# Highly computational efficiency identification algorithm for multiple-input multiple-output systems

Haoming Xing[1], Feng Ding[1], and Feng Pan[1]

[1]Jiangnan University College of Internet of Things Engineering

May 16, 2023

**Abstract**

The identification of multiple-input multiple-output (MIMO) systems is an important part of designing complex control systems. This paper studies an auxiliary model least squares iterative (AM-LSI) algorithm for MIMO systems. With the expansion of the system scale and limitations of the computer resources, there is an urgent need for an identification algorithm that provides higher computational efficiency. To address this issue, this paper further derives a hierarchical identification model and proposes a new auxiliary model hierarchical least squares iterative (AM-HLSI) algorithm for MIMO systems by applying the hierarchical identification principle. Through the analysis of the computational efficiency, the AM-HLSI algorithm has higher computational efficiency than the AM-LSI algorithm. Additionally, the feasibility of the AM-LSI and AM-HLSI algorithms is validated by a simulation example.

INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
*Int. J. Robust Nonlinear Control* 2023; **00**:1–15
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/acs

# Highly computational efficiency identification algorithm for multiple-input multiple-output systems

Haoming Xing, Feng Ding*, Feng Pan

*Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China.*

SUMMARY

The identification of multiple-input multiple-output (MIMO) systems is an important part of designing complex control systems. This paper studies an auxiliary model least squares iterative (AM-LSI) algorithm for MIMO systems. With the expansion of the system scale and limitations of the computer resources, there is an urgent need for an identification algorithm that provides higher computational efficiency. To address this issue, this paper further derives a hierarchical identification model and proposes a new auxiliary model hierarchical least squares iterative (AM-HLSI) algorithm for MIMO systems by applying the hierarchical identification principle. Through the analysis of the computational efficiency, the AM-HLSI algorithm has higher computational efficiency than the AM-LSI algorithm. Additionally, the feasibility of the AM-LSI and AM-HLSI algorithms is validated by a simulation example.


Received ...

KEY WORDS: Computational efficiency, Hierarchical identification, Multivariable system, Parameter estimation, Least squares

## 1. INTRODUCTION

Mathematical models are crucial in designing actual control systems [1, 2]. The analytical method [3–5] and the experimental method (i.e., system identification) [6–8] are used to establish the mathematical model of control systems. The analytical method relies on the physical or chemical laws to analyze the internal mechanism of the system and establish the mathematical model. However, it is often challenging to obtain the internal structures of actual control systems. Therefore, system identification is more generally used to establish the mathematical models of control systems. System identification methods have been developed to solve the identification problems of the linear or nonlinear scalar systems [9, 10]. For switched linear systems, Louis et al. proposed a new identification algorithm based on the statistical learning theory and illustrated the accuracy of the proposed method [11]. Alex et al. quantified the input and output of finite-impulse response systems into the binary values, estimated the joint probabilities of the unquantized signals from the binary signals, and inferred the system parameters [12].

In recent years, the scale of control systems has become larger and larger, and multiple-input multiple-output (MIMO) systems have gradually replaced scalar systems [13, 14]. Compared with scalar systems, MIMO systems have higher orders and more parameters, which brings more challenges to establish the mathematical model of MIMO systems. For the identification of

*Correspondence to: School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China. https://orcid.org/0000-0002-2721-2025
†Email: hmxing12@126.com; fding@jiangnan.edu.cn; pan_feng_63@163.com

MIMO systems, many scholars have studied the relevant identification algorithms [15–17]. For output-error linear MIMO models, Formentin et al. developed a novel theoretical framework for the control-oriented identification based on a Bayesian perspective on modeling and derived a Bayesian robust control design approach [18]. Cerone et al. studied the structured discrete-time nonlinear systems and proposed a single-stage set-membership identification algorithm to solve the problem in the context of the set-membership errors-in-variables identification [19]. For multivariate control systems with colored noise, Ma et al. proposed a filtering-based recursive generalized extended least squares algorithm and used the multi-innovation identification theory to improve the parameter estimation accuracy [20]. For multivariable equation-error systems, Xia et al. presented a maximum likelihood least squares-based iterative identification approach by combining the iterative identification technique with the maximum likelihood principle [21]. The algorithms mentioned above can effectively solve the identification problems of MIMO systems. However, there are still some issues in identifying MIMO systems, such as the heavy calculation burden and the low identification accuracy.

In order to reduce the calculation burden and improve the identification accuracy, this paper applies the hierarchical identification principle to identify MIMO systems [22, 23]. The hierarchical identification principle is a significant method to simplify the system structure and improve the identification calculation efficiency [24]. The operation of the hierarchical identification principle is to decompose the complex system into several subsystems and establish an identification model for each subsystem. There are frequently some associate items between these subsystems, so the coordination of these associate items needs to be considered during the identification process. For bilinear state-space systems with colored noise, Shahriari et al. studied a four-stage recursive least squares algorithm and a four-stage stochastic gradient algorithm by using the hierarchical identification principle [25]. Chaudhary et al. gave a fractional hierarchical gradient descent method by combining the hierarchical identification principle with the gradient descent method, which solved a nonlinear system identification problem [26]. Liu et al. considered the parameter estimation problems of two-input single-output Hammerstein output-error moving average systems, decomposed the system into two subsystems based on the hierarchical identification principle, and presented a hierarchical least squares algorithm [27].

The auxiliary model identification idea has been widely applied in system identification. Despite the fact that MIMO systems contain a large amount of measurement information, there are still some unmeasurable variables within the system. Therefore, it is necessary to establish auxiliary models to obtain estimated values for these unmeasurable variables. For dual-rate Hammerstein-Volterra systems, Zong et al. solved the problem of the incomplete identification data caused by the dual-rate sampling through using the auxiliary model method and proposed an auxiliary model-based hybrid particle swarm-gradient algorithm [28]. Chaudhary et al. studied an auxiliary model-based normalized variable initial value fractional least mean square algorithm for input nonlinear output-error system by using the auxiliary model identification idea [29].

This paper aims to propose a new highly computational efficiency algorithm for MIMO systems. The main contributions of this paper lie in the following aspects.

- Through using the auxiliary model for estimating the unknown fictitious output, we give an auxiliary model least squares iterative (AM-LSI) algorithm for identifying the MIMO systems.
- To improve the computational efficiency and the identification accuracy, an auxiliary model hierarchical least squares iterative (AM-HLSI) algorithm is further proposed by using the hierarchical identification principle.
- Moreover, the computational efficiency of the AM-LSI and AM-HLSI algorithms is analyzed in detail. The performance of the AM-LSI and AM-HLSI algorithms is tested in the numerical simulation example.

The organization of this paper is as follows. Section 2 defines some symbols and derives an identification model of the MIMO systems. In Section 3, the AM-LSI algorithm is presented by applying the auxiliary model identification idea. Section 4 transforms the identification model in



DOI: 10.1002/acs

Section 2 into a hierarchical identification model and further proposes a highly computational efficiency AM-HLSI algorithm. For comparison, Section 5 analyzes the computational efficiency of the AM-LSI and AM-HLSI algorithms. Moreover, the simulation example of these algorithms is given in Section 6. Finally, Section 7 gives some concluding remarks.

## 2. SYSTEM DESCRIPTION AND IDENTIFICATION MODELS

Let us introduce some notations first. '$X =: Y$' or '$Y := X$' represents '$Y$ is defined as $X$'; the norm of a matrix $\boldsymbol{X}$ is defined by $\|\boldsymbol{X}\| := \sqrt{\mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]}$; the symbol $\boldsymbol{I}_n$ denotes an identity matrix of size $n \times n$; the symbol $\mathbf{1}_n$ denotes a $n$-dimensional column vector whose elements are 1; the superscript T denotes the vector or matrix transpose; the symbol $z^{-1}$ represents the unit backward shift operator: $z^{-1}y(t) := y(t-1)$.

Consider the MIMO system described by the following output-error model,

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix} = \begin{bmatrix} \frac{B_{11}(z)}{A_{11}(z)} & \frac{B_{12}(z)}{A_{12}(z)} & \cdots & \frac{B_{1r}(z)}{A_{1r}(z)} \\ \frac{B_{21}(z)}{A_{21}(z)} & \frac{B_{22}(z)}{A_{22}(z)} & \cdots & \frac{B_{2r}(z)}{A_{2r}(z)} \\ \vdots & \vdots & & \vdots \\ \frac{B_{m1}(z)}{A_{m1}(z)} & \frac{B_{m2}(z)}{A_{m2}(z)} & \cdots & \frac{B_{mr}(z)}{A_{mr}(z)} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_r(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_m(t) \end{bmatrix}, \tag{1}$$

where the dimensions $m$ and $r$ are known, $y_i(t)$, $i = 1, 2, \cdots, m$, is the $i$th output of the system, $u_j(t)$, $j = 1, 2, \cdots, r$, is the $j$th input of the system, $v_i(t)$ is the white noise with zero mean, $A_{ij}(z)$ is the denominator of the transfer function from the $j$th input to the $i$th output, $B_{ij}(z)$ is the numerator of the transfer function from the $j$th input to the $i$th output. The polynomials $A_{ij}(z)$ and $B_{ij}(z)$ are defined as

$$A_{ij}(z) := 1 + a_{ij}(1)z^{-1} + a_{ij}(2)z^{-2} + \cdots + a_{ij}(n_j)z^{-n_j}, \tag{2}$$
$$B_{ij}(z) := b_{ij}(1)z^{-1} + b_{ij}(2)z^{-2} + \cdots + b_{ij}(n_j)z^{-n_j}. \tag{3}$$

Assume that the order $n_j$ is known. $a_{ij}(1), a_{ij}(2), \cdots, a_{ij}(n_j), b_{ij}(1), b_{ij}(2), \cdots, b_{ij}(n_j)$ are the unknown parameters to be estimated from the available input-output data $\{u_j(t), y_i(t)\}$. It is difficult to directly identify this MIMO system. One effective method is to decompose the MIMO system into $m$ multiple-input single-output subsystems,

$$\begin{aligned} y_i(t) &:= \frac{B_{i1}(z)}{A_{i1}(z)}u_1(t) + \frac{B_{i2}(z)}{A_{i2}(z)}u_2(t) + \cdots + \frac{B_{ir}(z)}{A_{ir}(z)}u_r(t) + v_i(t) \\ &= \sum_{j=1}^{r} \frac{B_{ij}(z)}{A_{ij}(z)}u_j(t) + v_i(t). \end{aligned} \tag{4}$$

Introduce the fictitious output:

$$x_{ij}(t) := \frac{B_{ij}(z)}{A_{ij}(z)}u_j(t). \tag{5}$$

Substituting (2)–(3) into (5), we have

$$\begin{aligned} x_{ij}(t) &= -[a_{ij}(1)z^{-1} + a_{ij}(2)z^{-2} + \cdots + a_{ij}(n_j)z^{-n_j}]x_{ij}(t) \\ &\quad +[b_{ij}(1)z^{-1} + b_{ij}(2)z^{-2} + \cdots + b_{ij}(n_j)z^{-n_j}]u_j(t) \\ &= -a_{ij}(1)x_{ij}(t-1) - a_{ij}(2)x_{ij}(t-2) - \cdots - a_{ij}(n_j)x_{ij}(t-n_j) \\ &\quad +b_{ij}(1)u_j(t-1) + b_{ij}(2)u_j(t-2) + \cdots + b_{ij}(n_j)u_j(t-n_j) \\ &= \boldsymbol{\varphi}_{ij}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_{ij}, \end{aligned} \tag{6}$$

where the sub-parameter vector $\boldsymbol{\vartheta}_{ij}$ and the sub-information vector $\boldsymbol{\varphi}_{ij}(t)$ are defined as

$$\boldsymbol{\vartheta}_{ij} := [a_{ij}(1), a_{ij}(2), \cdots, a_{ij}(n_j), b_{ij}(1), b_{ij}(2), \cdots, b_{ij}(n_j)]^{\mathrm{T}} \in \mathbb{R}^{2n_j}, \tag{7}$$



DOI: 10.1002/acs

$$\boldsymbol{\varphi}_{ij}(t) := [-x_{ij}(t-1), \cdots, -x_{ij}(t-n_j), u_j(t-1), \cdots, u_j(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{2n_j}. \tag{8}$$

Thus, Equation (4) can be expressed as

$$\begin{aligned} y_i(t) &= \sum_{j=1}^{r} x_{ij}(t) + v_i(t) \\ &= \boldsymbol{\varphi}_{i1}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_{i1} + \boldsymbol{\varphi}_{i2}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_{i2} + \cdots + \boldsymbol{\varphi}_{ir}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_{ir} + v_i(t) \\ &= \boldsymbol{\varphi}_{i}^{\mathrm{T}}(t)\boldsymbol{\vartheta}_{i} + v_i(t), \end{aligned} \tag{9}$$

where the parameter vector $\boldsymbol{\vartheta}_i$ and the information vector $\boldsymbol{\varphi}_i(t)$ are defined as

$$\boldsymbol{\vartheta}_i := [\boldsymbol{\vartheta}_{i1}^{\mathrm{T}}, \boldsymbol{\vartheta}_{i2}^{\mathrm{T}}, \cdots, \boldsymbol{\vartheta}_{ir}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n},\ n := n_1 + n_2 + \cdots + n_r, \tag{10}$$

$$\boldsymbol{\varphi}_i(t) := [\boldsymbol{\varphi}_{i1}^{\mathrm{T}}(t), \boldsymbol{\varphi}_{i2}^{\mathrm{T}}(t), \cdots, \boldsymbol{\varphi}_{ir}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n}. \tag{11}$$

Equation (9) is the identification model of the MIMO system in (1). The proposed AM-LSI algorithm in the next section is based on this identification model. This identification model involves all the unknown parameters $a_{ij}(1), a_{ij}(2), \cdots, a_{ij}(n_j), b_{ij}(1), b_{ij}(2), \cdots, b_{ij}(n_j)$ of the MIMO system in (1).

## 3. THE AM-LSI ALGORITHM

In this section, we propose an AM-LSI algorithm based on the identification model in (9).

Consider the data from $t=1$ to $t=L$ and define the stacked information matrix $\boldsymbol{\Xi}_i(L)$ and the stacked output vector $\boldsymbol{Y}_i(L)$ as

$$\boldsymbol{\Xi}_i(L) := [\boldsymbol{\varphi}_i(1), \boldsymbol{\varphi}_i(2), \cdots, \boldsymbol{\varphi}_i(L)]^{\mathrm{T}} \in \mathbb{R}^{L\times 2n}, \tag{12}$$

$$\boldsymbol{Y}_i(L) := [y_i(1), y_i(2), \cdots, y_i(L)]^{\mathrm{T}} \in \mathbb{R}^{L}. \tag{13}$$

According to the identification model in (9), define the cost function:

$$\begin{aligned} J_1(\boldsymbol{\vartheta}_i) &:= \frac{1}{2}\sum_{l=1}^{L}[y_i(l) - \boldsymbol{\varphi}_i(l)\boldsymbol{\vartheta}_i]^2 \\ &= \frac{1}{2}\|\boldsymbol{Y}_i(L) - \boldsymbol{\Xi}_i(L)\boldsymbol{\vartheta}_i\|^2. \end{aligned} \tag{14}$$

Letting $\hat{\boldsymbol{\vartheta}}_{i,k}$ be the $k$th estimate of the parameter vector $\boldsymbol{\vartheta}_i$ and minimizing the cost function $J_1(\boldsymbol{\vartheta}_i)$, we have

$$\hat{\boldsymbol{\vartheta}}_{i,k} := [\boldsymbol{\Xi}_{i,k}^{\mathrm{T}}(L)\boldsymbol{\Xi}_{i,k}(L)]^{-1}\boldsymbol{\Xi}_{i,k}^{\mathrm{T}}(L)\boldsymbol{Y}_i(L). \tag{15}$$

It is noteworthy that the information matrix $\boldsymbol{\Xi}_{i,k}(L)$ in (15) is formed by the information vector $\boldsymbol{\varphi}_i(t)$ with the unknown fictitious output $x_{ij}(t)$. Therefore, Equation (15) cannot be implemented directly. The approach is to establish an auxiliary model

$$\hat{x}_{ij,k}(t) := \hat{\boldsymbol{\varphi}}_{ij,k}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}_{ij,k} \in \mathbb{R} \tag{16}$$

for each unknown fictitious output by using the auxiliary model identification idea. Then, form the estimates $\hat{\boldsymbol{\varphi}}_{ij,k}(t)$ of $\boldsymbol{\varphi}_{ij}(t)$, $\hat{\boldsymbol{\varphi}}_{i,k}(t)$ of $\boldsymbol{\varphi}_i(t)$ and $\hat{\boldsymbol{\Xi}}_{i,k}(L)$ of $\boldsymbol{\Xi}_i(L)$ as

$$\hat{\boldsymbol{\varphi}}_{ij,k}(t) := [-\hat{x}_{ij}(t-1), \cdots, -\hat{x}_{ij}(t-n_j), u_j(t-1), \cdots, u_j(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{2n_j}, \tag{17}$$

$$\hat{\boldsymbol{\varphi}}_{i,k}(t) := [\hat{\boldsymbol{\varphi}}_{i1}^{\mathrm{T}}(t), \hat{\boldsymbol{\varphi}}_{i2}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\varphi}}_{ir}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n}, \tag{18}$$

$$\hat{\boldsymbol{\Xi}}_{i,k}(L) := [\hat{\boldsymbol{\varphi}}_i(1), \hat{\boldsymbol{\varphi}}_i(2), \cdots, \hat{\boldsymbol{\varphi}}_i(L)]^{\mathrm{T}} \in \mathbb{R}^{L\times 2n}. \tag{19}$$

Replacing the information matrix $\boldsymbol{\Xi}_i(L)$ in (15) with its estimate $\hat{\boldsymbol{\Xi}}_{i,k}(L)$, we can obtain the AM-LSI algorithm for estimating $\boldsymbol{\vartheta}_i$ in (9):

$$\hat{\boldsymbol{\vartheta}}_{i,k} = [\hat{\boldsymbol{\Xi}}_{i,k}^{\mathrm{T}}(L)\hat{\boldsymbol{\Xi}}_{i,k}(L)]^{-1}\hat{\boldsymbol{\Xi}}_{i,k}^{\mathrm{T}}(L)\boldsymbol{Y}_i(L), \tag{20}$$



DOI: 10.1002/acs

$$\boldsymbol{Y}_i(L) = [y_i(1), y_i(2), \cdots, y_i(L)]^{\mathrm{T}}, \tag{21}$$
$$\hat{\boldsymbol{\Xi}}_i(L) = [\hat{\boldsymbol{\varphi}}_i(1), \hat{\boldsymbol{\varphi}}_i(2), \cdots, \hat{\boldsymbol{\varphi}}_i(L)]^{\mathrm{T}}, \tag{22}$$
$$\hat{\boldsymbol{\varphi}}_i(t) = [\hat{\boldsymbol{\varphi}}_{i1}^{\mathrm{T}}(t), \hat{\boldsymbol{\varphi}}_{i2}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\varphi}}_{ir}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{23}$$
$$\hat{\boldsymbol{\varphi}}_{ij}(t) = [-\hat{x}_{ij}(t-1), \cdots, -\hat{x}_{ij}(t-n_j), u_j(t-1), \cdots, u_j(t-n_j)]^{\mathrm{T}}, \tag{24}$$
$$\hat{x}_{ij,k}(t) = \hat{\boldsymbol{\varphi}}_{ij,k}^{\mathrm{T}}(t)\hat{\boldsymbol{\vartheta}}_{ij,k}, \tag{25}$$
$$\hat{\boldsymbol{\vartheta}}_{ij,k} = [\hat{a}_{ij,k}(1), \hat{a}_{ij,k}(2), \cdots, \hat{a}_{ij,k}(n_j), \hat{b}_{ij,k}(1), \hat{b}_{ij,k}(2), \cdots, \hat{b}_{ij,k}(n_j)]^{\mathrm{T}}, \tag{26}$$
$$\hat{\boldsymbol{\vartheta}}_{i,k} = [\hat{\boldsymbol{\vartheta}}_{i1,k}^{\mathrm{T}}, \hat{\boldsymbol{\vartheta}}_{i2,k}^{\mathrm{T}}, \cdots, \hat{\boldsymbol{\vartheta}}_{ir,k}^{\mathrm{T}}]^{\mathrm{T}}. \tag{27}$$

The steps for implementing the AM-LSI algorithm in (20)–(27) are listed in the following.

1. Initialize: Set data length $L$ and the maximum iteration $k_{\max}$. Let $k = 1$, $\hat{\boldsymbol{\vartheta}}_{i,0} = \mathbf{1}_{2n}/p_0$, $\hat{x}_{ij,0}(t-s) =$ random number, $t = 1, 2, \cdots, L$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, r$, $s = 1, 2, \cdots, n_j$, $p_0 = 10^6$.
2. Collect the input data $\{u_j(t),\ t = 1, 2, \cdots, L\}$ and the output data $\{y_i(t),\ t = 1, 2, \cdots, L\}$, construct the stacked output vector $\boldsymbol{Y}_i(L)$ using (21).
3. Form the sub-information vector $\hat{\boldsymbol{\varphi}}_{ij,k}(t)$ by (24), the information vector $\hat{\boldsymbol{\varphi}}_{i,k}(t)$ by (23) and the information matrix $\hat{\boldsymbol{\Xi}}_{i,k}(L)$ by (22).
4. Update the parameter estimate $\hat{\boldsymbol{\vartheta}}_{i,k}$ through (20).
5. Compute $\hat{x}_{ij,k}(t)$ by (25).
6. If $i < m$, let $i := i + 1$ and go to Step 3; otherwise, let $i = 1$ and go to Step 7.
7. If $k < k_{\max}$, let $k := k + 1$ and go to Step 3; otherwise, terminate the procedure and obtain the parameter estimate $\hat{\boldsymbol{\vartheta}}_{i,k_{\max}}$.

The computational efficiency is an important property of an algorithm. For large-scale systems, such as the MIMO system in (1), its dimensions $m$ and $r$ and order $n_j$ are generally large numbers. The proposed AM-LSI algorithm based on the traditional methods is difficult to satisfy the requirements of high identification efficiency in practical applications. Therefore, this paper aims to further propose a new identification algorithm with high computational efficiency to address this issue.

## 4. THE AM-HLSI ALGORITHM

The hierarchical identification principle is an effective method to simplify the system structure and improve the computational efficiency. In this section, we derive further a hierarchical identification model for the MIMO system in (1) by using the hierarchical identification principle. Based on the hierarchical identification model, we further propose a highly computational efficiency AM-HLSI algorithm to estimate the unknown parameters of the MIMO system in (1).

According to the hierarchical identification principle, we can decompose the fictitious output in (6) as

$$x_{ij}(t) = \boldsymbol{\phi}_{ij}^{\mathrm{T}}(t)\boldsymbol{\alpha}_{ij} + \boldsymbol{\psi}_j^{\mathrm{T}}(t)\boldsymbol{\beta}_{ij}, \tag{28}$$

where the sub-parameter vectors $\boldsymbol{\alpha}_{ij}$ and $\boldsymbol{\beta}_{ij}$ and the sub-information vectors $\boldsymbol{\phi}_{ij}(t)$ and $\boldsymbol{\psi}_j(t)$ are defined as

$$\boldsymbol{\alpha}_{ij} := [a_{ij}(1), a_{ij}(2), \cdots, a_{ij}(n_j)]^{\mathrm{T}} \in \mathbb{R}^{n_j}, \tag{29}$$
$$\boldsymbol{\beta}_{ij} := [b_{ij}(1), b_{ij}(2), \cdots, b_{ij}(n_j)]^{\mathrm{T}} \in \mathbb{R}^{n_j}, \tag{30}$$
$$\boldsymbol{\phi}_{ij}(t) := [-x_{ij}(t-1), -x_{ij}(t-2), \cdots, -x_{ij}(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{n_j}, \tag{31}$$
$$\boldsymbol{\psi}_j(t) := [u_j(t-1), u_j(t-2), \cdots, u_j(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{n_j}. \tag{32}$$

Substituting (28) into (4), we have

$$y_i(t) = \sum_{j=1}^{r} x_{ij}(t) + v_i(t)$$

$$\begin{aligned}&=\boldsymbol{\phi}_{i1}^{\mathrm{T}}(t)\boldsymbol{\alpha}_{i1}+\boldsymbol{\phi}_{i2}^{\mathrm{T}}(t)\boldsymbol{\alpha}_{i2}+\cdots+\boldsymbol{\phi}_{ir}^{\mathrm{T}}(t)\boldsymbol{\alpha}_{ir}\\&\quad+\boldsymbol{\psi}_{1}^{\mathrm{T}}(t)\boldsymbol{\beta}_{i1}+\boldsymbol{\psi}_{2}^{\mathrm{T}}(t)\boldsymbol{\beta}_{i2}+\cdots+\boldsymbol{\psi}_{r}^{\mathrm{T}}(t)\boldsymbol{\beta}_{ir}+v_i(t)\\&=\boldsymbol{\phi}_{i}^{\mathrm{T}}(t)\boldsymbol{\alpha}_{i}+\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\beta}_{i}+v_i(t),\end{aligned}\tag{33}$$

where the parameter vectors $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ and the information vectors $\boldsymbol{\phi}_i(t)$ and $\boldsymbol{\psi}(t)$ are defined as

$$\boldsymbol{\alpha}_i:=[\boldsymbol{\alpha}_{i1}^{\mathrm{T}},\boldsymbol{\alpha}_{i2}^{\mathrm{T}},\cdots,\boldsymbol{\alpha}_{ir}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^n,\tag{34}$$
$$\boldsymbol{\beta}_i:=[\boldsymbol{\beta}_{i1}^{\mathrm{T}},\boldsymbol{\beta}_{i2}^{\mathrm{T}},\cdots,\boldsymbol{\beta}_{ir}^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^n,\tag{35}$$
$$\boldsymbol{\phi}_i(t):=[\boldsymbol{\phi}_{i1}^{\mathrm{T}}(t),\boldsymbol{\phi}_{i2}^{\mathrm{T}}(t),\cdots,\boldsymbol{\phi}_{ir}^{\mathrm{T}}(t)]^{\mathrm{T}}\in\mathbb{R}^n,\tag{36}$$
$$\boldsymbol{\psi}(t):=[\boldsymbol{\psi}_{1}^{\mathrm{T}}(t),\boldsymbol{\psi}_{2}^{\mathrm{T}}(t),\cdots,\boldsymbol{\psi}_{r}^{\mathrm{T}}(t)]^{\mathrm{T}}\in\mathbb{R}^n.\tag{37}$$

Equation (33) is the hierarchical identification model of the MIMO system in (1). The hierarchical identification model contains all the unknown parameters $a_{ij}(1),a_{ij}(2),\cdots,a_{ij}(n_j),b_{ij}(1),b_{ij}(2),\cdots,b_{ij}(n_j)$ of the MIMO system. According to the hierarchical identification model, define two identification submodels:

$$\mathrm{S}_1:\quad \gamma_{a,i}(t):=\boldsymbol{\phi}_i^{\mathrm{T}}(t)\boldsymbol{\alpha}_i+v_i(t),\tag{38}$$
$$\mathrm{S}_2:\quad \gamma_{b,i}(t):=\boldsymbol{\psi}^{\mathrm{T}}(t)\boldsymbol{\beta}_i+v_i(t).\tag{39}$$

Define the stacked information matrices $\boldsymbol{\Phi}_i(L)$ and $\boldsymbol{\Psi}(L)$, and the stacked output vectors $\boldsymbol{\Gamma}_{a,i}(L)$ and $\boldsymbol{\Gamma}_{b,i}(L)$ as

$$\boldsymbol{\Phi}_i(L):=[\boldsymbol{\phi}_i(1),\boldsymbol{\phi}_i(2),\cdots,\boldsymbol{\phi}_i(L)]^{\mathrm{T}}\in\mathbb{R}^{L\times n},\tag{40}$$
$$\boldsymbol{\Psi}(L):=[\boldsymbol{\psi}(1),\boldsymbol{\psi}(2),\cdots,\boldsymbol{\psi}(L)]^{\mathrm{T}}\in\mathbb{R}^{L\times n},\tag{41}$$
$$\begin{aligned}\boldsymbol{\Gamma}_{a,i}(L)&:=[\gamma_{a,i}(1),\gamma_{a,i}(2),\cdots,\gamma_{a,i}(L)]^{\mathrm{T}}\\&=\boldsymbol{Y}_i(L)-\boldsymbol{\Psi}(L)\boldsymbol{\beta}_i\in\mathbb{R}^L,\end{aligned}\tag{42}$$
$$\begin{aligned}\boldsymbol{\Gamma}_{b,i}(L)&:=[\gamma_{b,i}(1),\gamma_{b,i}(2),\cdots,\gamma_{b,i}(L)]^{\mathrm{T}}\\&=\boldsymbol{Y}_i(L)-\boldsymbol{\Phi}_i(L)\boldsymbol{\alpha}_i\in\mathbb{R}^L.\end{aligned}\tag{43}$$

According to the identification Submodels $\mathrm{S}_1$ and $\mathrm{S}_2$, define two cost functions:

$$\begin{aligned}J_2(\boldsymbol{\alpha}_i)&:=\frac{1}{2}\sum_{l=1}^{L}[\gamma_{a,i}(l)-\boldsymbol{\phi}_i^{\mathrm{T}}(l)\boldsymbol{\alpha}_i]^2\\&=\frac{1}{2}\|\boldsymbol{\Gamma}_{a,i}(L)-\boldsymbol{\Phi}_i(L)\boldsymbol{\alpha}_i\|^2,\end{aligned}\tag{44}$$
$$\begin{aligned}J_3(\boldsymbol{\beta}_i)&:=\frac{1}{2}\sum_{l=1}^{L}[\gamma_{b,i}(l)-\boldsymbol{\psi}^{\mathrm{T}}(l)\boldsymbol{\beta}_i]^2\\&=\frac{1}{2}\|\boldsymbol{\Gamma}_{b,i}(L)-\boldsymbol{\Psi}(L)\boldsymbol{\beta}_i\|^2.\end{aligned}\tag{45}$$

Letting $\hat{\boldsymbol{\alpha}}_{i,k}$ and $\hat{\boldsymbol{\beta}}_{i,k}$ be the $k$th iteration estimates of the parameter vectors $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ and minimizing the cost functions $J_2(\boldsymbol{\alpha}_i)$ and $J_3(\boldsymbol{\beta}_i)$, we have

$$\hat{\boldsymbol{\alpha}}_{i,k}:=[\boldsymbol{\Phi}_i^{\mathrm{T}}(L)\boldsymbol{\Phi}_i(L)]^{-1}\boldsymbol{\Phi}_i^{\mathrm{T}}(L)\boldsymbol{\Gamma}_{a,i}(L),\tag{46}$$
$$\hat{\boldsymbol{\beta}}_{i,k}:=[\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Gamma}_{b,i}(L).\tag{47}$$

The stacked output vectors $\boldsymbol{\Gamma}_{a,i}(L)$ and $\boldsymbol{\Gamma}_{b,i}(L)$ in (46) and (47) contain the unknown variables $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$. One may replace the unknown variables $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ by their $(k-1)$th iteration estimates $\hat{\boldsymbol{\alpha}}_{i,k-1}$ and $\hat{\boldsymbol{\beta}}_{i,k-1}$. Thus, we have

$$\hat{\boldsymbol{\alpha}}_{i,k}=[\boldsymbol{\Phi}_i^{\mathrm{T}}(L)\boldsymbol{\Phi}_i(L)]^{-1}\boldsymbol{\Phi}_i^{\mathrm{T}}(L)[\boldsymbol{Y}_i(L)-\boldsymbol{\Psi}(L)\hat{\boldsymbol{\beta}}_{i,k-1}],\tag{48}$$
$$\hat{\boldsymbol{\beta}}_{i,k}=[\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(L)[\boldsymbol{Y}_i(L)-\boldsymbol{\Phi}_i(L)\hat{\boldsymbol{\alpha}}_{i,k-1}].\tag{49}$$



*Prepared using* acsauth.cls
DOI: 10.1002/acs

Similarly, the information vector $\boldsymbol{\Phi}_i(L)$ in (48) and (49) is formed by the information vector $\boldsymbol{\phi}_i(L)$ with the unknown fictitious output $x_{ij}(t)$. Therefore, Equations (48) and (49) cannot be implemented directly. Use the same method to establish an auxiliary model

$$\hat{x}_{ij,k}(t) := \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{ij,k}(t)\hat{\boldsymbol{\alpha}}_{ij,k}(t) + \boldsymbol{\psi}^{\mathrm{T}}_j(t)\hat{\boldsymbol{\beta}}_{ij,k}(t) \in \mathbb{R} \tag{50}$$

for each unknown fictitious output. Then, form the estimates $\hat{\boldsymbol{\phi}}_{ij,k}(t)$ of $\boldsymbol{\phi}_{ij}(t)$, $\hat{\boldsymbol{\phi}}_{i,k}(t)$ of $\boldsymbol{\phi}_i(t)$ and $\hat{\boldsymbol{\Phi}}_{i,k}(L)$ of $\boldsymbol{\Phi}_i(L)$ as

$$\hat{\boldsymbol{\phi}}_{ij,k}(t) := [-\hat{x}_{ij,k}(t-1), -\hat{x}_{ij,k}(t-2), \cdots, -\hat{x}_{ij,k}(t-n_j)]^{\mathrm{T}} \in \mathbb{R}^{n_j}, \tag{51}$$

$$\hat{\boldsymbol{\phi}}_{i,k}(t) := [\hat{\boldsymbol{\phi}}^{\mathrm{T}}_{i1,k}(t), \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{i2,k}(t), \cdots, \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{ir,k}(t)]^{\mathrm{T}} \in \mathbb{R}^{n}, \tag{52}$$

$$\hat{\boldsymbol{\Phi}}_{i,k}(L) := [\hat{\boldsymbol{\phi}}_{i,k}(1), \hat{\boldsymbol{\phi}}_{i,k}(2), \cdots, \hat{\boldsymbol{\phi}}_{i,k}(L)]^{\mathrm{T}} \in \mathbb{R}^{L\times n}. \tag{53}$$

Replacing the information vector $\boldsymbol{\Phi}_i(L)$ in (48) and (49) with its estimate $\hat{\boldsymbol{\Phi}}_{i,k}(L)$, we can obtain the AM-HLSI algorithm for estimating $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ in (33):

$$\hat{\boldsymbol{\alpha}}_{i,k} = [\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{i,k}(L)\hat{\boldsymbol{\Phi}}_{i,k}(L)]^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{i,k}(L)[\boldsymbol{Y}_i(L) - \boldsymbol{\Psi}(L)\hat{\boldsymbol{\beta}}_{i,k-1}], \tag{54}$$

$$\hat{\boldsymbol{\beta}}_{i,k} = [\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)]^{-1}\boldsymbol{\Psi}^{\mathrm{T}}(L)[\boldsymbol{Y}_i(L) - \hat{\boldsymbol{\Phi}}_{i,k}(L)\hat{\boldsymbol{\alpha}}_{i,k-1}], \tag{55}$$

$$\boldsymbol{Y}_i(L) = [y_i(1), y_i(2), \cdots, y_i(L)]^{\mathrm{T}}, \tag{56}$$

$$\hat{\boldsymbol{\Phi}}_{i,k}(L) = [\hat{\boldsymbol{\phi}}_{i,k}(1), \hat{\boldsymbol{\phi}}_{i,k}(2), \cdots, \hat{\boldsymbol{\phi}}_{i,k}(L)]^{\mathrm{T}}, \tag{57}$$

$$\boldsymbol{\psi}(L) = [\boldsymbol{\psi}(1), \boldsymbol{\psi}(2), \cdots, \boldsymbol{\psi}(L)]^{\mathrm{T}}, \tag{58}$$

$$\hat{\boldsymbol{\phi}}_{i,k}(t) = [\hat{\boldsymbol{\phi}}^{\mathrm{T}}_{i1,k}(t), \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{i2,k}(t), \cdots, \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{ir,k}(t)]^{\mathrm{T}}, \tag{59}$$

$$\boldsymbol{\psi}(t) = [\boldsymbol{\psi}^{\mathrm{T}}_1(t), \boldsymbol{\psi}^{\mathrm{T}}_2(t), \cdots, \boldsymbol{\psi}^{\mathrm{T}}_r(t)]^{\mathrm{T}}, \tag{60}$$

$$\hat{\boldsymbol{\phi}}_{ij,k}(t) = [-\hat{x}_{ij,k}(t-1), -\hat{x}_{ij,k}(t-2), \cdots, -\hat{x}_{ij,k}(t-n_j)]^{\mathrm{T}}, \tag{61}$$

$$\boldsymbol{\psi}_j(t) = [u_j(t-1), u_j(t-2), \cdots, u_j(t-n_j)]^{\mathrm{T}}, \tag{62}$$

$$\hat{x}_{ij,k}(t) = \hat{\boldsymbol{\phi}}^{\mathrm{T}}_{ij,k}(t)\hat{\boldsymbol{\alpha}}_{ij,k} + \boldsymbol{\psi}^{\mathrm{T}}_j(t)\hat{\boldsymbol{\beta}}_{ij,k}, \tag{63}$$

$$\hat{\boldsymbol{\alpha}}_{ij,k} = [\hat{a}_{ij,k}(1), \hat{a}_{ij,k}(2), \cdots, \hat{a}_{ij,k}(n_j)]^{\mathrm{T}}, \tag{64}$$

$$\hat{\boldsymbol{\beta}}_{ij,k} = [\hat{b}_{ij,k}(1), \hat{b}_{ij,k}(2), \cdots, \hat{b}_{ij,k}(n_j)]^{\mathrm{T}}, \tag{65}$$

$$\hat{\boldsymbol{\alpha}}_{i,k} = [\hat{\boldsymbol{\alpha}}^{\mathrm{T}}_{i1,k}, \hat{\boldsymbol{\alpha}}^{\mathrm{T}}_{i2,k}, \cdots, \hat{\boldsymbol{\alpha}}^{\mathrm{T}}_{ir,k}]^{\mathrm{T}}, \tag{66}$$

$$\hat{\boldsymbol{\beta}}_{i,k} = [\hat{\boldsymbol{\beta}}^{\mathrm{T}}_{i1,k}, \hat{\boldsymbol{\beta}}^{\mathrm{T}}_{i2,k}, \cdots, \hat{\boldsymbol{\beta}}^{\mathrm{T}}_{ir,k}]^{\mathrm{T}}. \tag{67}$$

The steps for implementing the AM-HLSI algorithm in (54)–(67) are listed in the following.

1. Initialize: Set data length $L$ and the maximum iteration $k_{\max}$. Let $k = 1$, $\hat{\boldsymbol{\alpha}}_{i,0} = \mathbf{1}_n/p_0$, $\hat{\boldsymbol{\beta}}_{i,0} = \mathbf{1}_n/p_0$, $\hat{x}_{ij,0}(t-s) =$ random number, $t = 1, 2, \cdots, L$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, r$, $s = 1, 2, \cdots, n_j$, $p_0 = 10^6$.
2. Collect the input data $\{u_j(t),\ t = 1, 2, \cdots, L\}$ and the output data $\{y_i(t),\ t = 1, 2, \cdots, L\}$, construct the stacked output vector $\boldsymbol{Y}_i(L)$ using (56). Form the sub-information vector $\boldsymbol{\psi}_j(t)$ by (62), the information vector $\psi(t)$ by (60) and the information matrix $\boldsymbol{\Psi}(L)$ by (58). Let $i = 1$.
3. Form the sub-information vector $\hat{\boldsymbol{\phi}}_{ij,k}(t)$ by (61), the information vector $\hat{\boldsymbol{\phi}}_{i,k}(t)$ by (59) and the information matrix $\hat{\boldsymbol{\Phi}}_{i,k}(L)$ by (57).
4. Update the parameter estimates $\hat{\boldsymbol{\alpha}}_{i,k}$ through (54) and $\hat{\boldsymbol{\beta}}_{i,k}$ through (55).
5. Compute $\hat{x}_{ij,k}(t)$ by (63).
6. If $i < m$, let $i := i + 1$ and go to Step 3; otherwise, let $i = 1$ and go to Step 7.
7. If $k < k_{\max}$, let $k := k + 1$ and go to Step 3; otherwise, terminate the procedure and obtain the parameter estimates $\hat{\boldsymbol{\alpha}}_{i,k_{\max}}$ and $\hat{\boldsymbol{\beta}}_{i,k_{\max}}$.

This section proposes a highly computational efficiency AM-HLSI algorithm. In order to show the advantages of the AM-HLSI algorithm, we compare the computational efficiency of the AM-LSI and AM-HLSI algorithms in detail in the next section. The hierarchical identification principle



DOI: 10.1002/acs

plays an important role not only in finding the solutions for identifying MIMO systems, but also in deriving the parameter estimation algorithms for other linear [30–32] and nonlinear systems [33–35].

## 5. THE ANALYSIS OF THE COMPUTATIONAL EFFICIENCY

In this section, we analyze the computational efficiency of the AM-LSI and AM-HLSI algorithms. The computational efficiency is measured by the number of multiplications (division is treated as multiplication) and the number of additions (subtraction is treated as addition). An addition or multiplication operation is called one floating point operation, i.e., one flop. The computational efficiency of the AM-LSI and AM-HLSI algorithms at each iteration is listed in Table I–II.

Table I. The computational efficiency of the AM-LSI algorithm

| Expressions | Multiplications | Additions |
| --- | --- | --- |
| $\hat{\boldsymbol{\vartheta}}_{i,k}=\boldsymbol{\Theta}'_{i,k}(L)\hat{\boldsymbol{\Xi}}^{\mathrm{T}}_{i,k}(L)\boldsymbol{Y}_i(L)\in\mathbb{R}^n$ | $4Lmn^2+2Lmn$ | $4Lmn^2-2mn$ |
| $\boldsymbol{\Theta}_{i,k}(L):=\hat{\boldsymbol{\Xi}}^{\mathrm{T}}_{i,k}(L)\hat{\boldsymbol{\Xi}}_{i,k}(L)\in\mathbb{R}^{(2n)\times(2n)}$ | $4Lmn^2$ | $4Lmn^2-4mn^2$ |
| $\boldsymbol{\Theta}'_{i,k}(L):=\boldsymbol{\Theta}^{-1}_{i,k}(L)\in\mathbb{R}^{(2n)\times(2n)}$ | $8mn^3$ | $8mn^3-4mn^2$ |
| $\hat{x}_{ij,k}(t)=\hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{ij,k}(t)\hat{\boldsymbol{\vartheta}}_{ij,k}\in\mathbb{R}$ | $4Lmn$ | $2Lmn-Lmr$ |
| Subtotal flops | $8Lmn^2+8mn^3+6Lmn$ | $8Lmn^2+8mn^3+2Lmn$ $-Lmr-8mn^2-2mn$ |
| Total flops | $S_1=16Lmn^3+16mn^3+8Lmn-Lmr-8mn^2-2mn$ | |

Table II. The computational efficiency of the AM-HLSI algorithm

| Expressions | Multiplications | Additions |
| --- | --- | --- |
| $\hat{\boldsymbol{\alpha}}_{i,k}=\boldsymbol{\Lambda}'_{i,k}(L)\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{i,k}(L)\boldsymbol{\Omega}_i(L)\in\mathbb{R}^n$ | $Lmn^2+Lmn$ | $Lmn^2-mn$ |
| $\boldsymbol{\Lambda}_{i,k}(L):=\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{i,k}(L)\hat{\boldsymbol{\Phi}}_{i,k}(L)\in\mathbb{R}^{n\times n}$ | $Lmn^2$ | $Lmn^2-mn^2$ |
| $\boldsymbol{\Omega}_i(L):=\boldsymbol{Y}_i(L)-\boldsymbol{\Psi}(L)\hat{\boldsymbol{\beta}}_{i,k-1}\in\mathbb{R}^L$ | $Lmn$ | $Lmn$ |
| $\boldsymbol{\Lambda}'_{i,k}(L):=\boldsymbol{\Lambda}^{-1}_{i,k}(L)\in\mathbb{R}^{n\times n}$ | $mn^3$ | $mn^3-mn^2$ |
| $\hat{\boldsymbol{\beta}}_{i,k}=\boldsymbol{\Lambda}'(L)\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{i,k}(L)\boldsymbol{\Upsilon}_i(L)\in\mathbb{R}^n$ | $Lmn^2+Lmn$ | $Lmn^2-mn$ |
| $\boldsymbol{\Lambda}(L):=\boldsymbol{\Psi}^{\mathrm{T}}(L)\boldsymbol{\Psi}(L)\in\mathbb{R}^{n\times n}$ | $Ln^2$ | $Ln^2-n^2$ |
| $\boldsymbol{\Upsilon}_i(L):=\boldsymbol{Y}_i(L)-\hat{\boldsymbol{\Phi}}_{i,k}(L)\hat{\boldsymbol{\alpha}}_{i,k-1}\in\mathbb{R}^L$ | $Lmn$ | $Lmn$ |
| $\boldsymbol{\Lambda}'(L):=\boldsymbol{\Lambda}^{-1}(L)\in\mathbb{R}^{n\times n}$ | $n^3$ | $n^3-n^2$ |
| $\hat{x}_{ij,k}(t)=\hat{\boldsymbol{\phi}}^{\mathrm{T}}_{ij,k}(t)\hat{\boldsymbol{\alpha}}_{ij,k}+\boldsymbol{\varphi}^{\mathrm{T}}_j(t)\hat{\boldsymbol{\beta}}_{ij,k}\in\mathbb{R}$ | $2Lmn$ | $2Lmn-Lmr$ |
| Subtotal flops | $3Lmn^2+mn^3$ $+6Lmn+Ln^2+n^3$ | $3Lmn^2+mn^3+4Lmn$ $-Lmr+Ln^2-2mn^2$ $+n^3-2mn-2n^2$ |
| Total flops | $S_2=6Lmn^3+2mn^3+10Lmn-Lmr+2Ln^2$ $-2mn^2+2n^3-2mn-2n^2$ | |

According to Table I–II, the number of operations of the AM-LSI and AM-HLSI algorithms is

$$S_1:=16Lmn^3+16mn^3+8Lmn-Lmr-8mn^2-2mn,$$
$$S_2:=6Lmn^3+2mn^3+10Lmn-Lmr+2Ln^2-2mn^2+2n^3-2mn-2n^2.$$

Then, compare the number of operations between the AM-LSI algorithm and AM-HLSI algorithm,

$$\begin{aligned}S_1-S_2&=(16Lmn^3+16mn^3+8Lmn-Lmr-8mn^2-2mn)\\&\quad-(6Lmn^3+2mn^3+10Lmn-Lmr+2Ln^2-2mn^2+2n^3-2mn-2n^2)\\&=10Lmn^3+14mn^3-2Lmn-2Ln^2-6mn^2-2n^3+2n^2.\end{aligned}$$

The accuracy of the identification algorithm depends on the data length $L$, so the data length $L$ should be large enough, i.e., $L\gg m$ and $L\gg n$. Therefore, we have $S_1-S_2>0$. Compared with the AM-LSI algorithm, the AM-HLSI algorithm effectively improves the computational efficiency by using the hierarchical identification principle. For example, letting $L=2000$, $m=1000$, $n=$



DOI: 10.1002/acs

5000 and $r = 1000$, we have

$$\begin{aligned} S_1 - S_2 = & (4 \times 10^{18} + 2 \times 10^{15} + 8 \times 10^{10} - 2 \times 10^9 - 2 \times 10^{11} - 1 \times 10^7) \\ & -(1.5 \times 10^{18} + 2.5 \times 10^{14} + 1 \times 10^{11} - 2 \times 10^9 \\ & +1 \times 10^{11} - 5 \times 10^{10} + 2.5 \times 10^{11} - 6 \times 10^7) \\ = & 4.00199987799 \times 10^{18} - 1.50025037994 \times 10^{18} \\ = & 2.50174949805 \times 10^{18} \text{ flops.} \end{aligned}$$

In order to visually show the difference of the number of operations between the AM-LSI and AM-HLSI algorithms, we fix $r = 500$ and the data length $L = 1000$, and draw the number of operations versus $m$ and $n$ in Figure 1.

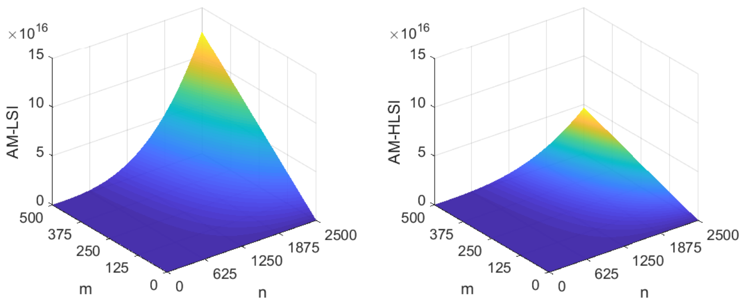


Figure 1. The AM-LSI and AM-HLSI flop numbers versus $m$ and $n$

## 6. EXAMPLE

In this section, the simulation example is provided to illustrate the efficiency of the AM-LSI and AM-HLSI algorithms. Consider the following two-input two-output system,

$$\begin{aligned} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} &= \begin{bmatrix} \frac{B_{11}(z)}{A_{11}(z)} & \frac{B_{12}(z)}{A_{12}(z)} \\ \frac{B_{21}(z)}{A_{21}(z)} & \frac{B_{22}(z)}{A_{22}(z)} \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} + \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}, \\ A_{11}(z) &= 1 + 0.46z^{-1} + 0.23z^{-2}, \ \ A_{12}(z) = 1 + 0.43z^{-1} + 0.28z^{-2}, \\ A_{21}(z) &= 1 + 0.54z^{-1} - 0.22z^{-2}, \ \ A_{22}(z) = 1 + 0.66z^{-1} + 0.37z^{-2}, \\ B_{11}(z) &= 0.58z^{-1} + 1.37z^{-2}, \ \ B_{12}(z) = 0.27z^{-1} + 0.32z^{-2}, \\ B_{21}(z) &= 0.68z^{-1} + 0.29z^{-2}, \ \ B_{22}(z) = 0.71z^{-1} + 0.57z^{-2}, \\ \boldsymbol{\alpha}_1 &= [a_{11}(1), a_{11}(2), a_{12}(1), a_{12}(2)]^{\mathrm{T}} = [0.46, 0.23, 0.43, 0.28]^{\mathrm{T}}, \\ \boldsymbol{\beta}_1 &= [b_{11}(1), b_{11}(2), b_{12}(1), b_{12}(2)]^{\mathrm{T}} = [0.58, 0.37, 0.27, 0.32]^{\mathrm{T}}, \\ \boldsymbol{\alpha}_2 &= [a_{21}(1), a_{21}(2), a_{22}(1), a_{22}(2)]^{\mathrm{T}} = [0.54, -0.22, 0.66, 0.37]^{\mathrm{T}}, \\ \boldsymbol{\beta}_2 &= [b_{21}(1), b_{21}(2), b_{22}(1), b_{22}(2)]^{\mathrm{T}} = [0.68, 0.29, 0.71, 0.57]^{\mathrm{T}}, \\ \boldsymbol{\vartheta}_1 &= [a_{11}(1), a_{11}(2), b_{11}(1), b_{11}(2), a_{12}(1), a_{12}(2), b_{12}(1), b_{12}(2)]^{\mathrm{T}} \\ &= [0.46, 0.23, 0.58, 0.37, 0.43, 0.28, 0.27, 0.32]^{\mathrm{T}}, \\ \boldsymbol{\vartheta}_2 &= [a_{21}(1), a_{21}(2), b_{21}(1), b_{21}(2), a_{22}(1), a_{22}(2), b_{22}(1), b_{22}(2)]^{\mathrm{T}} \\ &= [0.54, -0.22, 0.68, 0.29, 0.66, 0.37, 0.71, 0.57]^{\mathrm{T}}. \end{aligned}$$

In simulation, the inputs $u_1(t)$ and $u_2(t)$ are taken as uncorrelated persistent excitation signal sequences with zero mean and unit variance. The noises $v_1(t)$ and $v_2(t)$ are taken as uncorrelated white noise sequences with zero mean. When the noise variance $\sigma^2 = 1.00^2$, the 1st output channel




DOI: 10.1002/acs

noise-to-signal ratio of the system is $\delta_{\rm ns} = 208.57\%$ and the 2st output channel noise-to-signal ratio of the system is $\delta_{\rm ns} = 147.86\%$. Define the parameter estimation error:

$$\delta := \sum_{i=1}^{m} \|[\hat{\boldsymbol{\alpha}}_{i,k}^{\rm T}, \hat{\boldsymbol{\beta}}_{i,k}^{\rm T}] - [\boldsymbol{\alpha}_i^{\rm T}, \boldsymbol{\beta}_i^{\rm T}]\|/\|[\boldsymbol{\alpha}_i^{\rm T}, \boldsymbol{\beta}_i^{\rm T}]\| \times 100\%$$
$$= \sum_{i=1}^{m} \|\hat{\boldsymbol{\vartheta}}_{i,k} - \boldsymbol{\vartheta}_i\|/\|\boldsymbol{\vartheta}_i\| \times 100\%.$$

To test the performance of the AM-LSI algorithm in (20)–(27) and the AM-HLSI algorithm in (54)–(67). Take the noise variance $\sigma^2 = 0.10^2$ and the data length $L = 1000$, and apply the AM-LSI and AM-HLSI algorithms to identify this MIMO system. The AM-LSI and AM-HLSI parameter estimates versus $k$ are listed in Table III and the AM-LSI and AM-HLSI parameter estimation errors versus $k$ are shown in Figure 2.

Table III. The AM-LSI and AM-HLSI parameter estimates versus $k$ ($\sigma^2 = 0.10^2$)

| Algorithms | $k$ | 5 | 10 | 20 | 50 | 100 | True values |
|---|---|---|---|---|---|---|---|
| AM-LSI | $a_{11}(1)$ | 0.45958 | 0.46170 | 0.46224 | 0.46226 | 0.46226 | 0.46000 |
| | $a_{11}(2)$ | 0.22972 | 0.23149 | 0.23148 | 0.23148 | 0.23148 | 0.23000 |
| | $a_{12}(1)$ | 0.42369 | 0.42563 | 0.42548 | 0.42547 | 0.42547 | 0.43000 |
| | $a_{12}(2)$ | 0.28903 | 0.28890 | 0.28884 | 0.28884 | 0.28884 | 0.28000 |
| | $a_{21}(1)$ | 0.53567 | 0.54714 | 0.54059 | 0.54420 | 0.54674 | 0.54000 |
| | $a_{21}(2)$ | -0.22754 | -0.21903 | -0.22249 | -0.21991 | -0.21824 | -0.22000 |
| | $a_{22}(1)$ | 0.66600 | 0.66045 | 0.65990 | 0.65433 | 0.65077 | 0.66000 |
| | $a_{22}(2)$ | 0.36351 | 0.36728 | 0.36763 | 0.36364 | 0.36093 | 0.37000 |
| | $b_{11}(1)$ | 0.58808 | 0.58676 | 0.58684 | 0.58684 | 0.58684 | 0.58000 |
| | $b_{11}(2)$ | 0.37370 | 0.37439 | 0.37474 | 0.37475 | 0.37475 | 0.37000 |
| | $b_{12}(1)$ | 0.27443 | 0.27416 | 0.27414 | 0.27415 | 0.27415 | 0.27000 |
| | $b_{12}(2)$ | 0.31578 | 0.31634 | 0.31630 | 0.31630 | 0.31630 | 0.32000 |
| | $b_{21}(1)$ | 0.68912 | 0.68334 | 0.68513 | 0.68261 | 0.68134 | 0.68000 |
| | $b_{21}(2)$ | 0.29466 | 0.30126 | 0.29765 | 0.29884 | 0.29985 | 0.29000 |
| | $b_{22}(1)$ | 0.70446 | 0.71155 | 0.70842 | 0.71001 | 0.71022 | 0.71000 |
| | $b_{22}(2)$ | 0.57749 | 0.57346 | 0.57249 | 0.56914 | 0.56651 | 0.57000 |
| | $\delta$(%) | 1.29774 | 1.08921 | 0.94850 | 1.05014 | 1.24065 | |
| AM-HLSI | $a_{11}(1)$ | 0.12934 | 0.28466 | 0.42484 | 0.46209 | 0.46226 | 0.46000 |
| | $a_{11}(2)$ | 0.25647 | 0.24154 | 0.23253 | 0.23148 | 0.23148 | 0.23000 |
| | $a_{12}(1)$ | 0.37873 | 0.42204 | 0.42441 | 0.42546 | 0.42547 | 0.43000 |
| | $a_{12}(2)$ | 0.27357 | 0.28872 | 0.28889 | 0.28884 | 0.28884 | 0.28000 |
| | $a_{21}(1)$ | 0.11140 | 0.28929 | 0.47681 | 0.54140 | 0.54157 | 0.54000 |
| | $a_{21}(2)$ | -0.31551 | -0.28351 | -0.24193 | -0.22185 | -0.22209 | -0.22000 |
| | $a_{22}(1)$ | 0.27620 | 0.58454 | 0.66686 | 0.66288 | 0.65882 | 0.66000 |
| | $a_{22}(2)$ | 0.33655 | 0.34892 | 0.37177 | 0.37122 | 0.36928 | 0.37000 |
| | $b_{11}(1)$ | 0.58813 | 0.58815 | 0.58733 | 0.58685 | 0.58684 | 0.58000 |
| | $b_{11}(2)$ | 0.17927 | 0.27108 | 0.35299 | 0.37465 | 0.37475 | 0.37000 |
| | $b_{12}(1)$ | 0.27309 | 0.27333 | 0.27395 | 0.27415 | 0.27415 | 0.27000 |
| | $b_{12}(2)$ | 0.30062 | 0.31422 | 0.31579 | 0.31630 | 0.31630 | 0.32000 |
| | $b_{21}(1)$ | 0.68042 | 0.68322 | 0.68353 | 0.68344 | 0.68374 | 0.68000 |
| | $b_{21}(2)$ | 0.00373 | 0.12761 | 0.25430 | 0.29755 | 0.29775 | 0.29000 |
| | $b_{22}(1)$ | 0.70598 | 0.70717 | 0.71085 | 0.71261 | 0.71402 | 0.71000 |
| | $b_{22}(2)$ | 0.30289 | 0.52089 | 0.57983 | 0.57634 | 0.57372 | 0.57000 |
| | $\delta$(%) | 43.11816 | 20.29852 | 4.69544 | 0.98435 | 0.95935 | |

In order to study the influence of the different noise variances on the AM-LSI and AM-HLSI algorithms, we set the noise variances between $\sigma^2 = 0.20^2$ to $\sigma^2 = 1.20^2$ and the data length $L = 1000$, and use the AM-LSI and AM-HLSI algorithms to identify this system. Figure 3 and Figure 4 show the AM-LSI and AM-HLSI parameter estimation errors versus $k$ under the different noise variances.

To compare the influence of the different data lengths on the AM-LSI and AM-HLSI algorithms. First, we fix $L = 400$, $L = 800$ and $L = 1200$, and apply the AM-LSI algorithm to estimate this MIMO system. The AM-LSI parameter estimation errors versus $k$ are presented in Figure 5. Then, we take $L = 400$, $L = 600$, $L = 800$, $L = 1000$ and $L = 1200$, and adopt the AM-HLSI algorithm to estimate this MIMO system. The AM-HLSI parameter estimates versus $k$ are listed in Table IV and the AM-HLSI parameter estimation errors versus $k$ are presented in Figure 6.

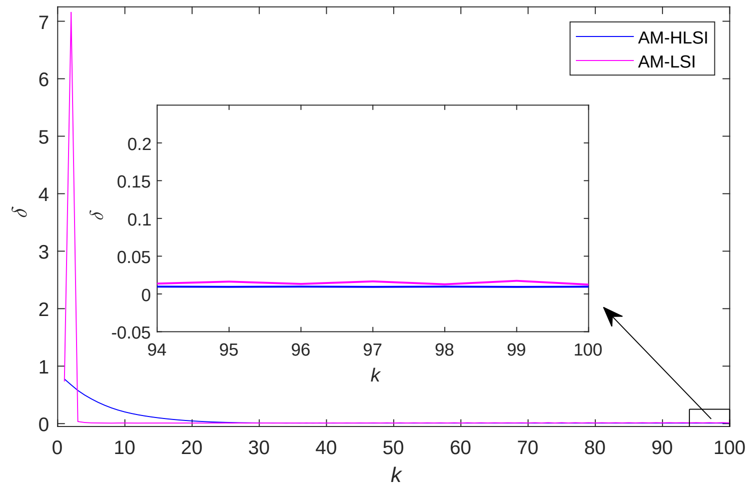


Figure 2. The AM-LSI and AM-HLSI parameter estimation errors $\delta$ versus $k$ ($\sigma^2 = 0.10^2$)

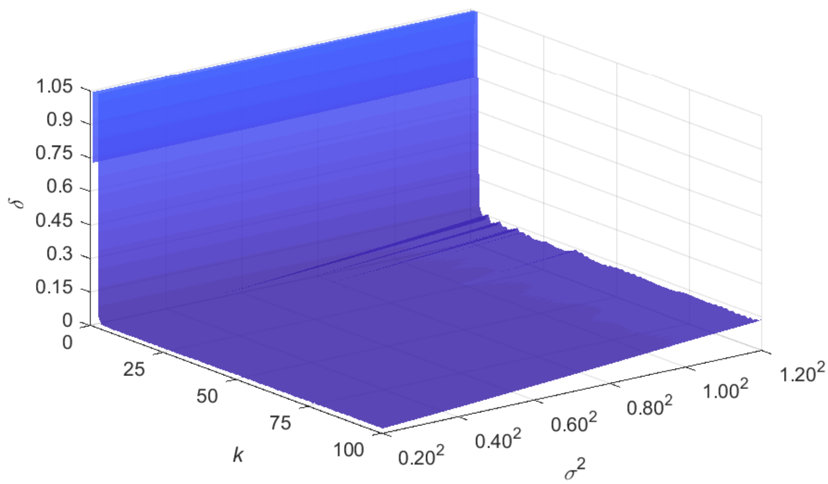


Figure 3. The AM-LSI parameter estimation errors $\delta$ versus $k$ under different $\sigma^2$

From Table III–IV and Figures 2–6, we can draw the following conclusions.

- The AM-LSI and AM-HLSI parameter estimation errors become smaller as the iteration $k$ increases. Thus, the proposed algorithms are effective for the MIMO systems.
- The AM-HLSI algorithm has higher parameter estimation accuracy than the AM-LSI algorithm under the same noise variance and data length.
- A lower noise variance leads to higher parameter estimation accuracy given by the AM-LSI and AM-HLSI algorithms under the same data length.
- As the data length increases, the parameter estimation errors for both AM-LSI and AM-HLSI become smaller under the same noise variance. However, these errors cannot be reduced infinitely. When the data length reaches 1000, the AM-HLSI estimation error stabilizes at approximately $1\%$. Therefore, it is important to select an appropriate data length for practical applications.

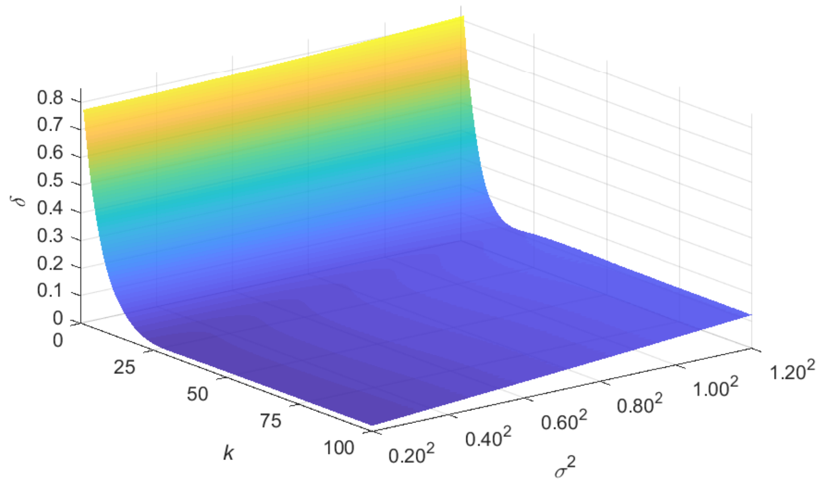


Figure 4. The AM-HLSI parameter estimation errors $\delta$ versus $k$ under different $\sigma^2$

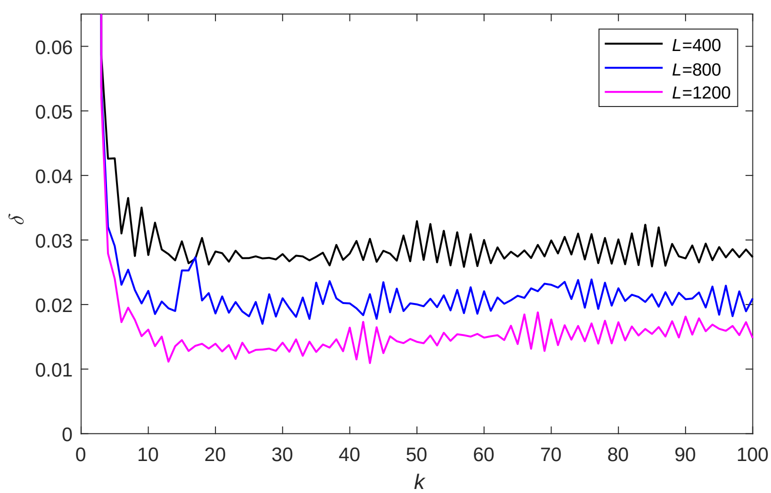


Figure 5. The AM-LSI parameter estimation errors $\delta$ versus $k$ under different $L$ ($\sigma^2 = 0.10^2$)

## 7. CONCLUSIONS

In this paper, we study the parameter identification of the MIMO systems described by the output-error model and give an AM-LSI algorithm by using the auxiliary model idea. In order to improve the computational efficiency, a highly computational efficiency AM-HLSI based on the hierarchical identification principle is further proposed in this paper. Through the analysis of the AM-LSI and AM-HLSI algorithms, the AM-HLSI algorithm has higher computational efficiency than the AM-HLS algorithm. The proposed AM-HLSI algorithm can effectively reduce the computational burden and improve the computational efficiency. The simulation example demonstrates that the AM-LSI and AM-HLSI algorithms are effective to identify the MIMO systems. Compared with the AM-LSI algorithm, the AM-HLSI algorithm has higher parameter estimation accuracy under the same conditions. The proposed AM-HLSI algorithm has excellent performance in identifying MIMO systems.

Table IV. The AM-HLSI parameter estimates versus $k$ under different $L$ ($\sigma^2 = 0.10^2$)

| $L$ | $k$ | 5 | 10 | 20 | 50 | 100 | True values |
|---|---|---|---|---|---|---|---|
| 400 | $a_{11}(1)$ | 0.14903 | 0.28911 | 0.40092 | 0.42830 | 0.42843 | 0.46000 |
| | $a_{11}(2)$ | 0.26238 | 0.24740 | 0.23802 | 0.23643 | 0.23642 | 0.23000 |
| | $a_{12}(1)$ | 0.36334 | 0.40313 | 0.40949 | 0.41082 | 0.41082 | 0.43000 |
| | $a_{12}(2)$ | 0.28621 | 0.28953 | 0.28765 | 0.28715 | 0.28715 | 0.28000 |
| | $a_{21}(1)$ | 0.17967 | 0.38857 | 0.51405 | 0.55297 | 0.55338 | 0.54000 |
| | $a_{21}(2)$ | -0.32668 | -0.26434 | -0.22251 | -0.20847 | -0.20932 | -0.22000 |
| | $a_{22}(1)$ | 0.35965 | 0.63988 | 0.66602 | 0.66320 | 0.65824 | 0.66000 |
| | $a_{22}(2)$ | 0.32711 | 0.35205 | 0.36782 | 0.36799 | 0.36446 | 0.37000 |
| | $b_{11}(1)$ | 0.58343 | 0.58606 | 0.58634 | 0.58620 | 0.58620 | 0.58000 |
| | $b_{11}(2)$ | 0.19585 | 0.27723 | 0.34173 | 0.35742 | 0.35749 | 0.37000 |
| | $b_{12}(1)$ | 0.28081 | 0.27707 | 0.27474 | 0.27404 | 0.27404 | 0.27000 |
| | $b_{12}(2)$ | 0.29309 | 0.30061 | 0.30103 | 0.30103 | 0.30112 | 0.32000 |
| | $b_{21}(1)$ | 0.65832 | 0.67697 | 0.68308 | 0.68254 | 0.68167 | 0.68000 |
| | $b_{21}(2)$ | 0.05949 | 0.19118 | 0.27313 | 0.29807 | 0.29815 | 0.29000 |
| | $b_{22}(1)$ | 0.73455 | 0.72015 | 0.71752 | 0.71705 | 0.71658 | 0.71000 |
| | $b_{22}(2)$ | 0.36481 | 0.55926 | 0.57592 | 0.57409 | 0.57055 | 0.57000 |
| | $\delta(\%)$ | 36.55720 | 14.71396 | 4.28540 | 2.67973 | 2.66678 | |
| 800 | $a_{11}(1)$ | 0.14934 | 0.29245 | 0.41904 | 0.45403 | 0.45422 | 0.46000 |
| | $a_{11}(2)$ | 0.25375 | 0.24092 | 0.23332 | 0.23239 | 0.23238 | 0.23000 |
| | $a_{12}(1)$ | 0.41917 | 0.43469 | 0.43351 | 0.43256 | 0.43255 | 0.43000 |
| | $a_{12}(2)$ | 0.26010 | 0.27719 | 0.27800 | 0.27798 | 0.27798 | 0.28000 |
| | $a_{21}(1)$ | 0.13717 | 0.30591 | 0.47912 | 0.53446 | 0.53436 | 0.54000 |
| | $a_{21}(2)$ | -0.29689 | -0.27888 | -0.24141 | -0.22420 | -0.22461 | -0.22000 |
| | $a_{22}(1)$ | 0.27085 | 0.56436 | 0.66229 | 0.66813 | 0.66510 | 0.66000 |
| | $a_{22}(2)$ | 0.34112 | 0.35251 | 0.37318 | 0.37347 | 0.37129 | 0.37000 |
| | $b_{11}(1)$ | 0.58322 | 0.58474 | 0.58477 | 0.58456 | 0.58455 | 0.58000 |
| | $b_{11}(2)$ | 0.19336 | 0.27726 | 0.35124 | 0.37160 | 0.37171 | 0.37000 |
| | $b_{12}(1)$ | 0.25599 | 0.25787 | 0.25981 | 0.26045 | 0.26045 | 0.27000 |
| | $b_{12}(2)$ | 0.31463 | 0.31863 | 0.31728 | 0.31665 | 0.31664 | 0.32000 |
| | $b_{21}(1)$ | 0.66773 | 0.67171 | 0.67394 | 0.67433 | 0.67513 | 0.68000 |
| | $b_{21}(2)$ | 0.02705 | 0.13717 | 0.24929 | 0.28578 | 0.28616 | 0.29000 |
| | $b_{22}(1)$ | 0.71137 | 0.71039 | 0.71089 | 0.71000 | 0.70895 | 0.71000 |
| | $b_{22}(2)$ | 0.28330 | 0.49708 | 0.57290 | 0.57836 | 0.57585 | 0.57000 |
| | $\delta(\%)$ | 41.64961 | 19.62713 | 4.82775 | 1.10666 | 0.97641 | |
| 1200 | $a_{11}(1)$ | 0.13418 | 0.28962 | 0.42958 | 0.46724 | 0.46742 | 0.46000 |
| | $a_{11}(2)$ | 0.25755 | 0.24288 | 0.23485 | 0.23410 | 0.23410 | 0.23000 |
| | $a_{12}(1)$ | 0.39774 | 0.42845 | 0.42745 | 0.42769 | 0.42769 | 0.43000 |
| | $a_{12}(2)$ | 0.26524 | 0.28103 | 0.28284 | 0.28340 | 0.28340 | 0.28000 |
| | $a_{21}(1)$ | 0.13765 | 0.32243 | 0.48886 | 0.53597 | 0.53615 | 0.54000 |
| | $a_{21}(2)$ | -0.32085 | -0.28922 | -0.24546 | -0.22967 | -0.22977 | -0.22000 |
| | $a_{22}(1)$ | 0.29563 | 0.58642 | 0.65986 | 0.65904 | 0.65636 | 0.66000 |
| | $a_{22}(2)$ | 0.34035 | 0.35405 | 0.37352 | 0.37349 | 0.37200 | 0.37000 |
| | $b_{11}(1)$ | 0.58540 | 0.58519 | 0.58462 | 0.58431 | 0.58431 | 0.58000 |
| | $b_{11}(2)$ | 0.18130 | 0.27193 | 0.35333 | 0.37518 | 0.37528 | 0.37000 |
| | $b_{12}(1)$ | 0.26942 | 0.26828 | 0.26749 | 0.26727 | 0.26727 | 0.27000 |
| | $b_{12}(2)$ | 0.30481 | 0.31479 | 0.31576 | 0.31619 | 0.31620 | 0.32000 |
| | $b_{21}(1)$ | 0.67417 | 0.67573 | 0.67748 | 0.67796 | 0.67823 | 0.68000 |
| | $b_{21}(2)$ | 0.02140 | 0.14459 | 0.25627 | 0.28789 | 0.28810 | 0.29000 |
| | $b_{22}(1)$ | 0.70925 | 0.70854 | 0.70904 | 0.70884 | 0.70896 | 0.71000 |
| | $b_{22}(2)$ | 0.30947 | 0.51547 | 0.56860 | 0.56780 | 0.56585 | 0.57000 |
| | $\delta(\%)$ | 41.26079 | 18.67385 | 4.05958 | 0.91669 | 0.94542 | |

## DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in this article.

### ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (No. 61873111) and the 111 Project (B12018).

### REFERENCES

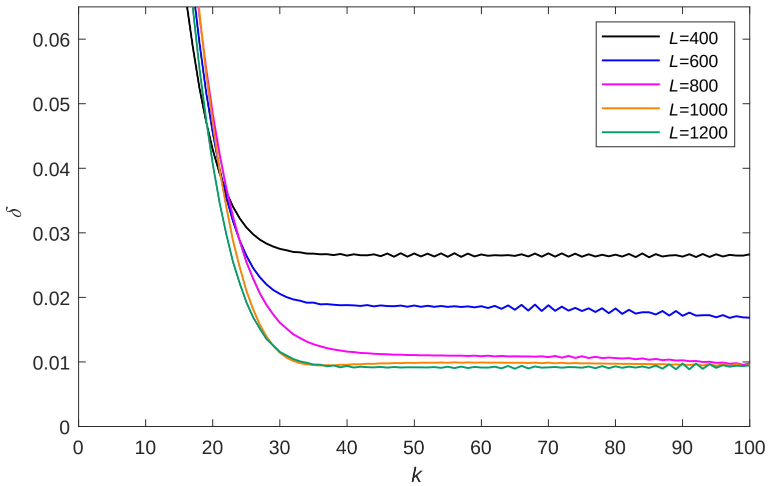


Figure 6. The AM-HLSI parameter estimation errors $\delta$ versus $k$ under different $L$ ($\sigma^2 = 0.10^2$)


1. Lorenzetti J, McClellan A, Farhat C, Pavone M. Linear reduced-order model predictive control. *IEEE Trans Automat Contr*. 2022; 67(11):5980-5995.
2. Ghasemi M, Chang S, Sivaloganathan S. Development of an integrated ultrasonic biofilm detachment model for biofilm thickness control in membrane aerated bioreactors. *Appl Math Model*. 2021;100:597-611.
3. Sikora BM, Pilat AK. Analytical modeling and experimental validation of the six pole axial active magnetic bearing. *Appl Math Model*. 2022;104:50-66.
4. Verma VK, Tripathi JN. Device parameters based analytical modeling of ground-bounce induced jitter in CMOS inverters. *IEEE Trans Electron Devices*. 2022;69(10):5462-5469.
5. Park JK, Kim YJ , Lu XN. New analytical model of microgrid frequency and voltage variations due to network reconfiguration. *IEEE Trans Smart Grid*. 2021;12(1):905-908.
6. Lauwers O, Vermeersch C, De Moor B. Cepstral identification of autoregressive systems. *Automatica*. 2022;139:110214.
7. Lee T, Lee CD, Park FC. Optimal excitation trajectories for mechanical systems identification. *Automatica*. 2021;131:109773.
8. Brouri A. Wiener-Hammerstein nonlinear system identification using spectral analysis. *Int J Robust Nonlinear Control*. 2022;32(10):6184-6204.
9. Morelli F, Bombois X, Hjalmarsson H, Bako L, Colin K. Least costly identification experiment for the identification of one module in a dynamic network. *Automatica*. 2021;133:109855.
10. Sadeqi A, Moradi S. Time-domain model identification based on reduced output data. *ISA Trans*. 2021;110:344-356.
11. Louis M, Fabien L, Marion G. A statistical learning perspective on switched linear system identification. *Automatica*. 2022;145:110532.
12. Alex SL, Erik W, Girish NN. Identification of FIR systems with binary input and output observations. *IEEE Trans Automat Contr*. 2021;66(3):1190-1198.
13. Shahid MI, King Q. Event-triggered distributed dynamic output-feedback dissipative control of multi-weighted and multi-delayed large-scale systems. *ISA Trans*. 2020;96:116-131.
14. Zimenko K, Polyakov A, Efimov D, Kremlev A. Homogeneity based finite/fixed-time observers for linear MIMO systems. *Int J Robust Nonlinear Control*. 2022;DOI: 10.1002/rnc.6284.
15. Ahmadipour M, Seydnejad SR, Barkhordari M. Subspace-based deterministic identification of MIMO linear state-delayed systems. *Circuits Syst Signal Process*. 2020;39(8):4067-4093.
16. Rekik O, Abed-Meraim K, Nait-Meziane M, Mokraoui A, Trung NL. Maximum likelihood based identification for nonlinear multichannel communications systems. *Signal Process*. 2021;189:108297.
17. Fazzi A, Grossmann B, Mercere G, Markovsky I. MIMO system identification using common denominator and numerators with known degrees. *Int J Adapt Control Signal Process*. 2022;36(4):870-881.
18. Formentin S, Chiuso A. Control-oriented regularization for linear system identification. *Automatica*. 2021;127:109539.
19. Cerone V, Razza V, Regruto D. A unified framework for the identification of a general class of multivariable nonlinear block-structured systems. *Int J Robust Nonlinear Control*. 2021;31(15):7344-7360.
20. Ma P, Wang L. Filtering-based recursive least squares estimation approaches for multivariate equation-error systems by using the multiinnovation theory. *Int J Adapt Control Signal Process*. 2021;35(9):1898-1915.
21. Xia HF, Xie L, Zhu QM. Maximum likelihood iterative identification approaches for multivariable equation-error moving average systems. *Int J Syst Science*. 2020;51(16):3285-3298.
22. Gu Y, Dai W, Zhu QM, Nouri H. Hierarchical multi-innovation stochastic gradient identification algorithm for estimating a bilinear state-space model with moving average noise. *J Comput Appl Math*. 2023;420:114794.

23. Liu HB, Wang JW, Meng XX. Hierarchical maximum likelihood generalized extended stochastic gradient algorithms for bilinear-in-parameter systems. *Optim Control Appl Methods*. 2022;43(2):402-417.
24. Pan J, Liu SD, Shu J, Wan XK. Hierarchical recursive least squares estimation algorithm for secondorder Volterra nonlinear systems. *Int J Control Automat Syst*. 2022;20(12):3940-3950.
25. Shahriari F, Arefi MM, Luo H, Yin S. Multistage parameter estimation algorithms for identification of bilinear systems. *Nonlinear Dyn*. 2022;110(3):2635-2655.
26. Chaudhary NI, Raja MAZ, Khan ZA, Mehmood A, Shah SM. Design of fractional hierarchical gradient descent algorithm for parameter estimation of nonlinear control autoregressive systems. *Chaos Solitons Fractals*. 2022;157:111913.
27. Liu J, Ji Y. Auxiliary model-based recursive least squares algorithm for two-input single-output Hammerstein output-error moving average systems by using the hierarchical identification principle. *Int J Robust Nonlinear Control*. 2022;32(8):7575-7593.
28. Zong TC, Li JH, Lu GP. Parameter identification of dual-rate Hammerstein-Volterra nonlinear systems by the hybrid particle swarm-gradient algorithm based on the auxiliary model. *Eng Appl Artif Intell*. 2023;117:105526.
29. Chaudhary NI, Khan ZA, Kiani AK, Raja MAZ, Chaudhary II, Pinto CMA. Design of auxiliary model based normalized fractional gradient algorithm for nonlinear output-error systems. *Chaos Solitons Fractals*. 2022;163:112611.
30. Ma H, Pan J, Ding W. Partially-coupled least squares based iterative parameter estimation for multi-variable output-error-like autoregressive moving average systems. *IET Control Theory Appl*. 2019;13(18):3040-3051.
31. Pan J, Ma H, Zhang X, et al. Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises. *IET Signal Process*. 2020;14(7):455-466.
32. Ji Y, Jiang AN. Filtering-based accelerated estimation approach for generalized time-varying systems with disturbances and colored noises. *IEEE Trans Circuits Syst II Express Briefs*. 2023;70(1):206-210.
33. Ji Y, Kang Z. Three-stage forgetting factor stochastic gradient parameter estimation methods for a class of nonlinear systems. *Int J Robust Nonlinear Control*. 2021;31(3):971-987.
34. Pan J, Liu YQ, Shu J. Gradient-based parameter estimation for an exponential nonlinear autoregressive time-series model by using the multi-innovation. *Int J Control Automat Syst*. 2023;21(1):140-150.
35. Ji Y, Jiang XK, Wan LJ. Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems. *J Franklin Inst*. 2020;357(8):5019-5032.