IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 2, 2017, accepted January 3, 2018, date of publication January 16, 2018, date of current version May 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2794396

Auxiliary Model Based Least Squares Iterative Algorithms for Parameter Estimation of Bilinear Systems Using Interval-Varying Measurements

MEIHANG LI AND XIMEI LIU
College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266042, China

Corresponding author: Ximei Liu (liuximei@qust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472195 and in part by the Taishan Scholar Project Fund of Shandong Province of China.

ABSTRACT This paper focuses on the parameter estimation of a class of bilinear systems, for which the input-output representation is derived by eliminating the state variables in the systems. Based on the obtained identification model and the hierarchical identification principle, a hierarchical auxiliary model based least squares iterative algorithm is derived, to improve the computation efficiency and the parameter estimation accuracy by using the auxiliary model identification idea and the interval-varying input-output data. For comparison, an auxiliary model based least squares iterative algorithm is presented. The simulation results show that the proposed algorithm has better performance in estimating the parameters of bilinear systems.

INDEX TERMS Parameter estimation, hierarchical principle, auxiliary model, least squares, bilinear system.

I. INTRODUCTION

Parameter estimation is the foundation of dynamic systems modeling and controller design, and has been widely used in diverse fields of science and engineering [1]–[3]. In many of these applications, such as the control of industrial processes and the nuclear fission, the systems to be identified show nonlinear behaviour. Therefore, the parameter estimation of nonlinear systems is of primary importance [4]–[6]. Many parameter estimation methods have been proposed for nonlinear systems [7]–[10], such as the subspace identification method [11], [12], the intelligent algorithm [13], [14], the data-driven identification method [15] and the EM method [16]. Besides, for improving the convergence speed, a linear filter based multi-innovation stochastic gradient algorithm [17] and a state observer based hierarchical multi-innovation stochastic gradient algorithm [18] were proposed for nonlinear systems.

One particularly meaningful model for the nonlinear system identification is the bilinear model, which has the advantages of simpleness of structure and similarity to the linear model [19]. The inputs and the states in the bilinear model are multiplicatively coupled, which makes it can economically characterise a wide class of nonlinear phenomena. Applications include control systems, modeling and control of industrial processes such as chemical processes or biological processes, heat exchange systems, nuclear engineering, and others [20]–[24]. However, it has difficulty in using these models effectively because of the parametric uncertainty associated with the models. Hence, the parameter estimation of bilinear models is necessary and has become a hot topic in the field of system control and identification.

Several identification methods for bilinear systems have appeared in the literature. For example, some classical methods based on orthogonal series approach [25], Walsh funcitons [26], block-pulse functions [27], Chebyshev polynomials [28], Legendre polynomials [29], Taylor polynomials [30], Galerkin methods [31] and Hartley-based modulating functions [32] have been studied for bilinear systems at the early time. Besides, Inagaki and Mochizuki [33] proposed a Volterra kernels estimation method for bilinear systems; Tsoulkas *et al.* [34] derived an estimation method based on cumulants for input-output bilinear systems. However, the computational complexity of the methods in [33] and [34] increases exponentially as the order of the bilinear system increases. Recently, the maximum likelihood methods [35], [36], the least squares methods [37] and the gradient methods [38], [39] have been developed for parameter estimation of bilinear systems.

Generally, the least squares based iterative methods have a faster convergence speed than the gradient based iterative methods. Therefore, the gradient based iterative methods proposed in [38] and [39] have a slower convergence speed. However, the least squares based iterative methods involved with the matrix inversion, which needs much computation burden. The least squares estimation algorithm of bilinear systems in [37] is recursive. Different from them, and for improving the computation efficiency of the least squares based iterative methods, this paper focuses on the problem of estimating the parameters of bilinear systems with colored noise and develops some new least squares based iterative identification algorithms based on the auxiliary model identification idea and the hierarchical identification principle. The main contributions are as follows.

- The difficulty of identifying the bilinear state space systems is that their model structure contains the products of the states and inputs, the input-output representation of a bilinear system is derived from a bilinear state space system for the identification by eliminating the state variables in the models.
- According to the hierarchical identification principle and the auxiliary model identification idea, an auxiliary model based least squares iterative (AM-LSI) algorithm and a hierarchical auxiliary model based least squares iterative (H-AM-LSI) algorithm are derived for bilinear systems by using the interval-varying input-output data.
- As the computation of the covariance matrix with large sizes needs large computational burden, the least squares based iterative methods have low computational efficiency for large scale systems. Therefore, the proposed H-AM-LSI algorithm can reduce much computation burden when the system orders are large.

The rest of the paper is organized as follows. Section II simply derives the identification model for bilinear systems with colored noise. Section III presents an AM-LSI identification algorithm for comparison. Section IV develops a H-AM-LSI identification algorithm based on the hierarchical identification principle. Section V provides a numerical example validating the algorithms proposed. Finally, we make some concluding remarks in Section VI.

II. SYSTEM DESCRIPTION AND IDENTIFICATION MODEL

Let us define some notation. "$A =: X$" or "$X := A$" stands for "A is defined as X". 1_n represents an n-dimensional column vector whose elements are 1. The superscript T denotes the matrix transpose. z^{-1} stands for a unit backward shift operator: $z^{-1}y(t) = y(t-1)$.

Consider the following bilinear system with the observability canonical form:

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

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

where $\boldsymbol{x}(t) := [x_1(t), x_2(t), \cdots, x_n(t)]^{\mathrm{T}}$ is the n-dimensional state vector, $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$ are the system input and output, respectively, $w(t) \in \mathbb{R}$ is a correlated noise with zero mean and may be a white noise process, an autoregressive (AR) process, a moving average (MA) process or an ARMA process, and $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b} \in \mathbb{R}^{1\times n}$, $\boldsymbol{g} \in \mathbb{R}^{n}$ and $\boldsymbol{h} \in \mathbb{R}^{1\times n}$ are constant matrices and vectors:

$$\boldsymbol{A} := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix},$$

$$\boldsymbol{B} := \begin{bmatrix} 0 \\ \boldsymbol{b} \end{bmatrix} \in \mathbb{R}^{n\times n},$$

$$\boldsymbol{b} := [-b_n, -b_{n-1}, -b_{n-2}, \cdots, -b_1],$$

$$\boldsymbol{g} := [g_1, g_2, \cdots, g_{n-1}, g_n]^{\mathrm{T}},$$

$$\boldsymbol{h} := [1, 0, \cdots, 0, 0].$$

From (1), we have the following representation

$$\begin{cases} x_1(t+1) = x_2(t) + g_1 u(t), \\ x_2(t+1) = x_3(t) + g_2 u(t), \\ \quad \vdots \\ x_{n-1}(t+1) = x_n(t) + g_{n-1}u(t), \\ x_n(t+1) = -a_n x_1(t) - a_{n-1}x_2(t) - a_{n-2}x_3(t) \\ \qquad - \cdots - a_1 x_n(t) - [b_n x_1(t) + b_{n-1}x_2(t) \\ \qquad + b_{n-2}x_3(t) + \cdots + b_1 x_n(t)]u(t) + g_n u(t). \end{cases} \tag{3}$$

Moving the terms gives

$$\begin{cases} x_2(t) = x_1(t+1) - g_1 u(t), \\ x_3(t) = x_2(t+1) - g_2 u(t) \\ \qquad = x_1(t+2) - g_1 u(t+1) - g_2 u(t), \\ x_4(t) = x_3(t+1) - g_3 u(t) \\ \qquad = x_1(t+3) - g_1 u(t+2) - g_2 u(t+1) - g_3 u(t), \\ \quad \vdots \\ x_n(t) = x_{n-1}(t+1) - g_{n-1}u(t) \\ \qquad = x_1(t+n-1) - g_1 u(t+n-2) \\ \qquad\quad - g_2 u(t+n-3) - \cdots - g_{n-1}u(t). \end{cases} \tag{4}$$

Multiplying both sides of the last equation in (4) by z gives

$$x_n(t+1) = x_1(t+n) - g_1 u(t+n-1) - g_2 u(t+n-2) - \cdots - g_{n-1}u(t+1). \tag{5}$$

Substituting (5) into the last equation in (3), we have the following relation,

$$-[a_n, a_{n-1}, a_{n-2}, \cdots, a_1] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$$-[b_n, b_{n-1}, b_{n-2}, \cdots, b_1]\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ x_n(t) \end{bmatrix} u(t) + g_n u(t)$$
$$= x_1(t+n) - g_1 u(t+n-1) - g_2 u(t+n-2) - \cdots - g_{n-1}u(t+1). \tag{6}$$

According to the matrix form of Equation (4) and from (6), and referring to the method in [37] and [40], we have

$$\begin{aligned}&(1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n})z^n x_1(t)\\ &\quad + [(b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z_{-n})z^n x_1(t)]u(t)\\ &= [g_n + a_{n-1}g_1 + a_{n-2}g_2 + \cdots + a_1 g_{n-1}, g_{n-1} + a_{n-2}g_1\\ &\quad + a_{n-3}g_2 + \cdots + a_1 g_{n-2}, \cdots, g_2 + a_1 g_1, g_1]\\ &\quad \times \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+n-1)\end{bmatrix} + \{[b_{n-1}g_1 + b_{n-2}g_2 + \cdots + b_1 g_{n-1},\\ &\quad b_{n-2}g_1 + b_{n-3}g_2 + \cdots\\ &\quad + b_1 g_{n-2}, \cdots, b_1 g_1, 0]\begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+n-1)\end{bmatrix}\}u(t). \end{aligned}\tag{7}$$

Define

$$\begin{aligned}&[c_n, \cdots, c_2, c_1]\\ &\quad := [g_n + a_{n-1}g_1 + a_{n-2}g_2 + \cdots\\ &\qquad + a_1 g_{n-1}, \cdots, g_2 + a_1 g_1, g_1] \in \mathbb{R}^{1\times n}, \end{aligned}\tag{8}$$
$$\begin{aligned}&[d_n, \cdots, d_3, d_2]\\ &\quad := [b_{n-1}g_1 + b_{n-2}g_2 + \cdots + b_1 g_{n-1}, \cdots,\\ &\qquad b_1 g_1] \in \mathbb{R}^{1\times(n-1)}. \end{aligned}\tag{9}$$

Then Equation (7) can be rewritten as

$$\begin{aligned}&(1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n})z^n x_1(t)\\ &\quad + [(b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n})z^n x_1(t)]u(t)\\ &= (c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n})z^n u(t)\\ &\quad + [(d_2 z^{-2} + d_3 z^{-3} + \cdots + d_n z^{-n})z^n u(t)]u(t). \end{aligned}\tag{10}$$

Define the following polynomials:

$$\begin{aligned}A(z) &:= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_n z^{-n}, \quad a_i \in \mathbb{R},\\ B(z) &:= b_1 z^{-1} + b_2 z^{-2} + \cdots + b_n z^{-n}, \quad b_i \in \mathbb{R},\\ C(z) &:= c_1 z^{-1} + c_2 z^{-2} + \cdots + c_n z^{-n}, \quad c_i \in \mathbb{R},\\ D(z) &:= d_2 z^{-2} + d_3 z^{-3} + \cdots + d_n z^{-n}, \quad d_i \in \mathbb{R}.\end{aligned}$$

Hence, Equation (10) can be rewritten as

$$A(z)z^n x_1(t) + u(t)[B(z)z^n x_1(t)] = C(z)z^n u(t) + u(t)[D(z)z^n u(t)].$$

Replacing t with $t-n$, we have

$$x_1(t) = \frac{C(z) + u(t-n)D(z)}{A(z) + u(t-n)B(z)}u(t).$$

Inserting $x_1(t)$ into (2), we can obtain the input-output representation of the bilinear state space system in (1)–(2) as

$$y(t) = \frac{C(z) + u(t-n)D(z)}{A(z) + u(t-n)B(z)}u(t) + w(t). \tag{11}$$

For studying the parameter estimation of bilinear systems, this paper considers $w(t)$ in (11) as an AR process. That is, $E(z)w(t) = v(t)$, where $v(t) \in \mathbb{R}$ is a white noise process with zero mean, $E(z)$ is a polynomial in z^{-1} and

$$E(z) := 1 + e_1 z^{-1} + e_2 z^{-2} + \cdots + e_{n_e} z^{-n_e}, \quad e_i \in \mathbb{R}.$$

Assume that the orders n and n_e are known and $u(t) = 0$, $y(t) = 0$ and $v(t) = 0$ for $t \leqslant 0$. The objective is to develop new least squares based iterative algorithms for estimating the parameters a_i, b_i, c_i, d_i and e_i from the observation data by using the auxiliary model identification idea.

Introduce the intermediate variable

$$\alpha(t) := \frac{C(z) + u(t-n)D(z)}{A(z) + u(t-n)B(z)}u(t) \in \mathbb{R}. \tag{12}$$

Define the parameter vectors

$$\begin{aligned}\boldsymbol{\theta} &:= [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}, \boldsymbol{c}^{\mathrm{T}}, \boldsymbol{d}^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{4n+n_e-1},\\ \boldsymbol{a} &:= [a_1, a_2, \cdots, a_n]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{b} &:= [b_1, b_2, \cdots, b_n]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{c} &:= [c_1, c_2, \cdots, c_n]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{d} &:= [d_2, d_3, \cdots, d_n]^{\mathrm{T}} \in \mathbb{R}^{n-1},\\ \boldsymbol{e} &:= [e_1, e_2, \cdots, e_{n_e}]^{\mathrm{T}} \in \mathbb{R}^{n_e},\end{aligned}$$

and the information vectors

$$\begin{aligned}\boldsymbol{\varphi}(t) &:= [\boldsymbol{\varphi}_1^{\mathrm{T}}(t), \boldsymbol{\varphi}_2^{\mathrm{T}}(t), \boldsymbol{\phi}_1^{\mathrm{T}}(t), \boldsymbol{\phi}_2^{\mathrm{T}}(t), \boldsymbol{\psi}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{4n+n_e-1},\\ \boldsymbol{\varphi}_1(t) &:= [-\alpha(t-1), -\alpha(t-2), \cdots, -\alpha(t-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{\varphi}_2(t) &:= [-u(t-n)\alpha(t-1), -u(t-n)\alpha(t-2), \cdots,\\ &\qquad - u(t-n)\alpha(t-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{\phi}_1(t) &:= [u(t-1), u(t-2), \cdots, u(t-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\ \boldsymbol{\phi}_2(t) &:= [u(t-n)u(t-2), u(t-n)u(t-3), \cdots,\\ &\qquad u(t-n)u(t-n)]^{\mathrm{T}} \in \mathbb{R}^{n-1},\\ \boldsymbol{\psi}(t) &:= [-w(t-1), -w(t-2), \cdots, -w(t-n_e)]^{\mathrm{T}} \in \mathbb{R}^{n_e}.\end{aligned}$$

Thus, $E(z)w(t) = v(t)$ can be written as

$$w(t) = [1 - E(z)]w(t) + v(t) = \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{e} + v(t). \tag{13}$$

Equation (13) is the noise model.

From (12), we have

$$\begin{aligned}\alpha(t) &= [1 - A(z) - u(t-n)B(z)]\alpha(t)\\ &\quad + [C(z) + u(t-n)D(z)]u(t)\\ &= -\sum_{i=1}^n a_i\alpha(t-i) - \sum_{i=1}^n b_i u(t-n)\alpha(t-i)\\ &\quad + \sum_{i=1}^n c_i u(t-i) + \sum_{i=2}^n d_i u(t-n)u(t-i)\\ &= \boldsymbol{\varphi}_1^{\mathrm{T}}(t)\boldsymbol{a} + \boldsymbol{\varphi}_2^{\mathrm{T}}(t)\boldsymbol{b} + \boldsymbol{\phi}_1^{\mathrm{T}}(t)\boldsymbol{c} + \boldsymbol{\phi}_2^{\mathrm{T}}(t)\boldsymbol{d}.\end{aligned}\tag{14}$$

Substituting (13) and (14) into (11) gives

$$\begin{aligned} y(t) &= \alpha(t) + w(t) \quad (15)\\ &= \boldsymbol{\varphi}_1^{\mathrm{T}}(t)\boldsymbol{a} + \boldsymbol{\varphi}_2^{\mathrm{T}}(t)\boldsymbol{b} + \boldsymbol{\phi}_1^{\mathrm{T}}(t)\boldsymbol{c} + \boldsymbol{\phi}_2^{\mathrm{T}}(t)\boldsymbol{d} + w(t)\\ &= \boldsymbol{\varphi}_1^{\mathrm{T}}(t)\boldsymbol{a} + \boldsymbol{\varphi}_2^{\mathrm{T}}(t)\boldsymbol{b} + \boldsymbol{\phi}_1^{\mathrm{T}}(t)\boldsymbol{c} + \boldsymbol{\phi}_2^{\mathrm{T}}(t)\boldsymbol{d} + \boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{e}\\ &\quad + v(t) \quad (16)\\ &= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t). \quad (17) \end{aligned}$$

Equation (17) is the identification model for the bilinear system in (11), the parameter vector $\boldsymbol{\theta}$ contains the parameters a_i, b_i, c_i and d_i of the system model and the parameter e_i of the noise model.

III. THE AUXILIARY MODEL BASED LEAST SQUARES ITERATIVE ALGORITHM

For the purpose of showing the advantages of the proposed algorithm, the following gives the auxiliary model based least squares iterative algorithm using interval-varying input-output data for comparisons.

Define an integer sequence $\{t_s : s = 0, 1, 2, \cdots\}$ which satisfies

$$\begin{aligned} &0 = t_0 < t_1 < t_2 < t_3 < \cdots < t_{s-1} < t_s < t_{s+1} < \cdots,\\ &1 \leq t_s^* := t_{s+1} - t_s. \quad (18) \end{aligned}$$

Replacing t in (14)-(17) with t_s gives

$$\begin{aligned} \alpha(t_s) &= \boldsymbol{\varphi}_1^{\mathrm{T}}(t_s)\boldsymbol{a} + \boldsymbol{\varphi}_2^{\mathrm{T}}(t_s)\boldsymbol{b} + \boldsymbol{\phi}_1^{\mathrm{T}}(t_s)\boldsymbol{c} + \boldsymbol{\phi}_2^{\mathrm{T}}(t_s)\boldsymbol{d}, \quad (19)\\ y(t_s) &= \alpha(t_s) + w(t_s) \quad (20)\\ &= \boldsymbol{\varphi}_1^{\mathrm{T}}(t_s)\boldsymbol{a} + \boldsymbol{\varphi}_2^{\mathrm{T}}(t_s)\boldsymbol{b} + \boldsymbol{\phi}_1^{\mathrm{T}}(t_s)\boldsymbol{c} + \boldsymbol{\phi}_2^{\mathrm{T}}(t_s)\boldsymbol{d}\\ &\quad + \boldsymbol{\psi}^{\mathrm{T}}(t_s)\boldsymbol{e} + v(t_s), \quad (21)\\ &= \boldsymbol{\varphi}^{\mathrm{T}}(t_s)\boldsymbol{\theta} + v(t_s) \quad (22) \end{aligned}$$

with

$$\begin{aligned} \boldsymbol{\varphi}(t_s) &= [\boldsymbol{\varphi}_1^{\mathrm{T}}(t_s), \boldsymbol{\varphi}_2^{\mathrm{T}}(t_s), \boldsymbol{\phi}_1^{\mathrm{T}}(t_s), \boldsymbol{\phi}_2^{\mathrm{T}}(t_s), \boldsymbol{\psi}^{\mathrm{T}}(t_s)]^{\mathrm{T}},\\ \boldsymbol{\varphi}_1(t_s) &= [-\alpha(t_s-1), -\alpha(t_s-2), \cdots, -\alpha(t_s-n)]^{\mathrm{T}},\\ \boldsymbol{\varphi}_2(t_s) &= [-u(t_s-n)\alpha(t_s-1), -u(t_s-n)\alpha(t_s-2), \cdots,\\ &\quad -u(t_s-n)\alpha(t_s-n)]^{\mathrm{T}},\\ \boldsymbol{\phi}_1(t_s) &= [u(t_s-1), u(t_s-2), \cdots, u(t_s-n)]^{\mathrm{T}},\\ \boldsymbol{\phi}_2(t_s) &:= [u(t_s-n)u(t_s-2), u(t_s-n)u(t_s-3), \cdots,\\ &\quad u(t_s-n)u(t_s-n)]^{\mathrm{T}},\\ \boldsymbol{\psi}(t_s) &:= [-w(t_s-1), -w(t_s-2), \cdots, -w(t_s-n_e)]^{\mathrm{T}}. \end{aligned}$$

According to the identification model in (22), define a quadratic criterion function

$$J_1(\boldsymbol{\theta}) := \sum_{j=0}^{t_s^*-1}[y(t_s+j) - \boldsymbol{\varphi}^{\mathrm{T}}(t_s+j)\boldsymbol{\theta}]^2.$$

Minimizing $J_1(\boldsymbol{\theta})$ and letting its partial derivative with respect to $\boldsymbol{\theta}$ be zero give

$$\frac{\partial J_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -2\sum_{j=0}^{t_s^*-1}\boldsymbol{\varphi}(t_s+j)[y(t_s+j) - \boldsymbol{\varphi}^{\mathrm{T}}(t_s+j)\boldsymbol{\theta}] = 0.$$

Assume that the information vector $\boldsymbol{\varphi}(t_s)$ is persistently exciting, that is, $\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\varphi}(t_s+j)\boldsymbol{\varphi}^{\mathrm{T}}(t_s+j)\right]$ is an invertible matrix. Then the least squares estimate of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}}(t_s) = \left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\varphi}(t_s+j)\boldsymbol{\varphi}^{\mathrm{T}}(t_s+j)\right]^{-1}\sum_{j=0}^{t_s^*-1}\boldsymbol{\varphi}(t_s+j)y(t_s+j). \quad (23)$$

However, some problems arise. The information vector $\boldsymbol{\varphi}(t_s)$ contains the unknown terms $\alpha(t_s-i)$ $(i = 1, 2, \cdots, n)$ and the unmeasured noise terms $w(t_s-i)$ $(i = 1, 2, \cdots, n_e)$, Equation (23) cannot compute the estimate $\hat{\boldsymbol{\theta}}(t_s)$ directly. The approach is based on the auxiliary model identification idea and the iterative principle. Let $\hat{\alpha}_k(t_s-i)$ and $\hat{w}_k(t_s-i)$ be the estimates of $\alpha(t_s-i)$ and $w(t_s-i)$. Define $\hat{\boldsymbol{\varphi}}_k(t_s)$, $\hat{\boldsymbol{\varphi}}_{1,k}(t_s)$, $\hat{\boldsymbol{\varphi}}_{2,k}(t_s)$ and $\hat{\boldsymbol{\psi}}_k(t_s)$ as the estimates of $\boldsymbol{\varphi}(t_s)$, $\boldsymbol{\varphi}_1(t_s)$, $\boldsymbol{\varphi}_2(t_s)$ and $\boldsymbol{\psi}(t_s)$:

$$\begin{aligned} \hat{\boldsymbol{\varphi}}_k(t_s) &:= [\hat{\boldsymbol{\varphi}}_{1,k}^{\mathrm{T}}(t_s), \hat{\boldsymbol{\varphi}}_{2,k}^{\mathrm{T}}(t_s), \boldsymbol{\phi}_1^{\mathrm{T}}(t_s), \boldsymbol{\phi}_2^{\mathrm{T}}(t_s), \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}}\\ &\quad \in \mathbb{R}^{4n+n_e-1},\\ \hat{\boldsymbol{\varphi}}_{1,k}(t_s) &:= [-\hat{\alpha}_{k-1}(t_s-1), -\hat{\alpha}_{k-1}(t_s-2), \cdots,\\ &\quad -\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\ \hat{\boldsymbol{\varphi}}_{2,k}(t_s) &:= [-u(t_s-n)\hat{\alpha}_{k-1}(t_s-1),\\ &\quad -u(t_s-n)\hat{\alpha}_{k-1}(t_s-2), \cdots,\\ &\quad -u(t_s-n)\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\ \hat{\boldsymbol{\psi}}_k(t_s) &:= [-\hat{w}_{k-1}(t_s-1), -\hat{w}_{k-1}(t_s-2), \cdots,\\ &\quad -\hat{w}_{k-1}(t_s-n_e)]^{\mathrm{T}} \in \mathbb{R}^{n_e}. \end{aligned}$$

Let $\hat{\boldsymbol{\theta}}_k(t_s) := [\hat{\boldsymbol{a}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{b}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{c}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{d}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{e}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}}$ be the estimates of $\boldsymbol{\theta} = [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}, \boldsymbol{c}^{\mathrm{T}}, \boldsymbol{d}^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}}]^{\mathrm{T}}$ at iteration k. Based on the auxiliary model identification idea, we define an auxiliary model $\hat{\alpha}_k(t_s) = \hat{\boldsymbol{\varphi}}_{1,k}^{\mathrm{T}}(t_s)\hat{\boldsymbol{a}}_k(t_s) + \hat{\boldsymbol{\varphi}}_{2,k}^{\mathrm{T}}(t_s)\hat{\boldsymbol{b}}_k(t_s) + \boldsymbol{\phi}_1^{\mathrm{T}}(t_s)\hat{\boldsymbol{c}}_k(t_s) + \boldsymbol{\phi}_2^{\mathrm{T}}(t_s)\hat{\boldsymbol{d}}_k(t_s)$. From (20), we have $w(t_s-i) = y(t_s-i) - \alpha(t_s-i)$. Replacing $\alpha(t_s)$ with its estimates $\hat{\alpha}_k(t_s)$ gives the estimate of $w(t_s)$ as $\hat{w}_k(t_s-i) = y(t_s-i) - \hat{\alpha}_k(t_s-i)$.

Replacing $\boldsymbol{\varphi}(t_s)$ in (23) with $\hat{\boldsymbol{\varphi}}_k(t_s)$, we have

$$\begin{aligned} \hat{\boldsymbol{\theta}}_k(t_s) &= \left[\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\varphi}}_k(t_s+j)\hat{\boldsymbol{\varphi}}_k^{\mathrm{T}}(t_s+j)\right]^{-1}\\ &\quad \times \sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\varphi}}_k(t_s+j)y(t_s+j). \quad (24) \end{aligned}$$

From the above derivations, we can summarize the auxiliary model based least squares iterative (AM-LSI) identification algorithm for the bilinear systems using interval-varying input-output data:

$$\begin{aligned} \hat{\boldsymbol{\theta}}_k(t_s) &= \left[\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\varphi}}_k(t_s+j)\hat{\boldsymbol{\varphi}}_k^{\mathrm{T}}(t_s+j)\right]^{-1}\\ &\quad \times \sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\varphi}}_k(t_s+j)y(t_s+j), \quad (25) \end{aligned}$$

TABLE 1. The computational efficiency of the AM-LSI algorithm.

Expressions	Number of multiplications	Number of additions
$\hat{\boldsymbol{\theta}}_k(t_s)=\boldsymbol{\alpha}_k'(t_s)\boldsymbol{\beta}_k(t_s)\in\mathbb{R}^{n_0}$	n_0^2	$n_0(n_0-1)$
$\boldsymbol{\alpha}_k'(t_s):=\boldsymbol{\alpha}_k^{-1}(t_s)\in\mathbb{R}^{n_0\times n_0}$	n_0^3	$n_0^3-n_0^2$
$\boldsymbol{\alpha}_k(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\varphi}_k(t_s+j)\hat{\varphi}_k^{\mathrm{T}}(t_s+j)\in\mathbb{R}^{n_0\times n_0}$	$n_0^2t_s^*$	$n_0^2(t_s^*-1)$
$\boldsymbol{\beta}_k(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\varphi}_k(t_s+j)y(t_s+j)\in\mathbb{R}^{n_0}$	$n_0t_s^*$	$n_0(t_s^*-1)$
$\hat{\alpha}_k(t_s)=\hat{\varphi}_{1,k}^{\mathrm{T}}(t_s)\hat{\boldsymbol{a}}_k(t_s)+\hat{\varphi}_{2,k}^{\mathrm{T}}(t_s)\hat{\boldsymbol{b}}_k(t_s)$ $+\boldsymbol{\phi}_1^{\mathrm{T}}(t_s)\hat{\boldsymbol{c}}_k(t_s)+\boldsymbol{\phi}_2^{\mathrm{T}}(t_s)\hat{\boldsymbol{d}}_k(t_s)\in\mathbb{R}$	$4n-1$	$4n-2$
$\hat{w}_k(t_s-i)=y(t_s-i)-\hat{\alpha}_k(t_s-i)\in\mathbb{R}$	0	1
Sum	$n_0^3+n_0^2(t_s^*+1)+n_0t_s^*+4n-1$	$n_0^3+n_0^2(t_s^*-1)+n_0(t_s^*-2)+4n-1$
Total flops	$N_1=2n_0^3+2n_0^2t_s^*+2n_0t_s^*-2n_e$	

$$\hat{\boldsymbol{\varphi}}_k(t_s)=[\hat{\boldsymbol{\varphi}}_{1,k}^{\mathrm{T}}(t_s),\hat{\boldsymbol{\varphi}}_{2,k}^{\mathrm{T}}(t_s),\boldsymbol{\phi}_1^{\mathrm{T}}(t_s),\boldsymbol{\phi}_2^{\mathrm{T}}(t_s),\hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}},\tag{26}$$

$$\hat{\boldsymbol{\varphi}}_{1,k}(t_s)=[-\hat{\alpha}_{k-1}(t_s-1),-\hat{\alpha}_{k-1}(t_s-2),\cdots,-\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}},\tag{27}$$

$$\hat{\boldsymbol{\varphi}}_{2,k}(t_s)=[-u(t_s-n)\hat{\alpha}_{k-1}(t_s-1),-u(t_s-n)\hat{\alpha}_{k-1}(t_s-2),\cdots,-u(t_s-n)\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}},\tag{28}$$

$$\boldsymbol{\phi}_1(t_s)=[u(t_s-1),u(t_s-2),\cdots,u(t_s-n)]^{\mathrm{T}},\tag{29}$$

$$\boldsymbol{\phi}_2(t_s)=[u(t_s-n)u(t_s-2),u(t_s-n)u(t_s-3),\cdots,u(t_s-n)u(t_s-n)]^{\mathrm{T}},\tag{30}$$

$$\hat{\boldsymbol{\psi}}_k(t_s)=[-\hat{w}_{k-1}(t_s-1),-\hat{w}_{k-1}(t_s-2),\cdots,-\hat{w}_{k-1}(t_s-n_e)]^{\mathrm{T}}\in\mathbb{R}^{n_e},\tag{31}$$

$$\hat{\alpha}_k(j)=\hat{\boldsymbol{\varphi}}_{1,k}^{\mathrm{T}}(j)\hat{\boldsymbol{a}}_k(t_s)+\hat{\boldsymbol{\varphi}}_{2,k}^{\mathrm{T}}(j)\hat{\boldsymbol{b}}_k(t_s)+\boldsymbol{\phi}_1^{\mathrm{T}}(j)\hat{\boldsymbol{c}}_k(t_s)+\boldsymbol{\phi}_2^{\mathrm{T}}(j)\hat{\boldsymbol{d}}_k(t_s),\quad j=t_s,t_s+1,,\cdots,t_{s+1}-1,\tag{32}$$

$$\hat{w}_k(t_s-i)=y(t_s-i)-\hat{\alpha}_k(t_s-i),\quad i=1,2,\cdots,n_e,\tag{33}$$

$$\hat{\boldsymbol{\theta}}_k(t_s)=[\hat{\boldsymbol{a}}_k^{\mathrm{T}}(t_s),\hat{\boldsymbol{b}}_k^{\mathrm{T}}(t_s),\hat{\boldsymbol{c}}_k^{\mathrm{T}}(t_s),\hat{\boldsymbol{d}}_k^{\mathrm{T}}(t_s),\hat{\boldsymbol{e}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}},\tag{34}$$

$$\hat{\boldsymbol{a}}_k(t_s)=[\hat{a}_{1,k}(t_s),\hat{a}_{2,k}(t_s),\cdots,\hat{a}_{n,k}(t_s)]^{\mathrm{T}},\tag{35}$$

$$\hat{\boldsymbol{b}}_k(t_s)=[\hat{b}_{1,k}(t_s),\hat{b}_{2,k}(t_s),\cdots,\hat{b}_{n,k}(t_s)]^{\mathrm{T}},\tag{36}$$

$$\hat{\boldsymbol{c}}_k(t_s)=[\hat{c}_{1,k}(t_s),\hat{c}_{2,k}(t_s),\cdots,\hat{c}_{n,k}(t_s)]^{\mathrm{T}},\tag{37}$$

$$\hat{\boldsymbol{d}}_k(t_s)=[\hat{d}_{2,k}(t_s),\hat{d}_{3,k}(t_s),\cdots,\hat{d}_{n,k}(t_s)]^{\mathrm{T}},\tag{38}$$

$$\hat{\boldsymbol{e}}_k(t_s)=[\hat{e}_{1,k}(t_s),\hat{e}_{2,k}(t_s),\cdots,\hat{e}_{n_e,k}(t_s)]^{\mathrm{T}}.\tag{39}$$

The steps involved in the AM-LSI algorithm for computing the parameter estimates $\hat{\boldsymbol{\theta}}_k(t_s)$ of bilinear systems using interval-varying input-output data are listed as follows.

1) Set $s=0$, $t_0=0$, and let $t_s^*=t_{s+1}-t_s$ be a random positive integer, and give a small $\varepsilon>0$.
2) Collect the input-output data $\{u(j),y(j),j=t_s,t_s+1,\cdots,t_{s+1}-1\}$.
3) Let $k=1$, and set the initial values: $\hat{\alpha}_0(j)$ is a random number, $j=t_s,t_s+1,,\cdots,t_{s+1}-1$, and $\hat{w}_0(t_s-i)$ is a random number, $i=1,2,\cdots,n_e$.
4) Form $\hat{\boldsymbol{\varphi}}_{1,k}(t_s)$, $\hat{\boldsymbol{\varphi}}_{2,k}(t_s)$, $\boldsymbol{\phi}_1(t_s)$, $\boldsymbol{\phi}_2(t_s)$ and $\hat{\boldsymbol{\psi}}_k(t_s)$ using (27) – (31), respectively, and form $\hat{\boldsymbol{\varphi}}_k(t_s)$ using (26).
5) Update the estimate $\hat{\boldsymbol{\theta}}_k(t_s)$ using (25) and read $\hat{\boldsymbol{a}}_k(t_s)$, $\hat{\boldsymbol{b}}_k(t_s)$, $\hat{\boldsymbol{c}}_k(t_s)$ and $\hat{\boldsymbol{d}}_k(t_s)$ from $\hat{\boldsymbol{\theta}}_k(t_s)$ in (34).
6) Compute $\hat{\alpha}_k(t_s)$ using (32), and compute $\hat{w}_k(t_s-i)$ using (33).
7) Compare $\hat{\boldsymbol{\theta}}_k(t_s)$ with $\hat{\boldsymbol{\theta}}_{k-1}(t_s)$, if $\|\hat{\boldsymbol{\theta}}_k(t_s)-\hat{\boldsymbol{\theta}}_{k-1}(t_s)\|>\varepsilon$, increase k by 1 and go to Step 4; otherwise, increase s by 1 and go to Step 2.

The computational efficiency of the AM-LSI algorithm is shown in Table 1. The floating point operations (flops) are used to evaluate the computational efficiency of the algorithm, and we define n_0 in Tables 1–2 as $n_0:=4n+n_e-1$.

IV. THE HIERARCHICAL AUXILIARY MODEL BASED LEAST SQUARES ITERATIVE ALGORITHM

As the computation of the covariance matrix with large sizes needs large computational burden, the AM-LSI algorithm has low computational efficiency for large scale systems [41], [42]. In the following, we divide the bilinear system into three subsystems based on the hierarchical identification principle, and develop a hierarchical AM-LSI algorithm for identifying the bilinear system in (22) using interval-varying input-output data.

Note that the identification model in (21) includes five sub-information vectors $\boldsymbol{\varphi}_1(t_s)$, $\boldsymbol{\varphi}_2(t_s)$, $\boldsymbol{\phi}_1(t_s)$, $\boldsymbol{\phi}_2(t_s)$ and $\boldsymbol{\psi}(t_s)$, where $\boldsymbol{\varphi}_1(t_s)$ and $\boldsymbol{\varphi}_2(t_s)$ contain the unmeasurable variables $\alpha(t_s-i)$ and are unknown, $\boldsymbol{\phi}_1(t_s)$ and $\boldsymbol{\phi}_2(t_s)$ are consisting of the observed data and are known, $\boldsymbol{\psi}(t_s)$ contains the unmeasurable noise terms $w(t_s-i)$ and is unknown. Define two new information vectors

$$\boldsymbol{\Gamma}(t_s):=[\boldsymbol{\varphi}_1^{\mathrm{T}}(t_s),\boldsymbol{\varphi}_2^{\mathrm{T}}(t_s)]^{\mathrm{T}}\in\mathbb{R}^{2n},$$
$$\boldsymbol{\Phi}(t_s):=[\boldsymbol{\phi}_1^{\mathrm{T}}(t_s),\boldsymbol{\phi}_2^{\mathrm{T}}(t_s)]^{\mathrm{T}}\in\mathbb{R}^{2n-1},$$

and the corresponding parameter vectors

$$\boldsymbol{\vartheta}:=[\boldsymbol{a}^{\mathrm{T}},\boldsymbol{b}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{2n},$$
$$\boldsymbol{\zeta}:=[\boldsymbol{c}^{\mathrm{T}},\boldsymbol{d}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{2n-1}.$$

Define three intermediate variables

$$y_1(t_s):=y(t_s)-\boldsymbol{\Phi}^{\mathrm{T}}(t_s)\boldsymbol{\zeta}-\boldsymbol{\psi}^{\mathrm{T}}(t_s)\boldsymbol{e},\tag{40}$$
$$y_2(t_s):=y(t_s)-\boldsymbol{\Gamma}^{\mathrm{T}}(t_s)\boldsymbol{\vartheta}-\boldsymbol{\psi}^{\mathrm{T}}(t_s)\boldsymbol{e},\tag{41}$$
$$y_3(t_s):=y(t_s)-\boldsymbol{\Phi}^{\mathrm{T}}(t_s)\boldsymbol{\zeta}-\boldsymbol{\Gamma}^{\mathrm{T}}(t_s)\boldsymbol{\vartheta}.\tag{42}$$

TABLE 2. The computational efficiency of the H-AM-LSI algorithm.

Expressions	Number of multiplications	Number of additions
$\hat{\boldsymbol{\vartheta}}_k(t_s)=\boldsymbol{\alpha}'_{1,k}(t_s)\boldsymbol{\beta}_{1,k}(t_s)\in\mathbb{R}^{2n}$	$(2n)^2$	$2n(2n-1)$
$\boldsymbol{\alpha}'_{1,k}(t_s):=\boldsymbol{\alpha}^{-1}_{1,k}(t_s)\in\mathbb{R}^{(2n)\times(2n)}$	$(2n)^3$	$(2n)^3-(2n)^2$
$\boldsymbol{\alpha}_{1,k}(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\Gamma}}_k(t_s+j)\hat{\boldsymbol{\Gamma}}^{\rm T}_k(t_s+j)\in\mathbb{R}^{(2n)\times(2n)}$	$(2n)^2t_s^*$	$(2n)^2(t_s^*-1)$
$\boldsymbol{\beta}_{1,k}(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\Gamma}}_k(t_s+j)[y(t_s+j)$		
$-\boldsymbol{\Phi}^{\rm T}(t_s+j)\hat{\boldsymbol{\zeta}}_{k-1}(t_s)-\hat{\boldsymbol{\psi}}^{\rm T}_k(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)]\in\mathbb{R}^{2n}$	$n_0t_s^*$	$n_0t_s^*-2n$
$\hat{\boldsymbol{\zeta}}_k(t_s)=\boldsymbol{\alpha}'_2(t_s)\boldsymbol{\beta}_{2,k}(t_s)\in\mathbb{R}^{2n-1}$	$(2n-1)^2$	$(2n-1)(2n-2)$
$\boldsymbol{\alpha}'_2(t_s):=\boldsymbol{\alpha}^{-1}_2(t_s)\in\mathbb{R}^{(2n-1)\times(2n-1)}$	$(2n-1)^3$	$(2n-1)^3-(2n-1)^2$
$\boldsymbol{\alpha}_2(t_s)=\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\rm T}(t_s+j)\in\mathbb{R}^{(2n-1)\times(2n-1)}$	$(2n-1)^2t_s^*$	$(2n-1)^2(t_s^*-1)$
$\boldsymbol{\beta}_{2,k}(t_s)=\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)[y(t_s+j)$		
$-\hat{\boldsymbol{\Gamma}}^{\rm T}_k(t_s+j)\hat{\boldsymbol{\vartheta}}_k(t_s)-\hat{\boldsymbol{\psi}}^{\rm T}_k(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)]\in\mathbb{R}^{2n-1}$	$n_0t_s^*$	$n_0t_s^*-2n+1$
$\hat{\boldsymbol{e}}_k(t_s)=\boldsymbol{\alpha}'_{3,k}(t_s)\boldsymbol{\beta}_{3,k}(t_s)\in\mathbb{R}^{n_e}$	n_e^2	$n_e(n_e-1)$
$\boldsymbol{\alpha}'_{3,k}(t_s):=\boldsymbol{\alpha}^{-1}_{3,k}(t_s)\in\mathbb{R}^{n_e\times n_e}$	n_e^3	$n_e^3-n_e^2$
$\boldsymbol{\alpha}_{3,k}(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\psi}}_k(t_s+j)\hat{\boldsymbol{\psi}}^{\rm T}_k(t_s+j)\in\mathbb{R}^{n_e\times n_e}$	$n_e^2t_s^*$	$n_e^2(t_s^*-1)$
$\boldsymbol{\beta}_{3,k}(t_s)=\sum_{j=0}^{t_s^*-1}\hat{\boldsymbol{\psi}}_k(t_s+j)[y(t_s+j)$		
$-\boldsymbol{\Phi}^{\rm T}(t_s+j)\hat{\boldsymbol{\zeta}}_k(t_s)-\hat{\boldsymbol{\Gamma}}^{\rm T}_k(t_s+j)\hat{\boldsymbol{\vartheta}}_k(t_s)]\in\mathbb{R}^{n_e}$	$n_0t_s^*$	$n_0t_s^*-n_e$
$\hat{\alpha}_k(t_s)=\hat{\boldsymbol{\Gamma}}^{\rm T}_k(t_s)\hat{\boldsymbol{\vartheta}}_k(t_s)+\boldsymbol{\Phi}^{\rm T}(t_s)\hat{\boldsymbol{\zeta}}_k(t_s)\in\mathbb{R}$	$4n-1$	$4n-2$
$\hat{w}_k(t_s-i)=y(t_s-i)-\hat{\alpha}_k(t_s-i)\in\mathbb{R}$	0	1
Sum	$(2n)^3+(2n-1)^3+n_e^3$ $+[(2n)^2+(2n-1)^2+n_e^2]$ $\times(t_s^*+1)+3n_0t_s^*+4n-1$	$(2n)^3+(2n-1)^3+n_e^3$ $+[(2n)^2+(2n-1)^2+n_e^2]$ $\times(t_s^*-1)+3n_0t_s^*-2n_0+4n-1$
Total flops	$N_2=2[(2n)^3+(2n-1)^3+n_e^3]$ $+2[(2n)^2+(2n-1)^2+n_e^2]t_s^*+6n_0t_s^*-2n_e$	

According to the hierarchical identification principle, Equation (21) can be decomposed into the following three fictitious subsystems:

$$y_1(t_s)=\boldsymbol{\Gamma}^{\rm T}(t_s)\boldsymbol{\vartheta}+v(t_s), \tag{43}$$
$$y_2(t_s)=\boldsymbol{\Phi}^{\rm T}(t_s)\boldsymbol{\zeta}+v(t_s), \tag{44}$$
$$y_3(t_s)=\boldsymbol{\psi}^{\rm T}(t_s)\boldsymbol{e}+v(t_s). \tag{45}$$

The parameter vectors $\boldsymbol{\vartheta}=[\boldsymbol{a}^{\rm T},\boldsymbol{b}^{\rm T}]^{\rm T}$, $\boldsymbol{\zeta}=[\boldsymbol{c}^{\rm T},\boldsymbol{d}^{\rm T}]^{\rm T}$ and $\boldsymbol{e}$ to be identified are included in the three subsystems, respectively.

According to the identification models in (43)–(45), define three quadratic criterion functions

$$J_2(\boldsymbol{\vartheta}):=\sum_{j=0}^{t_s^*-1}[y_1(t_s+j)-\boldsymbol{\Gamma}^{\rm T}(t_s+j)\boldsymbol{\vartheta}]^2,$$
$$J_3(\boldsymbol{\zeta}):=\sum_{j=0}^{t_s^*-1}[y_2(t_s+j)-\boldsymbol{\Phi}^{\rm T}(t_s+j)\boldsymbol{\zeta}]^2,$$
$$J_4(\boldsymbol{e}):=\sum_{j=0}^{t_s^*-1}[y_3(t_s+j)-\boldsymbol{\psi}^{\rm T}(t_s+j)\boldsymbol{e}]^2.$$

Minimizing $J_2(\boldsymbol{\vartheta})$, $J_3(\boldsymbol{\zeta})$ and $J_4(\boldsymbol{e})$, and letting their partial derivatives with respect to $\boldsymbol{\vartheta}$, $\boldsymbol{\zeta}$ and $\boldsymbol{e}$ be zero, respectively, we have

$$\frac{\partial J_2(\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}=-2\sum_{j=0}^{t_s^*-1}\boldsymbol{\Gamma}(t_s+j)[y_1(t_s+j)-\boldsymbol{\Gamma}^{\rm T}(t_s+j)\boldsymbol{\vartheta}]=0,$$
$$\frac{\partial J_3(\boldsymbol{\zeta})}{\partial\boldsymbol{\zeta}}=-2\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)[y_2(t_s+j)-\boldsymbol{\Phi}^{\rm T}(t_s+j)\boldsymbol{\zeta}]=0,$$
$$\frac{\partial J_4(\boldsymbol{e})}{\partial\boldsymbol{e}}=-2\sum_{j=0}^{t_s^*-1}\boldsymbol{\psi}(t_s+j)[y_3(t_s+j)-\boldsymbol{\psi}^{\rm T}(t_s+j)\boldsymbol{e}]=0.$$

Assume that the information vectors $\boldsymbol{\Gamma}(t_s)$, $\boldsymbol{\Phi}(t_s)$ and $\boldsymbol{\psi}(t_s)$ are persistently exciting, that is, $\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\Gamma}(t_s+j)\boldsymbol{\Gamma}^{\rm T}(t_s+j)\right]$, $\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\rm T}(t_s+j)\right]$ and $\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\psi}(t_s+j)\boldsymbol{\psi}^{\rm T}(t_s+j)\right]$ are invertible matrixes. Then the least squares estimates of $\boldsymbol{\vartheta}$, $\boldsymbol{\zeta}$ and $\boldsymbol{e}$ are

$$\hat{\boldsymbol{\vartheta}}(t_s)=\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\Gamma}(t_s+j)\boldsymbol{\Gamma}^{\rm T}(t_s+j)\right]^{-1}\times\sum_{j=0}^{t_s^*-1}\boldsymbol{\Gamma}(t_s+j)y_1(t_s+j), \tag{46}$$

$$\hat{\boldsymbol{\zeta}}(t_s)=\left[\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\rm T}(t_s+j)\right]^{-1}\times\sum_{j=0}^{t_s^*-1}\boldsymbol{\Phi}(t_s+j)y_2(t_s+j), \tag{47}$$

$$\hat{\boldsymbol{e}}(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\psi}(t_s+j)\boldsymbol{\psi}^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \boldsymbol{\psi}(t_s+j)y_3(t_s+j). \tag{48}$$

Substituting (40)–(42) into (46)–(48), respectively, we have the following relations,

$$\hat{\boldsymbol{\vartheta}}(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\Gamma}(t_s+j)\boldsymbol{\Gamma}^{\mathrm{T}}(t_s+j)\right]^{-1} \sum_{j=0}^{t_s^*-1} \boldsymbol{\Gamma}(t_s+j) \times [y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\boldsymbol{\zeta} - \boldsymbol{\psi}^{\mathrm{T}}(t_s+j)\boldsymbol{e}], \tag{49}$$

$$\hat{\boldsymbol{\zeta}}(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\right]^{-1} \sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j) \times [y(t_s+j) - \boldsymbol{\Gamma}^{\mathrm{T}}(t_s+j)\boldsymbol{\vartheta} - \boldsymbol{\psi}^{\mathrm{T}}(t_s+j)\boldsymbol{e}], \tag{50}$$

$$\hat{\boldsymbol{e}}(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\psi}(t_s+j)\boldsymbol{\psi}^{\mathrm{T}}(t_s+j)\right]^{-1} \sum_{j=0}^{t_s^*-1} \boldsymbol{\psi}(t_s+j) \times [y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\boldsymbol{\zeta} - \boldsymbol{\Gamma}^{\mathrm{T}}(t_s+j)\boldsymbol{\vartheta}]. \tag{51}$$

However, we can see that the information vectors $\boldsymbol{\varphi}_1(t_s)$ and $\boldsymbol{\varphi}_2(t_s)$ in $\boldsymbol{\Gamma}(t_s)$ contain the unknown terms $\alpha(t_s - i)$ $(i = 1, 2, \cdots, n)$, the information vector $\boldsymbol{\psi}(t_s)$ contains the unknown noise terms $w(t_s - i)$ $(i = 1, 2, \cdots, n_e)$, and the parameter vectors $\boldsymbol{\vartheta}$, $\boldsymbol{\zeta}$ and $\boldsymbol{e}$ in (49)–(51) are also unknown. Therefore, Equations (49)–(51) cannot compute $\hat{\boldsymbol{\vartheta}}(t_s)$, $\hat{\boldsymbol{\zeta}}(t_s)$ and $\hat{\boldsymbol{e}}(t_s)$ directly. The approach is based on the auxiliary model identification idea and the iterative principle. Let $\hat{\alpha}_k(t_s - i)$ and $\hat{w}_k(t_s - i)$ be the estimates of $\alpha(t_s - i)$ and $w(t_s - i)$. Define $\hat{\boldsymbol{\Gamma}}_k(t_s)$, $\hat{\boldsymbol{\varphi}}_{1,k}(t_s)$, $\hat{\boldsymbol{\varphi}}_{2,k}(t_s)$ and $\hat{\boldsymbol{\psi}}_k(t_s)$ as the estimates of $\boldsymbol{\Gamma}(t_s)$, $\boldsymbol{\varphi}_1(t_s)$, $\boldsymbol{\varphi}_2(t_s)$ and $\boldsymbol{\psi}(t_s)$:

$$\begin{aligned}
\hat{\boldsymbol{\Gamma}}_k(t_s) &:= [\hat{\boldsymbol{\varphi}}_{1,k}^{\mathrm{T}}(t_s), \hat{\boldsymbol{\varphi}}_{2,k}^{\mathrm{T}}(t_s)]^{\mathrm{T}} \in \mathbb{R}^{2n},\\
\hat{\boldsymbol{\varphi}}_{1,k}(t_s) &:= [-\hat{\alpha}_{k-1}(t_s-1), -\hat{\alpha}_{k-1}(t_s-2), \cdots,\\
&\quad -\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}} \in \mathbb{R}^{n},\\
\hat{\boldsymbol{\varphi}}_{2,k}(t_s) &:= [-u(t_s-n)\hat{\alpha}_{k-1}(t_s-1),\\
&\quad -u(t_s-n)\hat{\alpha}_{k-1}(t_s-2), \cdots,\\
&\quad -u(t_s-n)\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}} \in \mathbb{R}^{n},\\
\hat{\boldsymbol{\psi}}_k(t_s) &:= [-\hat{w}_{k-1}(t_s-1), -\hat{w}_{k-1}(t_s-2), \cdots,\\
&\quad -\hat{w}_{k-1}(t_s-n_e)]^{\mathrm{T}} \in \mathbb{R}^{n_e}.
\end{aligned}$$

Let $\hat{\boldsymbol{\vartheta}}_k(t_s) := [\hat{\boldsymbol{a}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{b}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}}$, $\hat{\boldsymbol{\zeta}}_k(t_s) := [\hat{\boldsymbol{c}}_k^{\mathrm{T}}(t_s), \hat{\boldsymbol{d}}_k^{\mathrm{T}}(t_s)]^{\mathrm{T}}$ and $\hat{\boldsymbol{e}}_k(t_s)$ be the estimates of $\boldsymbol{\vartheta} = [\boldsymbol{a}^{\mathrm{T}}, \boldsymbol{b}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\zeta} = [\boldsymbol{c}^{\mathrm{T}}, \boldsymbol{d}^{\mathrm{T}}]^{\mathrm{T}}$ and $\boldsymbol{e}$ at iteration k. Based on the auxiliary model identification idea, we define an auxiliary model $\hat{\alpha}_k(t_s) = \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s)\hat{\boldsymbol{\vartheta}}_k(t_s) + \boldsymbol{\Phi}^{\mathrm{T}}(t_s)\hat{\boldsymbol{\zeta}}_k(t_s)$. From (20), we have $w(t_s-i) = y(t_s-i) - \alpha(t_s-i)$. Replacing $\alpha(t_s)$ with its estimates $\hat{\alpha}_k(t_s)$ gives the estimate of $w(t_s)$ as $\hat{w}_k(t_s - i) = y(t_s - i) - \hat{\alpha}_k(t_s - i)$.

Replacing $\boldsymbol{\Gamma}(t_s)$, $\boldsymbol{\psi}(t_s)$, $\boldsymbol{\zeta}$ and $\boldsymbol{e}$ in (49) with their estimates $\hat{\boldsymbol{\Gamma}}_k(t_s)$, $\hat{\boldsymbol{\psi}}_k(t_s)$, $\hat{\boldsymbol{\zeta}}_{k-1}(t_s)$ and $\hat{\boldsymbol{e}}_{k-1}(t_s)$, and replacing $\boldsymbol{\Gamma}(t_s)$, $\boldsymbol{\psi}(t_s)$, $\boldsymbol{\vartheta}$ and $\boldsymbol{e}$ in (50) with their estimates $\hat{\boldsymbol{\Gamma}}_k(t_s)$, $\hat{\boldsymbol{\psi}}_k(t_s)$, $\hat{\boldsymbol{\vartheta}}_k(t_s)$ and $\hat{\boldsymbol{e}}_{k-1}(t_s)$, and replacing $\boldsymbol{\psi}(t_s)$, $\boldsymbol{\Gamma}(t_s)$, $\boldsymbol{\zeta}$ and $\boldsymbol{\vartheta}$ in (51) with their estimates $\hat{\boldsymbol{\psi}}_k(t_s)$, $\hat{\boldsymbol{\Gamma}}_k(t_s)$, $\hat{\boldsymbol{\zeta}}_k(t_s)$ and $\hat{\boldsymbol{\vartheta}}_k(t_s)$, we have

$$\hat{\boldsymbol{\vartheta}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\Gamma}}_k(t_s+j)\hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\Gamma}}_k(t_s+j)[y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{\zeta}}_{k-1}(t_s) - \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)], \tag{52}$$

$$\hat{\boldsymbol{\zeta}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j)[y(t_s+j) - \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{\vartheta}}_k(t_s) - \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)], \tag{53}$$

$$\hat{\boldsymbol{e}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\psi}}_k(t_s+j)\hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\psi}}_k(t_s+j)[y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{\zeta}}_k(t_s) - \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{\vartheta}}_k(t_s)]. \tag{54}$$

Therefore, we can summarize the hierarchial auxiliary model based least squares iterative (H-AM-LSI) identification algorithm for bilinear systems using interval-varying input-output data:

$$\hat{\boldsymbol{\vartheta}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\Gamma}}_k(t_s+j)\hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\Gamma}}_k(t_s+j)[y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j) \times\hat{\boldsymbol{\zeta}}_{k-1}(t_s) - \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)], \tag{55}$$

$$\hat{\boldsymbol{\zeta}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j)\boldsymbol{\Phi}^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \boldsymbol{\Phi}(t_s+j)[y(t_s+j) - \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j) \times\hat{\boldsymbol{\vartheta}}_k(t_s) - \hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{e}}_{k-1}(t_s)], \tag{56}$$

$$\hat{\boldsymbol{e}}_k(t_s) = \left[\sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\psi}}_k(t_s+j)\hat{\boldsymbol{\psi}}_k^{\mathrm{T}}(t_s+j)\right]^{-1} \times \sum_{j=0}^{t_s^*-1} \hat{\boldsymbol{\psi}}_k(t_s+j)[y(t_s+j) - \boldsymbol{\Phi}^{\mathrm{T}}(t_s+j) \times\hat{\boldsymbol{\zeta}}_k(t_s) - \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(t_s+j)\hat{\boldsymbol{\vartheta}}_k(t_s)], \tag{57}$$

$$\begin{aligned}\hat{\boldsymbol{\Gamma}}_k(t_s) = [&-\hat{\alpha}_{k-1}(t_s-1), -\hat{\alpha}_{k-1}(t_s-2), \cdots,\\ &-\hat{\alpha}_{k-1}(t_s-n), -u(t_s-n)\hat{\alpha}_{k-1}(t_s-1),\\ &-u(t_s-n)\hat{\alpha}_{k-1}(t_s-2), \cdots,\\ &-u(t_s-n)\hat{\alpha}_{k-1}(t_s-n)]^{\mathrm{T}}, \end{aligned} \quad (58)$$

$$\begin{aligned}\boldsymbol{\Phi}(t_s) = [&u(t_s-1), u(t_s-2), \cdots, u(t_s-n),\\ &u(t_s-n)u(t_s-2), u(t_s-n)u(t_s-3), \cdots,\\ &u(t_s-n)u(t_s-n)]^{\mathrm{T}}, \end{aligned} \quad (59)$$

$$\begin{aligned}\hat{\boldsymbol{\psi}}_k(t_s) = [&-\hat{w}_{k-1}(t_s-1), -\hat{w}_{k-1}(t_s-2), \cdots,\\ &-\hat{w}_{k-1}(t_s-n_e)]^{\mathrm{T}}, \end{aligned} \quad (60)$$

$$\begin{aligned}\hat{\alpha}_k(j) &= \hat{\boldsymbol{\Gamma}}_k^{\mathrm{T}}(j)\hat{\boldsymbol{\vartheta}}_k(t_s) + \boldsymbol{\Phi}^{\mathrm{T}}(j)\hat{\boldsymbol{\xi}}_k(t_s),\\ &j = t_s, t_s+1, \cdots, t_{s+1}-1, \end{aligned} \quad (61)$$

$$\hat{w}_k(t_s-i) = y(t_s-i) - \hat{\alpha}_k(t_s-i), \quad i = 1, 2, \cdots, n_e, \quad (62)$$

$$\begin{aligned}\hat{\boldsymbol{\vartheta}}_k(t_s) = [&\hat{a}_{1,k}(t_s), \hat{a}_{2,k}(t_s), \cdots, \hat{a}_{n,k}(t_s), \hat{b}_{1,k}(t_s),\\ &\hat{b}_{2,k}(t_s), \cdots, \hat{b}_{n,k}(t_s)]^{\mathrm{T}}, \end{aligned} \quad (63)$$

$$\begin{aligned}\hat{\boldsymbol{\xi}}_k(t_s) = [&\hat{c}_{1,k}(t_s), \hat{c}_{2,k}(t_s), \cdots, \hat{c}_{n,k}(t_s), \hat{d}_{2,k}(t_s),\\ &\hat{d}_{3,k}(t_s), \cdots, \hat{d}_{n,k}(t_s)]^{\mathrm{T}}, \end{aligned} \quad (64)$$

$$\hat{\boldsymbol{e}}_k(t_s) = [\hat{e}_{1,k}(t_s), \hat{e}_{2,k}(t_s), \cdots, \hat{e}_{n_e,k}(t_s)]^{\mathrm{T}}. \quad (65)$$

The identification steps of the H-AM-LSI algorithm in (55)–(65) to compute the parameter estimates $\hat{\boldsymbol{\vartheta}}_k(t)$, $\hat{\boldsymbol{\xi}}_k(t)$ and $\hat{\boldsymbol{e}}_k(t)$ for bilinear systems using interval-varying input-output data are listed as follows.

1) Set $s = 0$, $t_0 = 0$, and let $t_s^* = t_{s+1} - t_s$ be a random positive integer. Set $\hat{\boldsymbol{\xi}}_0(t_s) = \mathbf{1}_{2n-1}/p_0$ and $\hat{\boldsymbol{e}}_0(t_s) = \mathbf{1}_{n_e}/p_0$, and give a small $\varepsilon > 0$.
2) Collect the input-output data $\{u(j), y(j), j = t_s, t_s+1, \cdots, t_{s+1}-1\}$.
3) Let $k = 1$, and set the initial values: $\hat{\alpha}_0(j)$ is a random number, $j = t_s, t_s+1, \cdots, t_{s+1}-1$, and $\hat{w}_0(t_s-i)$ is a random number, $i = 1, 2, \cdots, n_e$.
4) Form $\hat{\boldsymbol{\Gamma}}_k(t_s)$, $\boldsymbol{\Phi}(t_s)$ and $\hat{\boldsymbol{\psi}}_k(t_s)$ using (58) – (60), respectively.
5) Update the estimates $\hat{\boldsymbol{\vartheta}}_k(t_s)$, $\hat{\boldsymbol{\xi}}_k(t_s)$ and $\hat{\boldsymbol{e}}_k(t_s)$ using (55) – (57), respectively.
6) Compute $\hat{\alpha}_k(t_s)$ using (61), and compute $\hat{w}_k(t_s-i)$ using (62).
7) If $\|\hat{\boldsymbol{\vartheta}}_k(t_s) - \hat{\boldsymbol{\vartheta}}_{k-1}(t_s)\| + \|\hat{\boldsymbol{\xi}}_k(t_s) - \hat{\boldsymbol{\xi}}_{k-1}(t_s)\| + \|\hat{\boldsymbol{e}}_k(t_s) - \hat{\boldsymbol{e}}_{k-1}(t_s)\| > \varepsilon$, increase k by 1 and go to Step 4; otherwise, obtain the iteration k and the parameter estimates $\hat{\boldsymbol{\xi}}_k(t_s)$ and $\hat{\boldsymbol{e}}_k(t_s)$, and set $\hat{\boldsymbol{\xi}}_0(t_{s+1}) = \hat{\boldsymbol{\xi}}_k(t_s)$ and $\hat{\boldsymbol{e}}_0(t_{s+1}) = \hat{\boldsymbol{e}}_k(t_s)$, increase s by 1 and go to Step 2.

The computational efficiency of the H-AM-LSI algorithm is given in Table 2. The difference between the total flops of the AM-LSI algorithm and the H-AM-LSI algorithm is given by

$$\begin{aligned}N_1 - N_2 &= 2n_0^3 + 2n_0^2t_s^* + 2n_0t_s^* - 2n_e - \{2[(2n)^3\\ &\quad + (2n-1)^3 + n_e^3] + 2[(2n)^2 + (2n-1)^2 + n_e^2]t_s^*\\ &\quad + 6n_0t_s^* - 2n_e\}\\ &= 2[n_0^3 - (2n)^3 - (2n-1)^3 - n_e^3] + 2[n_0^2 - (2n)^2\\ &\quad - (2n-1)^2 - n_e^2 - 4n_0]t_s^*\\ &= 2[n_0^3 - (2n)^3 - (2n-1)^3 - n_e^3] + 2[8n(n - \frac{9}{4})\\ &\quad + 2n(n_e-1) + 6n_e(n-1) + 4]t_s^*. \end{aligned} \quad (66)$$

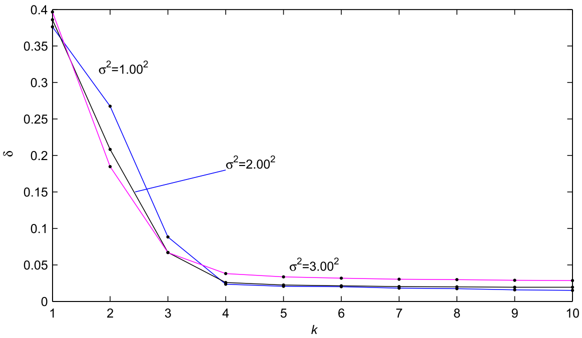

FIGURE 1. The AM-LSI estimation errors δ versus k with different σ^2.

Obviously, the value of $N_1 - N_2$ is positive when the system order n is large, that is, the H-AM-LSI algorithm has a higher computational efficiency than the AM-LSI algorithm for the large scale bilinear systems.

V. EXAMPLE

Consider a bilinear system:

$$\begin{aligned}&[A(z) + u(t-n)B(z)]y(t)\\ &\qquad = [C(z) + u(t-n)D(z)]u(t) + w(t),\\ E(z)w(t) &= v(t),\\ A(z) &= 1 + a_1z^{-1} + a_2z^{-2} = 1 + 0.60z^{-1} + 0.55z^{-2},\\ B(z) &= b_1z^{-1} + b_2z^{-2} = 0.20z^{-1} - 0.25z^{-2},\\ C(z) &= c_1z^{-1} + c_2z^{-2} = 0.10z^{-1} - 2.40z^{-2},\\ D(z) &= d_2z^{-2} = 0.11z^{-2},\\ E(z) &= 1 + e_1z^{-1} + e_2z^{-2} = 1 + 0.15z^{-1} - 0.10z^{-2}.\end{aligned}$$

The parameter vector to be estimated is given by

$$\begin{aligned}\boldsymbol{\theta} &= [a_1, a_2, b_1, b_2, c_1, c_2, d_2, e_1, e_2]^{\mathrm{T}}\\ &= [0.60, 0.55, 0.20, -0.25, 0.10,\\ &\qquad -2.40, 0.11, 0.15, -0.10]^{\mathrm{T}}.\end{aligned}$$

The example is simulated with MATLAB software. In the simulation, the input $\{u(t)\}$ is taken as a persistent excitation sequence with zero mean and unit variance, $\{v(t)\}$ is taken as an uncorrelated sequence noise with zero mean and variance $\sigma^2 = 1.00^2$, $\sigma^2 = 2.00^2$ and $\sigma^2 = 3.00^2$, respectively. Applying the AM-LSI algorithm and the H-AM-LSI algorithm to estimate the parameters of this system, the parameters and the estimation errors with different noise variances are shown in Tables 3–4. The estimation errors δ versus k are shown in Figures 1–2, where $\delta := \|\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$.

From the simulation results in Tables 3–4 and Figures 1–2, we can draw the following conclusions.

- The H-AM-LSI algorithm generates smaller estimation errors than the AM-LSI algorithm – see

TABLE 3. The AM-LSI estimates and errors with different σ^2.

σ^2	k	a_1	a_2	b_1	b_2	c_1	c_2	d_2	e_1	e_2	δ (%)
1.00^2	1	-0.08538	-0.02193	-0.00365	-0.00937	0.13542	-2.49258	0.10832	0.04823	0.02283	37.62630
	2	0.57369	0.47411	0.20231	-0.37222	0.07338	-2.43799	0.17370	0.50655	0.46209	26.74622
	3	0.56542	0.52317	0.19279	-0.23209	0.10649	-2.41116	0.13767	0.25709	0.09131	8.82880
	4	0.62053	0.56706	0.20154	-0.27635	0.09278	-2.39837	0.13760	0.15738	-0.06341	2.34517
	5	0.56608	0.52933	0.19849	-0.23497	0.10444	-2.40395	0.13045	0.14406	-0.07782	2.05783
	6	0.61874	0.56248	0.19969	-0.27513	0.09282	-2.39824	0.13482	0.15058	-0.07124	2.00173
	7	0.56935	0.53414	0.20016	-0.23690	0.10425	-2.40359	0.12974	0.14132	-0.08309	1.80526
	8	0.61559	0.55907	0.19877	-0.27260	0.09341	-2.39802	0.13409	0.14598	-0.07764	1.73556
	9	0.57331	0.53804	0.20102	-0.23961	0.10373	-2.40305	0.12975	0.13848	-0.08893	1.57522
	10	0.61208	0.55645	0.19841	-0.26972	0.09425	-2.39801	0.13350	0.14191	-0.08419	1.50150
2.00^2	1	-0.12186	-0.02864	-0.00010	-0.03187	0.13788	-2.49442	0.12439	0.03461	0.01596	38.60992
	2	0.55052	0.35767	0.21705	-0.38136	0.06891	-2.45475	0.20784	0.36704	0.30980	20.83298
	3	0.55344	0.53511	0.20389	-0.22779	0.10887	-2.41583	0.16841	0.22010	0.03428	6.70556
	4	0.61444	0.55277	0.19366	-0.28356	0.09008	-2.39699	0.16134	0.13978	-0.08691	2.58648
	5	0.56533	0.54105	0.20430	-0.24173	0.10509	-2.40316	0.15084	0.13455	-0.09798	2.24498
	6	0.60649	0.55152	0.19561	-0.27594	0.09289	-2.39671	0.15454	0.13538	-0.09549	2.14089
	7	0.57297	0.54417	0.20321	-0.24785	0.10301	-2.40175	0.15013	0.13308	-0.10117	2.02174
	8	0.60058	0.55016	0.19681	-0.27086	0.09471	-2.39739	0.15330	0.13384	-0.09882	1.99471
	9	0.57827	0.54606	0.20233	-0.25213	0.10158	-2.40090	0.15055	0.13253	-0.10242	1.93212
	10	0.59636	0.54914	0.19767	-0.26729	0.09600	-2.39793	0.15272	0.13307	-0.10057	1.92993
3.00^2	1	-0.15833	-0.03535	0.00345	-0.05437	0.14034	-2.49625	0.14046	0.02099	0.00908	39.66098
	2	0.52841	0.26560	0.23067	-0.38758	0.07217	-2.46373	0.23397	0.27596	0.18689	18.46176
	3	0.53050	0.52115	0.19881	-0.21931	0.11188	-2.42203	0.19740	0.20131	0.00914	6.68644
	4	0.62037	0.55782	0.19216	-0.30079	0.08643	-2.39775	0.18707	0.13993	-0.08808	3.79790
	5	0.54679	0.52981	0.20343	-0.23645	0.10765	-2.40363	0.17055	0.13439	-0.09771	3.35829
	6	0.61124	0.55743	0.19503	-0.29061	0.08956	-2.39583	0.17680	0.13601	-0.09549	3.18137
	7	0.55623	0.53449	0.20238	-0.24409	0.10498	-2.40225	0.16937	0.13305	-0.10051	3.02968
	8	0.60435	0.55571	0.19644	-0.28430	0.09171	-2.39647	0.17511	0.13461	-0.09825	2.96906
	9	0.56274	0.53761	0.20160	-0.24937	0.10325	-2.40133	0.16992	0.13250	-0.10174	2.87993
	10	0.59935	0.55419	0.19735	-0.27984	0.09324	-2.39697	0.17435	0.13383	-0.09979	2.85764
True values		0.60000	0.55000	0.20000	-0.25000	0.10000	-2.40000	0.11000	0.15000	-0.10000	

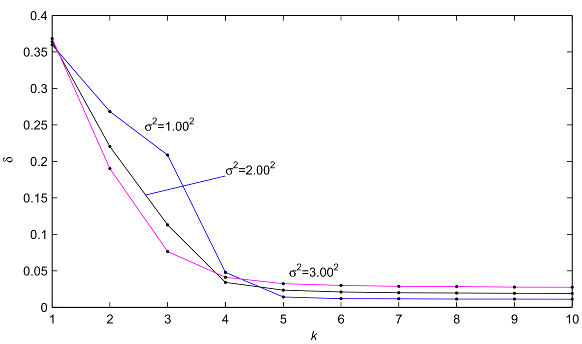

FIGURE 2. The H-AM-LSI estimation errors δ versus k with different σ^2.

Tables 3 – 4. In other words, the parameter estimates accuracy of the H-AM-LSI algorithm is higher than the AM-LSI algorithm.

- The parameter estimation errors of both the AM-LSI algorithm and the H-AM-LSI algorithm become small as the noise variance decreases – see Figures 1–2.
- The parameter estimation errors given by the AM-LSI algorithm and the H-AM-LSI algorithm become smaller with increasing iteration k – see Tables 3–4.

In order to validate the model, we use the AM-LSI estimates and the H-AM-LSI estimates to construct the estimated model, respectively, the predicted outputs are

$$\begin{aligned}\hat{y}_i(t_s) &= y(t_s) - v_i(t_s)\\ &= y(t_s) - \hat{E}_i(z)y(t_s) + \hat{E}_i(z)\hat{\alpha}_i(t_s),\end{aligned} \tag{67}$$

where $\hat{\alpha}_i(t_s) := \frac{\hat{C}_i(z)+u(t_s-n)\hat{D}_i(z)}{\hat{A}_i(z)+u(t_s-n)\hat{B}_i(z)}u(t_s)$, that is, $\hat{\alpha}_i(t_s) = [\hat{A}_i(z)-1+u(t_s-n)\hat{B}_i(z)]\hat{\alpha}_i(t_s) + [\hat{C}_i(z)+u(t_s-n)\hat{D}_i(z)]u(t_s)$.

Define $\hat{y}_{i,f}(t_s) := \hat{E}_i(z)y(t_s)$ and $\hat{\alpha}_{i,f}(t_s) := \hat{E}_i(z)\hat{\alpha}_i(t_s)$, then (67) can be expressed as

$$\hat{y}_i(t_s) = y(t_s) - \hat{y}_{i,f}(t_s) + \hat{\alpha}_{i,f}(t_s). \tag{68}$$

Using the AM-LSI parameter estimates in Table 3 with the noise variance $\sigma^2 = 1.00^2$ and $k = 10$ to construct the AM-LSI estimated model

$$\begin{aligned}\hat{y}_1(t_s) &= y(t_s) - \hat{y}_{1,f}(t_s) + \hat{\alpha}_{1,f}(t_s),\\ \hat{y}_{1,f}(t_s) &= (1+0.14191z^{-1} - 0.08416z^{-2})y(t_s),\\ \hat{\alpha}_{1,f}(t_s) &= (1+0.14191z^{-1} - 0.08416z^{-2})\hat{\alpha}_1(t_s),\\ \hat{A}_1(z) &= 1+0.61208z^{-1}+0.55645z^{-2},\\ \hat{B}_1(z) &= 0.19841z^{-1} - 0.26972z^{-2},\\ \hat{C}_1(z) &= 0.09425z^{-1} - 2.39801z^{-2},\\ \hat{D}_1(z) &= 0.13350z^{-2},\\ \hat{\alpha}_1(t_s) &= [\hat{A}_1(z) - 1 + u(t_s-n)\hat{B}_1(z)]\hat{\alpha}_1(t_s)\\ &\quad + [\hat{C}_1(z) + u(t_s-n)\hat{D}_1(z)]u(t_s).\end{aligned}$$

Using the H-AM-LSI parameter estimates in Table 4 with the noise variance $\sigma^2 = 1.00^2$ and $k = 10$ to construct the H-AM-LSI estimated model

$$\begin{aligned}\hat{y}_2(t_s) &= y(t_s) - \hat{y}_{2,f}(t_s) + \hat{\alpha}_{2,f}(t_s),\\ \hat{y}_{2,f}(t_s) &= (1+0.13271z^{-1} - 0.10039z^{-2})y(t_s),\\ \hat{\alpha}_{2,f}(t_s) &= (1+0.13271z^{-1} - 0.10039z^{-2})\hat{\alpha}_2(t_s),\end{aligned}$$

TABLE 4. The H-AM-LSI estimates and errors with different σ^2.

σ^2	k	a_1	a_2	b_1	b_2	c_1	c_2	d_2	e_1	e_2	δ (%)
1.00^2	1	-0.02364	-0.03492	-0.04049	-0.09402	0.13711	-2.48814	0.11066	0.03974	0.02965	35.97088
	2	0.55408	0.21759	0.21026	-0.37911	0.08579	-2.43230	0.17337	0.38613	0.43272	26.84590
	3	0.57158	0.56596	0.20478	-0.18428	0.09997	-2.41412	0.15211	0.43214	0.34656	20.86327
	4	0.60026	0.54644	0.19393	-0.27605	0.07941	-2.41654	0.15081	0.18481	0.00363	4.78122
	5	0.58185	0.54350	0.20220	-0.24617	0.10099	-2.41095	0.13432	0.13771	-0.09233	1.41276
	6	0.60216	0.55009	0.19805	-0.26231	0.09666	-2.39899	0.13280	0.13614	-0.09446	1.17924
	7	0.58733	0.54790	0.20175	-0.24968	0.10135	-2.40087	0.13040	0.13292	-0.10109	1.15640
	8	0.59950	0.54959	0.19851	-0.25974	0.09782	-2.39836	0.13132	0.13342	-0.09877	1.12719
	9	0.58997	0.54886	0.20132	-0.25175	0.10075	-2.40017	0.13028	0.13242	-0.10212	1.12443
	10	0.59735	0.54902	0.19890	-0.25800	0.09843	-2.39862	0.13100	0.13271	-0.10039	1.11472
2.00^2	1	-0.02318	-0.04759	-0.03446	-0.09541	0.14039	-2.49008	0.12714	0.02645	0.04740	36.39142
	2	0.53367	0.21193	0.21928	-0.37916	0.08847	-2.43839	0.19907	0.31274	0.28420	22.03206
	3	0.55148	0.54800	0.20143	-0.19948	0.09861	-2.42652	0.18495	0.27982	0.13624	11.29236
	4	0.60506	0.54661	0.19375	-0.28740	0.07688	-2.41832	0.17649	0.14677	-0.07009	3.41450
	5	0.56599	0.53875	0.20345	-0.24334	0.10397	-2.41009	0.15404	0.13501	-0.09731	2.34824
	6	0.60460	0.55157	0.19618	-0.27457	0.09380	-2.39703	0.15448	0.13544	-0.09582	2.09850
	7	0.57447	0.54443	0.20290	-0.24890	0.10272	-2.40135	0.15031	0.13281	-0.10117	1.99674
	8	0.59944	0.55037	0.19719	-0.26981	0.09555	-2.39707	0.15258	0.13359	-0.09902	1.95411
	9	0.57927	0.54632	0.20214	-0.25280	0.10154	-2.40040	0.15036	0.13236	-0.10218	1.91100
	10	0.59559	0.54930	0.19792	-0.26657	0.09668	-2.39753	0.15202	0.13290	-0.10050	1.89950
3.00^2	1	-0.02271	-0.06026	-0.02842	-0.09681	0.14368	-2.49202	0.14362	0.01316	0.06515	36.84747
	2	0.51293	0.20652	0.22826	-0.37843	0.09178	-2.44547	0.22513	0.25443	0.16236	19.01832
	3	0.53584	0.53299	0.19798	-0.21407	0.10059	-2.43077	0.21124	0.21342	0.03195	7.63231
	4	0.60702	0.54949	0.19330	-0.29711	0.07799	-2.41551	0.19828	0.13974	-0.08672	4.11001
	5	0.55390	0.53344	0.20323	-0.24262	0.10539	-2.40789	0.17368	0.13383	-0.09921	3.23204
	6	0.60448	0.55444	0.19557	-0.28481	0.09183	-2.39618	0.17578	0.13488	-0.09743	2.99824
	7	0.56316	0.53914	0.20250	-0.24920	0.10347	-2.40128	0.17011	0.13252	-0.10145	2.86936
	8	0.59862	0.55306	0.19683	-0.27924	0.09376	-2.39639	0.17369	0.13365	-0.09959	2.82529
	9	0.56862	0.54156	0.20169	-0.25374	0.10206	-2.40034	0.17028	0.13220	-0.10216	2.76552
	10	0.59434	0.55174	0.19765	-0.27547	0.09504	-2.39684	0.17304	0.13311	-0.10066	2.75219
True values		0.60000	0.55000	0.20000	-0.25000	0.10000	-2.40000	0.11000	0.15000	-0.10000	

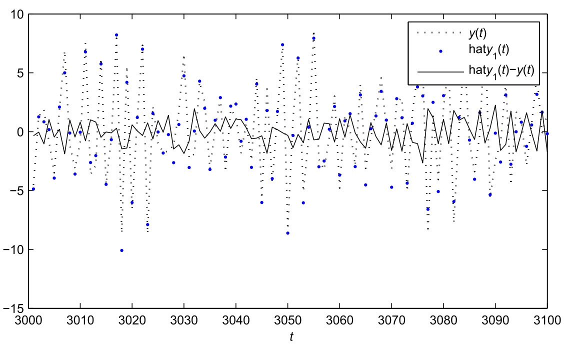

FIGURE 3. The predicted outputs $\hat{y}_1(t)$, actual outputs $y(t)$ and their errors $\hat{y}_1(t) - y(t)$ versus t based on the AM-LSI estimates.

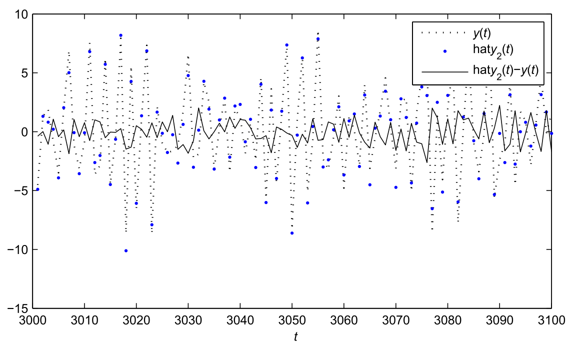

FIGURE 4. The predicted outputs $\hat{y}_2(t)$, actual outputs $y(t)$ and their errors $\hat{y}_2(t) - y(t)$ versus t based on the H-AM-LSI estimates.

$$
\begin{aligned}
\hat{A}_2(z) &= 1 + 0.59735z^{-1} + 0.54902z^{-2},\\
\hat{B}_2(z) &= 0.19890z^{-1} - 0.25800z^{-2},\\
\hat{C}_2(z) &= 0.09843z^{-1} - 2.39862z^{-2},\\
\hat{D}_2(z) &= 0.13100z^{-2},\\
\hat{\alpha}_2(t_s) &= [\hat{A}_2(z) - 1 + u(t_s - n)\hat{B}_2(z)]\hat{\alpha}_2(t_s)\\
&\quad + [\hat{C}_2(z) + u(t_s - n)\hat{D}_2(z)]u(t_s).
\end{aligned}
$$

In order to validate these estimated models, we use the rest $t_r = 100$ data from $t = t_s^* + 1$ to $t = t_s^* + t_r$ to compute the predicted outputs $\hat{y}_i(t)$ of the system. Their predicted outputs $\hat{y}_i(t)$, actual outputs $y(t)$ and their errors $\hat{y}_i(t) - y(t)$ are plotted in Figures 3–4 for the AM-LSI and H-AM-LSI algorithms. Using the estimated outputs to compute the root-mean-square errors (RMSEs)

$$
Error(\hat{y}_1) := \left[\frac{1}{t_r}\sum_{j=t_s^*+1}^{t_s^*+t_r}[\hat{y}_1(t) - y(t)]^2\right]^{1/2} = 1.07019,
$$

$$
Error(\hat{y}_2) := \left[\frac{1}{t_r}\sum_{j=t_s^*+1}^{t_s^*+t_r}[\hat{y}_2(t) - y(t)]^2\right]^{1/2} = 1.06931.
$$

From Figures 3–4, we can see that the predicted outputs $\hat{y}_1(t)$ and $\hat{y}_2(t)$ are very close to the true outputs $y(t)$, and the RMSEs of the two algorithms are very close to the noise standard deviation $\sigma = 1.00$. In other words, the estimated model can capture the dynamics of the system.

VI. CONCLUSIONS

This work mainly presents an auxiliary model based least squares iterative (AM-LSI) algorithm and a hierarchical auxiliary model based least squares iterative (H-AM-LSI) algorithm for the bilinear systems with an AR process by using interval-varying input-output data. Compared with the AM-LSI algorithm, the H-AM-LSI algorithm has a lower computational burden and a higher parameter estimation accuracy. The simulation results shown that all the proposed algorithms can identify bilinear systems well and can give accurate parameter estimates for bilinear systems. However, the convergence analyses of the proposed method is very difficult and worth further studying. The proposed methods can be extended to other linear nonlinear systems and nonlinear systems with different structure and colored noise, and can be applied to other fields.

ACKNOWLEDGMENTS

The authors are grateful to Professor Feng Ding at the Jiangnan University for his helpful suggestions and the main idea of this work comes from him and his book "*System Identification – Auxiliary Model Identification Idea and Methods*, Science Press, Beijing, 2017".

REFERENCES

[1] F. Ding, F. F. Wang, L. Xu, T. Hayat, and A. Alsaedi, "Parameter estimation for pseudo-linear systems using the auxiliary model and the decomposition technique," *IET Control Theory Appl.*, vol. 11, no. 3, pp. 390–400, Feb. 2017.

[2] F. Ding, F. Wang, L. Xu, and M. Wu, "Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering," *J. Franklin Inst.*, vol. 354, no. 3, pp. 1321–1339, Feb. 2017.

[3] F. Ding, X. Wang, L. Mao, and L. Xu, "Joint state and multi-innovation parameter estimation for time-delay linear systems and its convergence based on the Kalman filtering," *Digit. Signal Process.*, vol. 62, pp. 211–223, Mar. 2017.

[4] F. Ding and X. Wang, "Hierarchical stochastic gradient algorithm and its performance analysis for a class of bilinear-in-parameter systems," *Circuits Syst. Signal Process.*, vol. 36, no. 4, pp. 1393–1405, Apr. 2017.

[5] D.-Q. Wang, Z. Zhang, and J.-Y. Yuan, "Maximum likelihood estimation method for dual-rate Hammerstein systems," *Int. J. Control Autom. Syst.*, vol. 15, no. 2, pp. 698–705, Apr. 2017.

[6] D. Wang, L. Mao, and F. Ding, "Recasted models-based hierarchical extended stochastic gradient method for MIMO nonlinear systems," *IET Control Theory Appl.*, vol. 11, no. 4, pp. 476–485, Feb. 2017.

[7] D. Wang and F. Ding, "Parameter estimation algorithms for multivariable Hammerstein CARMA systems," *Inf. Sci.*, vols. 355–356, pp. 237–248, Aug. 2016.

[8] D. Wang, "Hierarchical parameter estimation for a class of MIMO Hammerstein systems based on the reframed models," *Appl. Math. Lett.*, vol. 57, pp. 13–19, Jul. 2016.

[9] S. Peng, C. Chen, H. Shi, and Z. Yao, "State of charge estimation of battery energy storage systems based on adaptive unscented Kalman filter with a noise statistics estimator," *IEEE Access*, vol. 5, pp. 13202–13212, 2017.

[10] L. Wan, G. Han, L. Shu, N. Feng, C. Zhu, and J. Lloret, "Distributed parameter estimation for mobile wireless sensor network based on cloud computing in battlefield surveillance system," *IEEE Access*, vol. 3, pp. 1729–1739, Oct. 2015.

[11] D. Westwick and M. Verhaegen, "Identifying MIMO Wiener systems using subspace model identification methods," *Signal Process.*, vol. 52, no. 2, pp. 235–258, Jul. 1996.

[12] V. Verdult and M. Verhaegen, "Subspace identification of multivariable linear parameter-varying systems," *Automatica*, vol. 38, no. 5, pp. 805–814, May 2002.

[13] Y.-X. Zheng and Y. Liao, "Parameter identification of nonlinear dynamic systems using an improved particle swarm optimization," *Optik-Int. J. Light Electron Opt.*, vol. 127, no. 19, pp. 7865–7874, Oct. 2016.

[14] S.-H. Tsai and Y.-W. Chen, "A novel fuzzy identification method based on ant colony optimization algorithm," *IEEE Access*, vol. 4, pp. 3747–3756, 2016.

[15] P. T. Brewick and S. F. Masri, "An evaluation of data-driven identification strategies for complex nonlinear dynamic systems," *Nonlinear Dyn.*, vol. 85, no. 2, pp. 1297–1318, Jul. 2016.

[16] W. Xiong, X. Yang, L. Ke, and B. Xu, "EM algorithm-based identification of a class of nonlinear Wiener systems with missing output data," *Nonlinear Dyn.*, vol. 80, nos. 1–2, pp. 329–339, Apr. 2015.

[17] Y. Mao, F. Ding, A. Alsaedi, and T. Hayat, "Adaptive filtering parameter estimation algorithms for Hammerstein nonlinear systems," *Signal Process.*, vol. 128, pp. 417–425, Nov. 2016.

[18] X. Wang and F. Ding, "Recursive parameter and state estimation for an input nonlinear state space system using the hierarchical identification principle," *Signal Process.*, vol. 117, pp. 208–218, Dec. 2015.

[19] J. C. Ralston and B. Boashash, "Identification of bilinear systems using bandlimited regression," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5. Apr. 1997, pp. 3925–3928.

[20] R. Mohler and W. Kolodziej, "An overview of stochastic bilinear control processes," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-10, no. 12, pp. 913–918, Dec. 1980.

[21] R. Mohler and W. Kolodziej, "An overview of bilinear system theory and applications," *IEEE Trans. Syst., Man, Cybern.*, vol. 10, no. 10, pp. 683–688, Oct. 1980.

[22] Y. Yu and K.-H. Jo, "Output feedback fault-tolerant control for a class of discrete-time fuzzy bilinear systems," *Int. J. Control, Autom. Syst.*, vol. 14, no. 2, pp. 486–494, Apr. 2016.

[23] Y. Marrero-Ponce, E. Contreras-Torres, C. R. García-Jacas, S. J. Barigye, N. Cubillán, and Y. J. Alvarado, "Novel 3D bio-macromolecular bilinear descriptors for protein science: Predicting protein structural classes," *J. Theor. Biol.*, vol. 374, pp. 125–137, Jun. 2015.

[24] I. D. Christoskov and P. T. Petkov, "A practical procedure of bilinear weighted core kinetics parameters computation for the purpose of experimental reactivity determination," *Ann. Nucl. Energy*, vol. 29, no. 9, pp. 1041–1054, Jun. 2002.

[25] P. N. Paraskevopoulos and K. G. Arvanitis, "A new orthogonal series approach to state space analysis of bilinear systems," *IEEE Trans. Autom. Control*, vol. 39, no. 4, pp. 793–797, Apr. 1994.

[26] V. R. Karanam, P. A. Frick, and R. R. Mohler, "Bilinear system identification by Walsh functions," *IEEE Trans. Autom. Control*, vol. 23, no. 4, pp. 709–713, Aug. 1978.

[27] B. Cheng and N.-S. Hsu, "Analysis and parameter estimation of bilinear systems via block-pulse functions," *Int. J. Control*, vol. 36, no. 1, pp. 53–65, Jul. 1982.

[28] C.-C. Liu and Y.-P. Shih, "Analysis and parameter estimation of bilinear systems via Chebyshev polynomials," *J. Franklin Inst.*, vol. 317, no. 6, pp. 373–382, Jun. 1984.

[29] C. Hwang and M. Y. Chen, "Analysis and parameter identification of bilinear systems via shifted Legendre polynomials," *Int. J. Control*, vol. 44, no. 2, pp. 351–362, 1986.

[30] H.-Y. Chung and Y.-Y. Sun, "Analysis and parameter estimation of bilinear systems using Taylor operational matrices," *IEEE Trans. Syst., Man, Cybern.*, vol. 17, no. 6, pp. 1068–1071, Nov. 1987.

[31] C. Hwang and M. Y. Chen, "Parameter identification of bilinear systems using the Galerkin method," *Int. J. Syst. Sci.*, vol. 16, no. 5, pp. 641–648, May 1985.

[32] S. Daniel-Berhe and H. Unbehauen, "Bilinear continuous-time systems identification via Hartley-based modulating functions," *Automatica*, vol. 34, no. 4, pp. 499–503, Apr. 1998.

[33] M. Inagaki and H. Mochizuki, "Bilinear system identification by Volterra kernels estimation," *IEEE Trans. Autom. Control*, vol. 29, no. 8, pp. 746–749, Aug. 1984.

[34] V. Tsoulkas, P. Koukoulas, and N. Kalouptsidis, "Identification of input-output bilinear systems using cumulants," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2753–2761, Nov. 2001.

[35] S. Gibson, A. Wills, and B. Ninness, "Maximum-likelihood parameter estimation of bilinear systems," *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1581–1596, Oct. 2005.

[36] M. Li, X. Liu, and F. Ding, "The maximum likelihood least squares based iterative estimation algorithm for bilinear systems with autoregressive moving average noise," *J. Franklin Inst.*, vol. 354, no. 12, pp. 4861–4881, Aug. 2017.

[37] D. Meng, "Recursive least squares and multi-innovation gradient estimation algorithms for bilinear stochastic systems," *Circuits Syst. Signal Process.*, vol. 36, no. 3, pp. 1052–1065, Mar. 2017.

[38] M. Li, X. Liu, and F. Ding, "The gradient-based iterative estimation algorithms for bilinear systems with autoregressive noise," *Circuits Syst. Signal Process.*, vol. 36, no. 11, pp. 4541–4568, Nov. 2017.

[39] M. Li, X. Liu, and F. Ding, "Least-squares-based iterative and gradient-based iterative estimation algorithms for bilinear systems," *Nonlinear Dyn.*, vol. 89, no. 1, pp. 197–211, Jul. 2017.

[40] H. Dai and N. K. Sinha, "Robust recursive least-squares method with modified weights for bilinear system identification," *IEE Proc. D, Control Theory Appl.*, vol. 136, no. 3, pp. 122–126, May 1989.

[41] M. A. Z. Raja and N. I. Chaudhary, "Two-stage fractional least mean square identification algorithm for parameter estimation of CARMA systems," *Signal Process.*, vol. 107, pp. 327–339, Feb. 2015.

[42] A. Janot, P. O. Vandanjon, and M. Gautier, "Identification of physical parameters and instrumental variables validation with two-stage least squares estimator," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1386–1393, Jul. 2013.

MEIHANG LI received the M.Sc. degree from the College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao, China, in 2014, where he is currently pursuing the Ph.D. degree. His current research interests include system identification and process control.

XIMEI LIU received the M.Sc. and Ph.D. degrees in automatic control from the East China University of Science and Technology, Shanghai, China, in 1990 and 2009, respectively. Since 1982, she has been with the College of Automation and Electronic Engineering, Qingdao University of Science and Technology, where she is currently a Professor. Her current research interests include intelligent control, fault detection, and system identification.