & \cdot & & \\ \cdot & & & -a_{12}(n_{12}) & 0 & \\ -a_1(n_1) & & & 0 & & \\ 1 & & & 0 & & \\ -a_{21}(2) & \cdots & 0 & -a_2(2) & \cdots & 1 \\ -a_{21}(3) & \cdots & & -a_2(3) & \cdots & \\ \cdot & & & \cdot & & \\ -a_{21}(n_{21}) & 0 & & \cdot & & \\ 0 & & & -a_2(n_2) & & \\ 0 & & & 1 & & \end{bmatrix} B = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{n_1} \\ s_1 \\ s_2 \\ \vdots \\ s_{n_2} \end{bmatrix}. \quad (43)$$

Define the large matrix:

$$W := \left[\begin{array}{ccc|ccc} -a_1(2) & \cdots & 1 & -a_{12}(2) & \cdots & 0 \\ -a_1(3) & \cdots & & -a_{12}(3) & \cdots & \\ \cdot & & & \cdot & & \\ \cdot & & & -a_{12}(n_{12}) & 0 & \\ -a_1(n_1) & & & 0 & & \\ 1 & & & 0 & & \\ \hline -a_{21}(2) & \cdots & 0 & -a_2(2) & \cdots & 1 \\ -a_{21}(3) & \cdots & & -a_2(3) & \cdots & \\ \cdot & & & \cdot & & \\ -a_{21}(n_{21}) & 0 & & \cdot & & \\ 0 & & & -a_2(n_2) & & \end{array} \right],$$

$$K := [g_1^T, g_2^T, \dots, g_{n_1}^T, h_1^T, h_2^T, \dots, h_{n_2}^T]^T,$$

$$J := [r_1^T, r_2^T, \dots, r_{n_1}^T, s_1^T, s_2^T, \dots, s_{n_2}^T]^T.$$

Equations (37) and (41) can be expressed as $WF = K$, $WB = J$. Applying the estimates $\hat{a}_i$, $\hat{F}(t)$ and $\hat{B}(t)$ to set up: $\hat{F}(t) = \hat{W}^{-1}(t)\hat{K}(t)$, $\hat{B}(t) = \hat{W}^{-1}(t)\hat{J}(t)$. Replacing $t - n$ in (7) and (8) to t yields

$$x_i(t - n_i) = Y_i(t) - M_i x(t - d) - Q_i U_i(t) - V_i(t), \quad i = 1, 2.$$

Using $\hat{M}_1$, $\hat{M}_2$, $\hat{Q}_1$, $\hat{Q}_2$, $\hat{V}_1$ and $\hat{V}_2$ to replace M_1 , M_2 , Q_1 , Q_2 , V_1 and V_2 in the above equations, we get the estimates of state vectors:

$$\hat{x}_i(t - n_i) = Y_i(t) - \hat{M}_i(t)\hat{x}(t - d) - \hat{Q}_i(t)U_i(t) - \hat{V}_i(t), \quad i = 1, 2.$$

Under the known $\hat{\theta}_1(t)$ and $\hat{\theta}_2(t)$, according to the least squares principle, the state estimation algorithm of the two-input two-output systems with time delay is summarized as follows:

$$\hat{x}_i(t - n_i) = Y_i(t) - \hat{M}_i(t)\hat{x}(t - d) - \hat{Q}_i(t)U_i(t) - \hat{V}_i(t), \quad (44)$$

$$Y_i(t) = [y_i(t - n_i), y_i(t - n_i + 1), \dots, y_i(t - 1)]^T, \quad (45)$$

$$U_i(t) = [u^T(t - n_i), u^T(t - n_i + 1), \dots, u^T(t - 1)]^T, \quad (46)$$

$$\hat{V}_i(t) = [\hat{v}_i(t - n_i), \hat{v}_i(t - n_i + 1), \dots, \hat{v}_i(t - 1)]^T, \quad (47)$$

$$\hat{X}(t - d + n_i) = [\hat{x}^T(t - d), \hat{x}^T(t - d + 1), \dots, \hat{x}^T(t - d + n_i - 1)]^T, \quad (48)$$

$$\hat{M}_i(t) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \hat{b}_{i1}(t) & \cdots & 0 & 0 \\ \hat{b}_{i2}(t) & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & 0 \\ \hat{b}_{i,n_i-1}(t) & \cdots & \hat{b}_{i1}(t) & 0 \end{bmatrix}, \quad (49)$$

$$\hat{Q}_i(t) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \hat{f}_{i1}(t) & \cdots & 0 & 0 \\ \hat{f}_{i2}(t) & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & 0 \\ \hat{f}_{i,n_i-1}(t) & \cdots & \hat{f}_{i1}(t) & 0 \end{bmatrix}, i = 1, 2, \quad (50)$$

$$\hat{F}(t) = \hat{W}^{-1}(t) \hat{K}(t), \quad (51)$$

$$\hat{B}(t) = \hat{W}^{-1}(t) \hat{J}(t), \quad (52)$$

$$\hat{\theta}_1(t) = \begin{bmatrix} \hat{\theta}_{11}(t) \\ \hat{\theta}_{12}(t) \\ \hat{\theta}_{13}(t) \end{bmatrix}, \quad \hat{\theta}_2(t) = \begin{bmatrix} \hat{\theta}_{21}(t) \\ \hat{\theta}_{22}(t) \\ \hat{\theta}_{23}(t) \end{bmatrix}, \quad (53)$$

$$\hat{\theta}_{11}(t) = [\hat{a}_1(1), \hat{a}_1(2), \dots, \hat{a}_1(n_1), \hat{a}_{12}(1), \hat{a}_{12}(2), \dots, \hat{a}_{12}(n_{12}), 0, \dots, 0]^T, \quad (54)$$

$$\hat{\theta}_{12}(t) = [\hat{g}_1^T(t), \hat{g}_2^T(t), \dots, \hat{g}_{n_1}^T(t)]^T, \quad (55)$$

$$\hat{\theta}_{13}(t) = [\hat{r}_1^T(t), \hat{r}_2^T(t), \dots, \hat{r}_{n_1}^T(t)]^T, \quad (56)$$

$$\hat{\theta}_{21}(t) = [\hat{a}_{21}(1), \hat{a}_{21}(2), \dots, \hat{a}_{21}(n_{21}), 0, \dots, 0, \hat{a}_2(1), \hat{a}_2(2), \dots, \hat{a}_2(n_2)]^T, \quad (57)$$

$$\hat{\theta}_{22}(t) = [\hat{h}_1^T(t), \hat{h}_2^T(t), \dots, \hat{h}_{n_2}^T(t)]^T, \quad (58)$$

$$\hat{\theta}_{23}(t) = [\hat{s}_1^T(t), \hat{s}_2^T(t), \dots, \hat{s}_{n_2}^T(t)]^T, \quad (59)$$

$$\hat{W}(t) = \left[\begin{array}{ccc|ccc} -\hat{a}_1(2) & \cdots & 1 & -\hat{a}_{12}(2) & \cdots & 0 \\ -\hat{a}_1(3) & \cdots & & -\hat{a}_{12}(3) & \cdots & \\ \vdots & & & \vdots & & \\ \vdots & & & -\hat{a}_{12}(n_{12}) & 0 & \\ -\hat{a}_1(n_1) & & & 0 & & \\ 1 & & & 0 & & \\ \hline -\hat{a}_{21}(2) & \cdots & 0 & -\hat{a}_2(2) & \cdots & 1 \\ -\hat{a}_{21}(3) & \cdots & & -\hat{a}_2(3) & \cdots & \\ \vdots & & & \vdots & & \\ -\hat{a}_{21}(n_{21}) & 0 & & \vdots & & \\ 0 & & & -\hat{a}_2(n_2) & & \end{array} \right], \quad (60)$$

$$\hat{K}(t) = [\hat{g}_1^T(t), \dots, \hat{g}_{n_1}^T(t), \hat{h}_1^T(t), \dots, \hat{h}_{n_2}^T(t)]^T, \quad (61)$$

$$\hat{J}(t) = [\hat{r}_1^T(t), \dots, \hat{r}_{n_1}^T(t), \hat{s}_1^T(t), \dots, \hat{s}_{n_2}^T(t)]^T. \quad (62)$$

The main steps of the proposed algorithm are summarized as follows.

- Input and output:** $u(t), y(t)$. Form the information vector $\hat{\phi}_1(t)$ and $\hat{\phi}_2(t)$, further get the output $Y_1(t), Y_2(t)$ and the input $U_1(t), U_2(t)$.
- Initialization** Let $t = 1$, set the initial values $\hat{\theta}_1(0) = \mathbf{1}_n/p_0$, $\hat{\theta}_2(0) = \mathbf{1}_n/p_0$, the initial covariance matrix $P_1(0) = p_0 \mathbf{I}, P_2(0) = p_0 \mathbf{I}$, $p_0 = 10^6$, $u(t) = 0, y(t) = 0$ and $\hat{v}(t) = 1/p_0$ for $t \leq 0$.
- Compute the parameter estimates** Calculate the gain vector $L_1(t), L_2(t)$ and the covariance matrix $P_1(t), P_2(t)$. Update the parameter estimate $\hat{\theta}_1(t)$ and $\hat{\theta}_2(t)$. Compute $\hat{v}_1(t), \hat{v}_2(t)$ and form $\hat{V}_1(t), \hat{V}_2(t)$. If $\|\hat{\theta}_1(t) - \hat{\theta}_1(t-1)\| \leq \varepsilon, \|\hat{\theta}_2(t) - \hat{\theta}_2(t-1)\| \leq \varepsilon, \varepsilon > 0$, then terminate the procedure and obtain the estimates $\hat{\theta}_1(t)$ and $\hat{\theta}_2(t)$; otherwise, increase t by 1 and go to Step 1.
- Compute the states** Compute $\hat{a}_i(t), \hat{b}_i(t), \hat{f}_i(t)$, and form $\hat{M}_i(t), \hat{Q}_i(t)$. Compute $\hat{x}_i(t - n_i)$.

Remark 4. According to the state equation at different time t , the state vector is represented by measurable input and output variables, and the identification model of the system is derived. Then, the single-input single-output model algorithm is generalized, and its corresponding residual-based augmented least squares algorithm is derived. The estimated parameters are used to identify system status. The proposed algorithm is computationally intensive and highly accurate.

Remark 5. The convergence properties of the proposed algorithm can be analyzed by means of the martingale convergence theorem [58–60].

5 | EXAMPLE

Consider the two-input two-output state space system with 2-step state-delay, the parameters are

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0.32 & 0.49 & 0.55 & 0.01 \\ 0 & 0 & 0 & 1 \\ -0.40 & -0.86 & -0.60 & -0.65 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.20 & -0.25 & 0.01 & 0.01 \\ 0.10 & 0.10 & 0.23 & 0.14 \\ 0.30 & 0.01 & 0.20 & 0.10 \\ 0.10 & 0.50 & 0.10 & 0.20 \end{bmatrix},$$

TABLE 1 The parameter estimates and errors ($\sigma^2 = 0.10^2$)

t	100	200	500	1000	2000	3000
$\theta_{11}(1) = 0.32000$	0.33323	0.33133	0.32130	0.32189	0.31938	0.32011
$\theta_{11}(2) = 0.49000$	0.55353	0.55688	0.56101	0.53621	0.52862	0.51793
$\theta_{11}(3) = 0.55000$	0.47000	0.47997	0.53305	0.53626	0.54508	0.54398
$\theta_{11}(4) = 0.01000$	-0.10662	-0.09320	-0.04877	-0.03429	-0.01833	-0.01297
$\theta_{12}(1) = 0.46100$	0.33646	0.34914	0.39761	0.41544	0.43096	0.43595
$\theta_{12}(2) = -1.45000$	-1.78539	-1.80924	-1.81610	-1.68662	-1.64324	-1.58898
$\theta_{12}(3) = 0.10000$	0.09481	0.09715	0.09261	0.09561	0.09820	0.09904
$\theta_{12}(4) = 5.00000$	5.01411	5.01028	5.00554	5.00143	5.00036	5.00066
$\theta_{13}(1) = -0.00100$	-0.03347	-0.02524	-0.02297	-0.01680	-0.01204	-0.00967
$\theta_{13}(2) = 0.22240$	0.27315	0.26559	0.23827	0.23503	0.22993	0.22954
$\theta_{13}(3) = 0.22310$	0.24304	0.23747	0.22903	0.22507	0.22503	0.22468
$\theta_{13}(4) = 0.13410$	0.05497	0.07462	0.10181	0.11434	0.12335	0.12490
$\theta_{13}(5) = 0.20000$	0.11593	0.12433	0.14428	0.16035	0.17142	0.17731
$\theta_{13}(6) = -0.25000$	-0.28450	-0.29954	-0.27371	-0.26484	-0.26212	-0.25808
$\theta_{13}(7) = 0.01000$	-0.00505	-0.02683	-0.01839	-0.01375	-0.00950	-0.00370
$\theta_{13}(8) = 0.01000$	-0.00178	-0.00225	-0.01584	-0.01028	-0.00560	0.00020
$\delta_1(\%)$	7.79630	7.94004	7.38806	4.83661	3.88061	2.81674
t	100	200	500	1000	2000	3000
$\theta_{11}(1) = -0.40000$	-0.37980	-0.41822	-0.39763	-0.39283	-0.39233	-0.39148
$\theta_{11}(2) = -0.86000$	-0.80657	-0.68193	-0.78550	-0.82997	-0.84960	-0.86341
$\theta_{11}(3) = -0.60000$	-0.88774	-0.67550	-0.69762	-0.69267	-0.67007	-0.66176
$\theta_{11}(4) = -0.65000$	-0.89629	-0.76787	-0.75996	-0.73791	-0.71229	-0.69996
$\theta_{12}(1) = 0.43600$	0.21078	0.31174	0.32703	0.34737	0.37277	0.38543
$\theta_{12}(2) = 5.30000$	5.01916	4.39057	4.91191	5.14152	5.24458	5.31549
$\theta_{12}(3) = -1.00000$	-0.99814	-0.99074	-0.99320	-0.99918	-0.99897	-0.99938
$\theta_{12}(4) = 0.00000$	0.00723	-0.02473	-0.01129	-0.00755	-0.00353	-0.00085
$\theta_{13}(1) = 0.46700$	0.33712	0.31694	0.35822	0.39296	0.41551	0.42825
$\theta_{13}(2) = 0.29150$	0.36207	0.27650	0.28655	0.29273	0.29063	0.29260
$\theta_{13}(3) = 0.23860$	0.15888	0.11444	0.14511	0.17147	0.18971	0.20111
$\theta_{13}(4) = 0.27360$	0.17982	0.23558	0.23143	0.23078	0.24020	0.24426
$\theta_{13}(5) = 0.30000$	0.13147	0.14948	0.19789	0.23163	0.25606	0.26937
$\theta_{13}(6) = 0.01000$	0.09225	0.12605	0.11173	0.08131	0.06317	0.05272
$\theta_{13}(7) = 0.20000$	0.05266	0.04479	0.10487	0.13806	0.16192	0.17608
$\theta_{13}(8) = 0.10000$	0.13306	0.01138	0.05407	0.07719	0.08686	0.09461
$\delta_2(\%)$	10.85096	17.82920	8.79871	4.91667	3.01024	2.27523

$$\mathbf{x}F = \begin{bmatrix} 0.1 & 5 \\ 0.5 & 1 \\ -1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

In simulation, the input $\{\nu(t)\}$ is generated from uniform distribution and is taken as an uncorrelated persistent excitation signal sequence with zero mean and unit variance, and $\{\nu(t)\}$ as a white noise sequence is generated from Gaus-

sian distribution with zero mean and variances $\sigma^2 = 0.10^2$ and $\sigma^2 = 0.50^2$. Applying the estimation algorithm in (17)–(21) and the state estimation algorithm in (44)–(62) to estimate the parameter vector and the state vector, the simulation results for different noise variances are shown in Tables 1–2 and the parameter estimation errors δ versus t are shown in Figures 1–2, where $\delta := \|\hat{\theta}(t) - \theta\|/\|\theta\|$, the state estimates and estimation errors versus t are shown in Figures 3–6 (Solid line: State true value $x(t)$, Dot line: State estimated value $\hat{x}(t)$).

TABLE 2 The parameter estimates and errors ($\sigma^2 = 0.50^2$)

t	100	200	500	1000	2000	3000
$\theta_{11}(1) = 0.32000$	0.35824	0.35174	0.32679	0.33053	0.32039	0.32089
$\theta_{11}(2) = 0.49000$	0.56783	0.62383	0.63961	0.57088	0.57876	0.55134
$\theta_{11}(3) = 0.55000$	0.37275	0.41366	0.50710	0.49560	0.52760	0.53498
$\theta_{11}(4) = 0.01000$	-0.23034	-0.18695	-0.11321	-0.09020	-0.05915	-0.03801
$\theta_{12}(1) = 0.46100$	0.15183	0.18228	0.30506	0.35143	0.37912	0.39998
$\theta_{12}(2) = -1.45000$	-1.89288	-2.23143	-2.26346	-1.88623	-1.89789	-1.75371
$\theta_{12}(3) = 0.10000$	0.13195	0.14960	0.10786	0.10041	0.10083	0.10245
$\theta_{12}(4) = 5.00000$	5.06358	5.02870	5.00778	4.99859	4.99795	5.00106
$\theta_{13}(1) = -0.00100$	-0.01936	-0.04388	-0.04573	-0.02174	-0.01536	-0.00829
$\theta_{13}(2) = 0.22240$	0.39880	0.34557	0.28545	0.27757	0.25987	0.25313
$\theta_{13}(3) = 0.22310$	0.29878	0.25992	0.24702	0.24213	0.24361	0.24201
$\theta_{13}(4) = 0.13410$	0.04245	0.04720	0.09592	0.11043	0.12364	0.12610
$\theta_{13}(5) = 0.20000$	0.05028	0.05727	0.08788	0.12666	0.13952	0.15643
$\theta_{13}(6) = -0.25000$	-0.40911	-0.38944	-0.32104	-0.31158	-0.30945	-0.29278
$\theta_{13}(7) = 0.01000$	-0.02743	-0.06997	-0.05712	-0.04923	-0.04426	-0.02234
$\theta_{13}(8) = 0.01000$	0.02134	-0.00932	-0.02996	-0.01028	-0.01329	-0.00056
$\delta_1(\%)$	13.16158	17.26615	16.39064	9.21619	9.09722	6.20602
t	100	200	500	1000	2000	3000
$\theta_{11}(1) = -0.40000$	-0.45308	-0.44816	-0.44920	-0.44329	-0.43244	-0.41958
$\theta_{11}(2) = -0.86000$	-0.52383	-0.46546	-0.50568	-0.59307	-0.66158	-0.73180
$\theta_{11}(3) = -0.60000$	-0.64117	-0.66040	-0.67489	-0.66275	-0.63280	-0.63739
$\theta_{11}(4) = -0.65000$	-0.79551	-0.86462	-0.84811	-0.80126	-0.75338	-0.73133
$\theta_{12}(1) = 0.43600$	0.28401	0.21731	0.26339	0.28629	0.32767	0.35017
$\theta_{12}(2) = 5.30000$	3.56959	3.22069	3.44162	3.91283	4.29220	4.65195
$\theta_{12}(3) = -1.00000$	-1.04865	-0.99878	-0.98843	-0.99868	-0.99514	-0.99810
$\theta_{12}(4) = 0.00000$	-0.06058	-0.02795	-0.02279	-0.01566	-0.00584	0.00137
$\theta_{13}(1) = 0.46700$	0.20094	0.21362	0.24797	0.29672	0.31582	0.33734
$\theta_{13}(2) = 0.29150$	0.14114	0.22112	0.26196	0.26052	0.24329	0.24893
$\theta_{13}(3) = 0.23860$	0.02026	0.04474	0.08310	0.09669	0.10066	0.11759
$\theta_{13}(4) = 0.27360$	0.21213	0.19479	0.23864	0.24027	0.24118	0.23474
$\theta_{13}(5) = 0.30000$	0.06717	0.02043	0.04688	0.10141	0.15251	0.19085
$\theta_{13}(6) = 0.01000$	0.20522	0.13866	0.13498	0.13232	0.14350	0.14225
$\theta_{13}(7) = 0.20000$	-0.04823	-0.09095	-0.08206	-0.03265	0.03308	0.08847
$\theta_{13}(8) = 0.10000$	-0.14198	-0.13627	-0.08112	-0.04237	-0.03059	-0.00158
$\delta_2(\%)$	33.40628	39.52957	35.27017	26.57254	19.58267	13.07450

From Tables 1–2 and Figures 1–6, we can draw the following conclusions.

- The parameter estimation errors become smaller (in general) with the increasing of t .
- In the case of the same zero mean variance, the parameter estimation accuracy improves as the data length t increases.
- The data converges faster when the noise variance is lower.

- The state estimates are close to their true values with t increasing.

6 | CONCLUSIONS

The basic algorithm derivation principle of this paper is similar to the corresponding multi-input single-output model, but the number of parameters of this model is large, the dimension

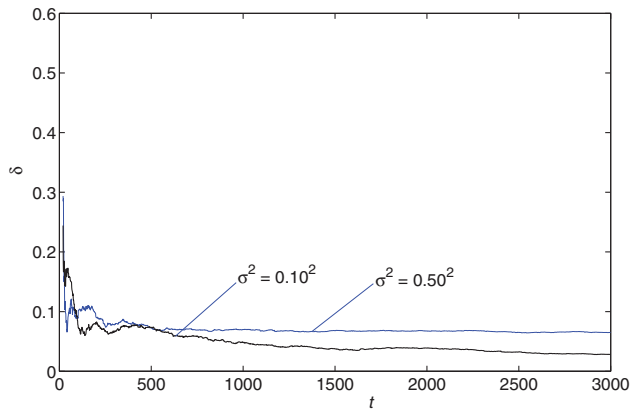


FIGURE 1 The parameter estimation errors δ_1 versus t

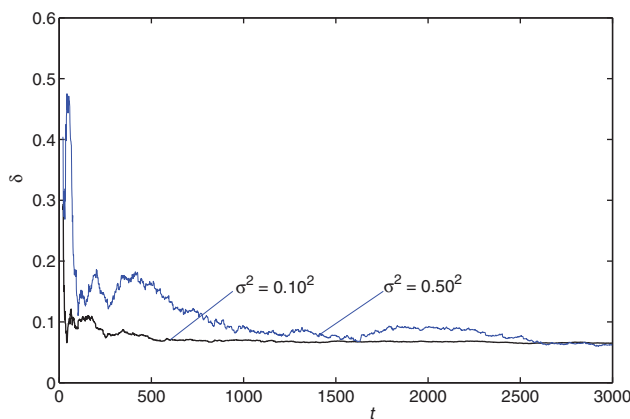


FIGURE 2 The parameter estimation errors δ_2 versus t

is complex, and there is coupling. When calculating the system state according to the hierarchical identification principle, it is necessary to combine the identification of two subsystems. The parameters make the identification more difficult, and because the recursive algorithm calculates the inverse of the matrix in the calculation process, the calculation amount is relatively large,

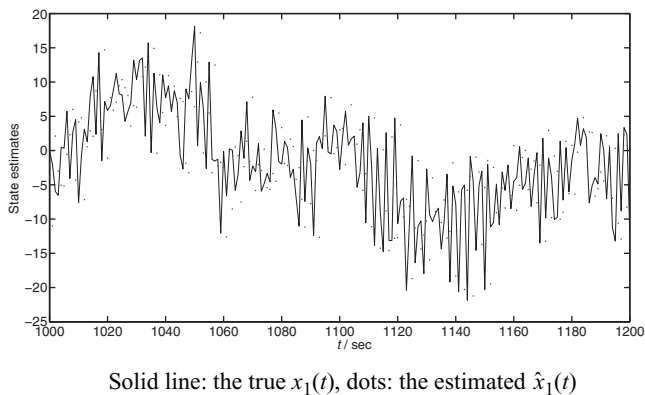


FIGURE 3 Solid line: the true $x_1(t)$, dots: the estimated $\hat{x}_1(t)$
The state and state estimate $\hat{x}_1(t)$ versus t

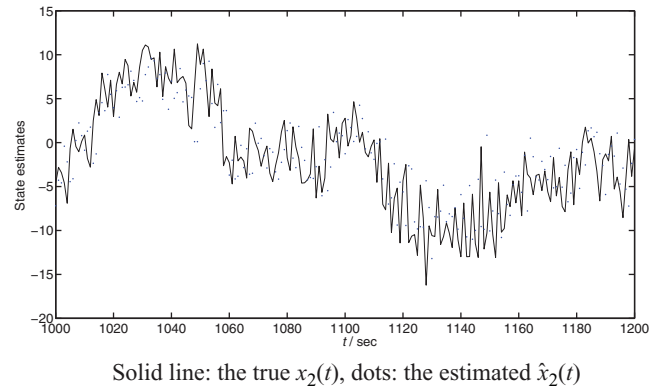


FIGURE 4 Solid line: the true $x_2(t)$, dots: the estimated $\hat{x}_2(t)$
The state and state estimate $\hat{x}_2(t)$ versus t

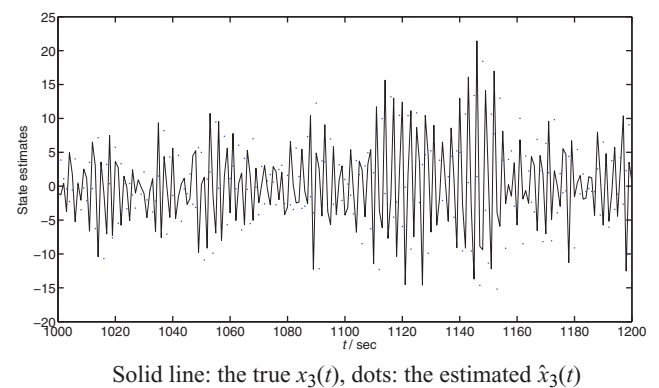


FIGURE 5 Solid line: the true $x_3(t)$, dots: the estimated $\hat{x}_3(t)$
The state and state estimate $\hat{x}_3(t)$ versus t

which affects the identification accuracy. This paper starts with a bivariate model with few dimensions to study the recursive least squares algorithm based on residuals. The proposed algorithms here can combine other estimation methods [61–66] to study parameter identification problems of different systems [67–72] and can be applied to other fields.

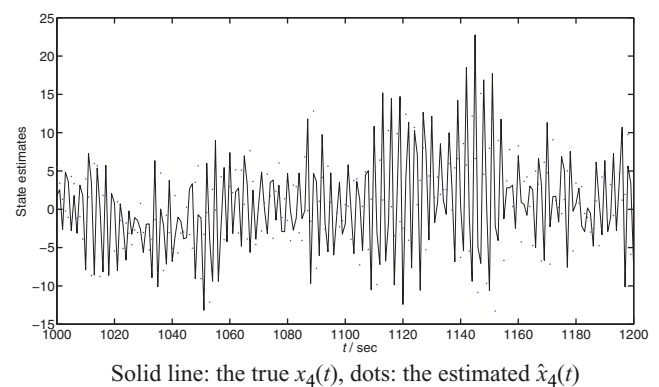


FIGURE 6 Solid line: the true $x_4(t)$, dots: the estimated $\hat{x}_4(t)$
The state and state estimate $\hat{x}_4(t)$ versus t

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (No. 61903050 and 61803049), the Natural Science Foundation of Jiangsu Province (No. BK20181033) and Natural Science Fundamental Research Project of Colleges and Universities in Jiangsu province (No. 19KJB470009 and 17KJB510002). The first author would like to show her gratitude to Professor Feng Ding and this work was completed under his supervision.

ORCID

Yu Gu  <https://orcid.org/0000-0003-2337-5043>

REFERENCES

- Xu, L.: Hierarchical multi-innovation generalised extended stochastic gradient methods for multivariable equation-error autoregressive moving average systems. *IET Control Theory Appl.* 14(10), 1276–1286 (2020)
- Xu, L.: Separable multi-innovation stochastic gradient estimation algorithm for the nonlinear dynamic responses of systems. *Int. J. Adapt. Control.* 34(7), 937–954 (2020)
- Zhang, X.: Recursive parameter estimation and its convergence for bilinear systems. *IET Control Theory Appl.* 14(5), 677–688 (2020)
- Zhang, X.: State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl.* 12(12), 1704–1713 (2018)
- Wang, D.Q., et al.: Model recovery for Hammerstein systems using the auxiliary model based orthogonal matching pursuit method. *Appl. Math. Model.* 54, 537–550 (2018)
- Ding, F., et al.: Gradient estimation algorithms for the parameter identification of bilinear systems using the auxiliary model. *J. Comput. Appl. Math.* 369, 112575 (2020)
- Chen, J., et al.: Variational Bayesian approach for ARX systems with missing observations and varying time-delays. *Automatica* 94, 194–204 (2018)
- Gan, M., Li, H.X., Peng, H.: A variable projection approach for efficient estimation of RBF-ARX model. *IEEE Trans. Cybernet.* 45(3), 462–471 (2015)
- Pan, J., et al.: A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems. *Int. J. Control Autom. Syst.* 15(3), 1189–1197 (2017)
- Zhang, L., Tang, S.Y., Lv, L.L.: An finite iterative algorithm for solving periodic Sylvester bimatrix equations. *J. Franklin Inst.* 357(15), 10757–10772 (2020)
- Zhang, L., et al.: Improved Dota2 lineup recommendation model based on a bidirectional LSTM. *Tsinghua Sci. Technol.* 25(6), 712–720 (2020)
- Xu, G., et al.: Fixed time synchronization control for bilateral teleoperation mobile manipulator with nonholonomic constraint and time delay. *IEEE Transactions on Circuits and Systems II: Express Briefs* 67(12), 3452–3456 (2020)
- Ding, F., et al.: A hierarchical least squares identification algorithm for Hammerstein nonlinear systems using the key term separation. *J. Franklin Inst.* 355(8), 3737–3752 (2018)
- Lv, L.L., et al.: A numerical solution of a class of periodic coupled matrix equations. *J. Franklin Inst.* 358(3), 2039–2059 (2021)
- Ansari, A., Bernstein, D.S.: Deadbeat unknown-input state estimation and input reconstruction for linear discrete-time systems. *Automatica* 103, 11–19 (2019)
- Xu, L., et al.: Hierarchical Newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses. *Int. J. Syst. Sci.* 50(1), 141–151 (2019)
- Xia, H.F., et al.: Improved least-squares identification for multiple-output non-linear stochastic systems. *IET Control Theory Appl.* 14(7), 964–971 (2020)
- Telmoudi, A.J., et al.: Modeling and state of health estimation of nickel-metal hydride battery using an EPSO-based fuzzy c-regression model. *Soft Comput.* 24(10), 7265–7279 (2020)
- Telmoudi, A.J., et al.: Parameter estimation of nonlinear systems using a robust possibilistic c-regression model algorithm. *Proceed. Inst. Mech. Eng.* 234(1), 134–143 (2020)
- Rafal, S., Marek, R., Krzysztof, L.: Modeling of discrete-time fractional-order state space systems using the balanced truncation method. *J. Franklin Inst.* 354(7), 3008–3020 (2017)
- Zhang, X., et al.: State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors. *Int. J. Adapt. Control.* 33(7), 1157–1173 (2019)
- Zhang, X.: Hierarchical parameter and state estimation for bilinear systems. *Int. J. Syst. Sci.* 51(2), 275–290 (2020)
- Zhang, X.: Adaptive parameter estimation for a general dynamical system with unknown states. *Int. J. Robust Nonlin.* 30(4), 1351–1372 (2020)
- Zhang, X.: Recursive parameter estimation methods and convergence analysis for a special class of nonlinear systems. *Int. J. Robust Nonlin.* 30(4), 1373–1393 (2020)
- Ding, F.: State filtering and parameter estimation for state space systems with scarce measurements. *Signal Process.* 104, 369–380 (2014)
- Ding, F.: Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Model.* 38(1), 403–412 (2014)
- Zhuang, L.F., et al.: Parameter and state estimation algorithm for single-input single-output linear systems using the canonical state space models. *Appl. Math. Model.* 36(8), 3454–3463 (2012)
- Meurer, F., Neubert, M., Werner, N.: Nonlinear state estimation for the Czochralski process based on the weighing signal using an extended Kalman filter. *J. Cryst. Growth* 419(1), 57–63 (2015)
- Alessandri, A., Gaggero, M.: Fast moving horizon state estimation for discrete-time systems using single and multi iteration descent methods. *IEEE Trans. Automat. Control* 62(9), 4499–4511 (2017)
- Chen, J., Zhu, Q.M., Li, J.: Biased compensation recursive least squares-based threshold algorithm for time-delay rational models via redundant rule. *Nonlinear Dyn.* 91(2), 797–807 (2018)
- Ding, F., et al.: Gradient based and least squares based iterative estimation algorithms for multi-input multi-output systems. *Proc. Inst. Mech. Eng., Part I: J. Syst. Control Eng.* 226(1), 43–55 (2012)
- Yang, X.Q., et al.: Robust identification of Wiener time-delay system with expectation-maximization algorithm. *J. Franklin Inst.* 354(13), 5678–5693 (2017)
- Sanz, R., Garcia, P., Krstic, M.: Observation and stabilization of LTV systems with time-varying measurement delay. *Automatica* 103, 573–579 (2019)
- Stojanovic, S.B.: Robust finite-time stability of discrete time systems with interval time-varying delay and nonlinear perturbations. *J. Franklin Inst.* 354(11), 4549–4572 (2017)
- Gu, Y., Lu, X.L., Ding, R.: Parameter and state estimation algorithm for a state space model with a one-unit state delay. *Circuits Syst. Signal Process.* 32(5), 2267–2280 (2013)
- Gu, Y., et al.: Parameter estimation for an input nonlinear state space system with time delay. *J. Franklin Inst.* 351(12), 5326–5339 (2014)
- Gu, Y., et al.: States based iterative parameter estimation for a state space model with multi-state delays using decomposition. *Signal Process.* 106, 294–300 (2015)
- Xu, L., et al.: Parameter estimation for control systems based on impulse responses. *Int. J. Control Autom. Syst.* 15(6), 2471–2479 (2017)
- Xu, L., et al.: The parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle. *IET Signal Process.* 11(2), 228–237 (2017)
- Ji, Y., et al.: Parameter estimation for block-oriented nonlinear systems using the key term separation. *Int. J. Robust Nonlin.* 30(9), 3727–3752 (2020)
- Ji, Y., Jiang, X.K., Wan, L.J.: Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems. *J. Franklin Inst.* 357(8), 5019–5032 (2020)

42. Xu, L., et al.: A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay. *Signal Process.* 140, 97–103 (2017)
43. Fan, Y.M., Liu, X.M.: Two-stage auxiliary model gradient-based iterative algorithm for the input nonlinear controlled autoregressive system with variable-gain nonlinearity. *Int. J. Robust Nonlin.* 30(14), 5492–5509 (2020)
44. Liu, X.M., Fan, Y.M.: Maximum likelihood extended gradient-based estimation algorithms for the input nonlinear controlled autoregressive moving average system with variable-gain nonlinearity. *Int. J. Robust Nonlin.* 31, 4017–4036 (2021)
45. Ding, F., et al.: Performance analysis of the generalized projection identification for time-varying systems. *IET Control Theory Appl.* 10(18), 2506–2514 (2016)
46. Xu, L., Song, G.L.: A recursive parameter estimation algorithm for modeling signals with multi-frequencies. *Circuits Syst. Signal Process.* 39(8), 4198–4224 (2020)
47. Li, M.H., Liu, X.M.: The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique. *Signal Process.* 147, 23–34 (2018)
48. Li, M.H., Liu, X.M.: Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique. *Int. J. Control Autom. Syst.* 18(6), 1581–1592 (2020)
49. Ding, F., et al.: Joint state and multi-innovation parameter estimation for time-delay linear systems and its convergence based on the Kalman filtering. *Digit. Signal Process.* 62, 211–223 (2017)
50. Ma, H., et al.: Partially-coupled least squares based iterative parameter estimation for multi-variable output-error-like autoregressive moving average systems. *IET Control Theory Appl.* 13(18), 3040–3051 (2019)
51. Xu, L., et al.: Separable recursive gradient algorithm for dynamical systems based on the impulse response signals. *Int. J. Control Autom. Syst.* 18(12), 3167–3177 (2020)
52. Pan, J., et al.: Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises. *IET Signal Process.* 14(7), 455–466 (2020)
53. Ding, F., et al.: Parameter estimation for pseudo-linear systems using the auxiliary model and the decomposition technique. *IET Control Theory Appl.* 11(3), 390–400 (2017)
54. Li, M.H., Liu, X.M.: Filtering-based maximum likelihood gradient iterative estimation algorithm for bilinear systems with autoregressive moving average noise. *Circuits Syst. Signal Process.* 37(11), 5023–5048 (2018)
55. Li, M.H., Liu, X.M.: The filtering-based maximum likelihood iterative estimation algorithms for a special class of nonlinear systems with autoregressive moving average noise using the hierarchical identification principle. *Int. J. Adapt. Control.* 33(7), 1189–1211 (2019)
56. Ding, F., et al.: Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique. *IET Control Theory Appl.* 12(7), 892–899 (2018)
57. Li, M.H., Liu, X.M.: Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems. *Int. J. Adapt. Control.* 35(2), 240–261 (2021)
58. Gu, Y., Zhu, Q., Nouri, H.: Bias compensation-based parameter and state estimation for a class of time-delay nonlinear state-space models. *IET Control Theory Appl.* 14(15), 2176–2185 (2020)
59. Gu, Y., et al.: State space model identification of multirate processes with time-delay using the expectation maximization. *J. Franklin Inst.* 356(3), 1623–1639 (2019)
60. Gu, Y., et al.: Moving horizon estimation for multirate systems with time-varying time-delays. *J. Franklin Inst.* 356(4), 2325–2345 (2019)
61. Zhang, X., et al.: State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle. *IET Control Theory Appl.* 12(12), 1704–1713 (2018)
62. Zhang, X., et al.: Recursive parameter identification of the dynamical models for bilinear state space systems. *Nonlinear Dyn.* 89(4), 2415–2429 (2017)
63. Xu, L., et al.: Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits Syst. Signal Process.* 36(4), 1735–1753 (2017)
64. Zhang, X., et al.: Highly computationally efficient state filter based on the delta operator. *Int. J. Adapt. Control.* 33(6), 875–889 (2019)
65. Xu, L.: The parameter estimation algorithms based on the dynamical response measurement data. *Adv. Mech. Eng.* 9(11), 1687814017730003 (2017)
66. Xu, L., et al.: Iterative parameter estimation for signal models based on measured data. *Circuits Syst. Signal Process.* 37(7), 3046–3069 (2018)
67. Ding, F., et al.: Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering. *J. Franklin Inst.* 354(3), 1321–1339 (2017)
68. Ma, H., et al.: Partially-coupled gradient-based iterative algorithms for multivariable output-error-like systems with autoregressive moving average noises. *IET Control Theory Appl.* 14(17), 2613–2627 (2020)
69. Xu, L., et al.: Hierarchical parameter estimation for the frequency response based on the dynamical window data. *Int. J. Control Autom. Syst.* 16(4), 1756–1764 (2018)
70. Xu, L., et al.: Auxiliary model multi-innovation stochastic gradient parameter estimation methods for nonlinear sandwich systems. *Int. J. Robust Nonlin.* 31(1), 148–165 (2021)
71. Wu, M.H., et al.: Object detection based on RGC mask R-CNN. *IET Image Process.* 14(8), 1502–1508 (2020)
72. Wan, X.K., et al.: Heartbeat classification algorithm based on one-dimensional convolution neural network. *J. Mech. Med. Biol.* 20(7), 2050046 (2020)

How to cite this article: Gu, Y., et al.: State filtering and parameter estimation for two-input two-output systems with time delay. *IET Control Theory Appl.* 15, 2053–2066 (2021). <https://doi.org/10.1049/cth2.12161>