International Journal of Control, Automation and Systems 18(4) (2020) 886-896
http://dx.doi.org/10.1007/s12555-019-0140-3
ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Two-stage Gradient-based Iterative Estimation Methods for Controlled Autoregressive Systems Using the Measurement Data

Feng Ding*, Lei Lv, Jian Pan, Xiangkui Wan, and Xue-Bo Jin

Abstract: This paper considers the parameter identification problems of controlled autoregressive systems using observation information. According to the hierarchical identification principle, we decompose the controlled autoregressive system into two subsystems by introducing two fictitious output variables. Then a two-stage gradient-based iterative algorithm is proposed by means of the iterative technique. In order to improve the performance of the tracking the time-varying parameters, we derive a two-stage multi-innovation gradient-based iterative algorithm based on the multi-innovation identification theory. Finally, an example is provided to illustrate the effectiveness of the proposed algorithms.

Keywords: Gradient search, hierarchical identification, iterative technique, mathematical modeling, multi-innovation identification, parameter estimation.

1. INTRODUCTION

Mathematical models are useful for system analysis and control [1,2] and can be applied in many areas such as network model fitting [3, 4]. Basically, there are two ways of constructing mathematical models [5–8]. The first one is the analytic approach, which uses basic laws from physics to describe the dynamical behavior of a process. Due to the difficult of the first way, system identification, which is an experimental way, becomes the common choice to establish mathematical models [9]. Some identification methods have been used for signal modeling [10–12] and time-series modeling [13].

The identification methods can be divided into two main types, one of which is the on-line methods and the other is off-line methods [14–17]. The on-line methods, which are also called the recursive methods, compute the parameter estimates in real-time. The off-line methods, i.e., the iterative methods, update the parameter estimates by using batch data [18, 19]. By eliminating the state variables, Li and Liu gave the input-output representation for a class of bilinear systems and proposed a filtering based least squares iterative algorithm by means of the hierarchical principle [20] and the data filtering technique [21].

The multi-innovation identification is an important branch of system identification [22, 23]. The innovation refers to the useful information that can improve the parameter estimation accuracy. The main idea is to expand the dimension of the innovation and to make full use of the system information. In this literature, for the sine combination signals and periodic signals, a multi-innovation stochastic gradient algorithm was derived by expanding the scalar innovation into the innovation vector [24]. In this paper, we expand the multi-innovation theory into the iterative identification algorithm. In particular, we update the parameter estimates by using the data in a moving data window which moves forward with time increasing.

In this paper, the parameter estimation problems of controlled autoregressive systems are investigated using the gradient search [25] and the hierarchical identification principle [26]. On the basis of the gradient-based iterative parameter estimation algorithms for dynamical systems [27, 28], this work is to divide the linear autoregressive identification model into two sub-identification models, and to identify the parameter vectors of these two models, respectively. The main contributions of this paper are as follows:

- A two-stage identification model is deduced for controlled autoregressive systems by adopting the hierarchical identification principle.
- Based on the gradient search, a two-stage gradient-

Manuscript received March 1, 2019; revised July 22, 2019; accepted August 15, 2019. Recommended by Associate Editor Yongping Pan under the direction of Editor Young IL Lee. This work was supported by the National Natural Science Foundation of China (no. 61571182).

Feng Ding, Lei Lv, Jian Pan, and Xiangkui Wan are with the Hubei Key Laboratory for High-efficiency Utilization of Solar Energy and Operation Control of Energy Storage System, the School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, P. R. China (e-mails: fding@qust.edu.cn, lvlei9076@163.com, jpan@163.com, wanxiangkui@163.com). Feng Ding is with the College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, P. R. China. Xue-Bo Jin is with the School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, P. R. China (e-mail: jinxuebo@btbu.edu.cn).
* Corresponding author.

based iterative algorithm is presented for identifying the parameters of the controlled autoregressive systems.

- A two-stage multi-innovation gradient-based iterative algorithm is derived by means of the multi-innovation identification theory.

The rest of this paper is organized as follows: Section 2 offers some notations and derives the two-stage identification model of controlled autoregressive systems. A two-stage gradient-based iterative algorithm is presented in Section 3. By employing the multi-innovation identification theory, Section 4 deduces a two-stage multi-innovation gradient-based iterative algorithm. Section 5 gives the gradient-based iterative algorithm for comparison. Section 6 offers an example to illustrate the effectiveness of the proposed algorithms. Finally, Section 7 gives some concluding remarks.

2. SYSTEM DESCRIPTION AND IDENTIFICATION MODELS

Let us start with the necessary notations. "$A =: X$" or "$X := A$" stands for "A is defined as X"; the superscript T stands for the vector/matrix transpose; The symbol I_n denotes an identity matrix of size $n \times n$; $\mathbf{1}_n$ stands for an n-dimensional column vector whose elements are 1; the norm of a matrix (or a column vector) X is defined by $\|X\|^2 := \mathrm{tr}[XX^{\mathrm{T}}]$. $\lambda_{\max}[X]$ is the largest eigenvalue of the symmetric matrix X.

Consider the following controlled autoregressive (CAR) system:

$$A(z)y(t) = B(z)u(t) + v(t), \tag{1}$$

where $u(t)$ is the input of the system and $y(t)$ is the output of the system, $v(t)$ is a white noise with zero mean, $A(z)$ and $B(z)$ are polynomials in the unit backward shift operator $[z^{-1}y(t) = y(t-1), zy(t) = y(t+1)]$, and defined as

$$\begin{aligned} A(z) &:= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}, \\ B(z) &:= b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}. \end{aligned}$$

Assume that n_a and n_b are known, and $y(t) = 0$, $u(t) = 0$ and $v(t) = 0$ for $t \leqslant 0$.

Define the parameter vectors:

$$\begin{aligned} a &:= [a_1, a_2, \cdots, a_{n_a}]^{\mathrm{T}} \in \mathbb{R}^{n_a}, \\ b &:= [b_1, b_2, \cdots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^{n_b}, \end{aligned}$$

and the corresponding information vectors:

$$\begin{aligned} \varphi_a(t) &:= [-y(t-1), -y(t-2), \cdots, -y(t-n_a)]^{\mathrm{T}} \in \mathbb{R}^{n_a}, \\ \varphi_b(t) &:= [u(t-1), u(t-2), \cdots, u(t-n_b)]^{\mathrm{T}} \in \mathbb{R}^{n_b}. \end{aligned}$$

Through the above definitions, the system in (1) can be rewritten as

$$\begin{aligned} y(t) &= [1 - A(z)]y(t) + B(z)u(t) + v(t) \\ &= \varphi_a^{\mathrm{T}}(t)a + \varphi_b^{\mathrm{T}}(t)b + v(t). \end{aligned} \tag{2}$$

For the identification model in (2), the commonly used method is to combine the information vectors $\varphi_a(t)$ and $\varphi_b(t)$ into a large information vector $\varphi(t) := \begin{bmatrix} \varphi_a(t) \\ \varphi_b(t) \end{bmatrix}$, and to define a corresponding parameter vector. Then equation (2) can be transformed to the linear autoregressive identification model. Here, we provide another idea to dispose the identification model in (2) by adopting the hierarchical identification principle. Hierarchical identification is the decomposition based identification. The key idea is to decompose the identification model into several subsystems, such that the scale of the optimization problem can be decreased and the computational efficiency of the identification algorithm can be improved.

Define two fictitious output variables:

$$\begin{aligned} y_1(t) &:= y(t) - \varphi_b^{\mathrm{T}}(t)b \in \mathbb{R}, \\ y_2(t) &:= y(t) - \varphi_a^{\mathrm{T}}(t)a \in \mathbb{R}. \end{aligned}$$

Then the identification model in (2) can be decomposed into two sub-identification models:

$$y_1(t) = \varphi_a^{\mathrm{T}}(t)a + v(t), \tag{3}$$

$$y_2(t) = \varphi_b^{\mathrm{T}}(t)b + v(t). \tag{4}$$

Equations (3)-(4) can also be called the two-stage identification model for the system in (1). Equation (3) contains the parameter vector a and the information vector $\varphi_a(t)$, and (4) includes the parameter vector b and the information vector $\varphi_b(t)$. Here, we can find that there are two coupled variables a and b between these two sub-identification models. The objective of this paper is to use the gradient search principle and the multi-innovation identification theory to propose two-stage algorithms, coordinating the associate items between sub-identification models based on the hierarchical identification principle.

3. THE TWO-STAGE GRADIENT-BASED ITERATIVE ALGORITHM

The iterative algorithm uses batch data to update the parameter estimates. In this section, we define two cost functions to present a two-stage gradient-based iterative algorithm by using the gradient search principle. In addition, we give a brief discussion about how to choose the iterative step-sizes in the algorithm.

Let L be the data length. According to the two-stage identification model in (3)-(4), define two gradient criterion functions as

$$J_1(a) := \frac{1}{2}\sum_{j=1}^{L}[y_1(j) - \varphi_a^{\mathrm{T}}(j)a]^2,$$

$$J_2(b) := \frac{1}{2}\sum_{j=1}^{L}[y_2(j)-\varphi_b^{\mathrm{T}}(j)b]^2.$$

Define the stacked output vector $Y(L)$, the stacked fictitious output vectors $Y_1(L)$ and $Y_2(L)$, and the stacked information matrices $\Phi_a(L)$ and $\Phi_b(L)$ as

$$Y(L) := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(L) \end{bmatrix} \in \mathbb{R}^L,$$

$$Y_1(L) := \begin{bmatrix} y_1(1) \\ y_1(2) \\ \vdots \\ y_1(L) \end{bmatrix} = Y(L)-\Phi_b(L)b \in \mathbb{R}^L,$$

$$Y_2(L) := \begin{bmatrix} y_2(1) \\ y_2(2) \\ \vdots \\ y_2(L) \end{bmatrix} = Y(L)-\Phi_a(L)a \in \mathbb{R}^L,$$

$$\Phi_a(L) := \begin{bmatrix} \varphi_a^{\mathrm{T}}(1) \\ \varphi_a^{\mathrm{T}}(2) \\ \vdots \\ \varphi_a^{\mathrm{T}}(L) \end{bmatrix} \in \mathbb{R}^{L\times n_a},$$

$$\Phi_b(L) := \begin{bmatrix} \varphi_b^{\mathrm{T}}(1) \\ \varphi_b^{\mathrm{T}}(2) \\ \vdots \\ \varphi_b^{\mathrm{T}}(L) \end{bmatrix} \in \mathbb{R}^{L\times n_b}.$$

Then the cost functions $J_1(a)$ and $J_2(b)$ can be equivalently expressed as

$$J_3(a) := \frac{1}{2}\|Y_1(L)-\Phi_a(L)a\|^2,$$
$$J_4(b) := \frac{1}{2}\|Y_2(L)-\Phi_b(L)b\|^2.$$

Let $k = 1,2,3,\cdots$ be an iterative variable, $\hat{a}_k \in \mathbb{R}^{n_a}$ and $\hat{b}_k \in \mathbb{R}^{n_b}$ be the iterative estimates of a and b at iteration k, and $\mu_1 \geqslant 0$ and $\mu_2 \geqslant 0$ be the step-sizes, which are also known as the convergence factors. Using the negative gradient search to minimize the criterion function $J_3(a)$ and $J_4(b)$ results in the following gradient-based iterative relations,

$$\begin{aligned}
\hat{a}_k =& \hat{a}_{k-1}-\mu_1\mathrm{grad}[J_3(\hat{a}_{k-1})] \\
=& \hat{a}_{k-1}+\mu_1\Phi_a^{\mathrm{T}}(L)[Y_1(L)-\Phi_a(L)\hat{a}_{k-1}] \\
=& \hat{a}_{k-1}+\mu_1\Phi_a^{\mathrm{T}}(L) \\
& \times [Y(L)-\Phi_b(L)b-\Phi_a(L)\hat{a}_{k-1}] \qquad (5) \\
=& [I_{n_a}-\mu_1\Phi_a^{\mathrm{T}}(L)\Phi_a(L)]\hat{a}_{k-1} \\
& +\mu_1\Phi_a^{\mathrm{T}}(L)[Y(L)-\Phi_b(L)b], \qquad (6) \\
\hat{b}_k =& \hat{b}_{k-1}-\mu_2\mathrm{grad}[J_4(\hat{b}_{k-1})] \\
=& \hat{b}_{k-1}+\mu_2\Phi_b^{\mathrm{T}}(L)[Y_2(L)-\Phi_b(L)\hat{b}_{k-1}] \\
=& \hat{b}_{k-1}+\mu_2\Phi_b^{\mathrm{T}}(L) \\
& \times [Y(L)-\Phi_a(L)a-\Phi_b(L)\hat{b}_{k-1}] \qquad (7) \\
=& [I_{n_b}-\mu_2\Phi_b^{\mathrm{T}}(L)\Phi_b(L)]\hat{b}_{k-1} \\
& +\mu_2\Phi_b^{\mathrm{T}}(L)[Y(L)-\Phi_a(L)a]. \qquad (8)
\end{aligned}$$

Equations (6) and (8) can be seen as the discrete-time systems of the state variables $\hat{a}_k$ and $\hat{b}_k$. In order to make sure of the convergence of the parameter estimation vectors $\hat{a}_k$ and $\hat{b}_k$, all the eigenvalues of matrices $[I_{n_a}-\mu_1\Phi_a^{\mathrm{T}}(L)\Phi_a(L)]$ and $[I_{n_b}-\mu_2\Phi_b^{\mathrm{T}}(L)\Phi_b(L)]$ must be in the unit circle. In other words, μ_1 and μ_2 should satisfy

$$-I_{n_a} \leq I_{n_a}-\mu_1\Phi_a^{\mathrm{T}}(L)\Phi_a(L) \leq I_{n_a},$$
$$-I_{n_b} \leq [I_{n_b}-\mu_2\Phi_b^{\mathrm{T}}(L)\Phi_b(L)] \leq I_{n_b}.$$

Therefore, one conservative choice is

$$\mu_1 \leq \frac{2}{\lambda_{\max}[\Phi_a^{\mathrm{T}}(L)\Phi_a(L)]} = 2\lambda_{\max}^{-1}[\Phi_a^{\mathrm{T}}(L)\Phi_a(L)], \quad (9)$$
$$\mu_2 \leq \frac{2}{\lambda_{\max}[\Phi_b^{\mathrm{T}}(L)\Phi_b(L)]} = 2\lambda_{\max}^{-1}[\Phi_b^{\mathrm{T}}(L)\Phi_b(L)]. \quad (10)$$

Because of the complexity of the eigenvalue computation, the convergence factors can also be simply and conservatively taken as

$$\mu_1 \leq \frac{2}{\|\Phi_a(L)\|^2} = 2\|\Phi_a(L)\|^{-2}, \qquad (11)$$
$$\mu_2 \leq \frac{2}{\|\Phi_b(L)\|^2} = 2\|\Phi_b(L)\|^{-2}. \qquad (12)$$

Remark 1: Compared with that of (9) and (10), the computation of the convergence factors μ_1 and μ_2 in (11) and (12) does not involve the complicated eigenvalues of square matrices. Instead, we use the traces of the matrices $\Phi_a(L)$ and $\Phi_b(L)$, such that the computational cost of finding the convergence factors can be reduced.

However, equations (6), (8) and (9)-(12) still cannot generate the estimates $\hat{a}_k$ and $\hat{b}_k$. Because there are unknown variables b and a on the right-hand side of (6) and (8). To cope with this problem, we use the hierarchical identification principle to coordinate the associated terms. That is, use the estimates $\hat{b}_{k-1}$ and $\hat{a}_{k-1}$ at iteration $k-1$ to replace the unknown variables b and a in (6) and (8). Combining (9)-(12), we can derive a two-stage gradient-based iterative (2S-GI) algorithm to estimate the parameter vectors a and b for the CAR system:

$$\begin{aligned}
\hat{a}_k =& \hat{a}_{k-1}+\mu_1\Phi_a^{\mathrm{T}}(L) \\
& \times [Y(L)-\Phi_a(L)\hat{a}_{k-1}-\Phi_b(L)\hat{b}_{k-1}], \qquad (13) \\
\mu_1 =& \lambda_{\max}^{-1}[\Phi_a^{\mathrm{T}}(L)\Phi_a(L)], \qquad (14) \\
\hat{b}_k =& \hat{b}_{k-1}+\mu_2\Phi_b^{\mathrm{T}}(L) \\
& \times [Y(L)-\Phi_a(L)\hat{a}_{k-1}-\Phi_b(L)\hat{b}_{k-1}], \qquad (15)
\end{aligned}$$

$$\mu_2 = \lambda_{\max}^{-1}[\Phi_b^{\mathrm{T}}(L)\Phi_b(L)], \tag{16}$$
$$Y(L) = [y(1), y(2), \cdots, y(L)]^{\mathrm{T}}, \tag{17}$$
$$\Phi_a(L) = [\varphi_a(1), \varphi_a(2), \cdots, \varphi_a(L)]^{\mathrm{T}}, \tag{18}$$
$$\Phi_b(L) = [\varphi_b(1), \varphi_b(2), \cdots, \varphi_b(L)]^{\mathrm{T}}, \tag{19}$$
$$\varphi_a(t) = [-y(t-1), -y(t-2), \cdots, -y(t-n_a)]^{\mathrm{T}}, \tag{20}$$
$$\varphi_b(t) = [u(t-1), u(t-2), \cdots, u(t-n_b)]^{\mathrm{T}}, \tag{21}$$
$$\hat{a}_k = [\hat{a}_{1,k}, \hat{a}_{2,k}, \cdots, \hat{a}_{n_a,k}]^{\mathrm{T}}, \tag{22}$$
$$\hat{b}_k = [\hat{b}_{1,k}, \hat{b}_{2,k}, \cdots, \hat{b}_{n_b,k}]^{\mathrm{T}}. \tag{23}$$

The steps of computing $\hat{a}_k$ and $\hat{b}_k$ involved in the 2S-GI algorithm in (13)-(23) are summarized in the following:

1) For $t \leqslant 0$, all the variables are set to be zero. Let $k = 1$, give the data length L ($L \gg n_a + n_b$) and set the initial values: $\hat{a}_0 = \mathbf{1}_{n_a}/p_0$, $\hat{b}_0 = \mathbf{1}_{n_b}/p_0$, $p_0 = 10^6$, and the parameter estimation accuracy ε.
2) Collect the input and output data $u(t)$ and $y(t)$, $t = 1, 2, \cdots, L$. Form the information vectors $\varphi_a(t)$ and $\varphi_b(t)$ using (20)-(21).
3) Construct the stacked output vector $Y(L)$ by (17) and the stacked information matrices $\Phi_a(L)$ and $\Phi_b(L)$ by (18)-(19), compute the step-sizes μ_1 and μ_2 according to (14) and (16).
4) Update the parameter estimation vectors $\hat{a}_k$ and $\hat{b}_k$ by using (13) and (15). Read out the estimates $\hat{a}_{i,k}$ ($i = 1, 2, \cdots, n_a$) and $\hat{b}_{j,k}$ ($j = 1, 2, \cdots, n_b$) from (22) and (23).
5) If $\|\hat{a}_k - \hat{a}_{k-1}\| + \|\hat{b}_k - \hat{b}_{k-1}\| > \varepsilon$, increase k by 1 and go to Step 4; otherwise, obtain the iteration k and the parameter estimation vectors $\hat{a}_k$ and $\hat{b}_k$, terminate this computational procedure.

Remark 2: The 2S-GI algorithm in (13)-(23) can also be called the hierarchical gradient-based iterative (HGI) algorithm. The difference is that the 2S-GI algorithm is just one kind of the HGI algorithm. The three-stage gradient-based iterative algorithm and the multi-stage gradient-based iterative algorithm are all known as the HGI algorithm.

4. THE TWO-STAGE MULTI-INNOVATION GRADIENT-BASED ITERATIVE ALGORITHM

The main idea of the multi-innovation identification theory is to use the data in a moving data window to update the parameter estimates. In addition, the moving data window moves forward as time t increases. In this section, we adopt the multi-innovation identification theory to deal with the CAR system in (1) and propose a two-stage multi-innovation gradient-based iterative algorithm which has a constant window length.

Consider the newest p data from $j = t - p + 1$ to $j = t$ ($p \gg n_a + n_b$), and define the stacked output vector $Y(p,t)$, the stacked fictitious output vectors $Y_1(p,t)$ and $Y_2(p,t)$, and the stacked information matrices $\Phi_a(p,t)$ and $\Phi_b(p,t)$:

$$Y(p,t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \in \mathbb{R}^p,$$

$$Y_1(p,t) := \begin{bmatrix} y_1(t) \\ y_1(t-1) \\ \vdots \\ y_1(t-p+1) \end{bmatrix} = Y(p,t) - \Phi_b(p,t)b \in \mathbb{R}^p,$$

$$Y_2(p,t) := \begin{bmatrix} y_2(t) \\ y_2(t-1) \\ \vdots \\ y_2(t-p+1) \end{bmatrix} = Y(p,t) - \Phi_a(p,t)a \in \mathbb{R}^p,$$

$$\Phi_a(p,t) := \begin{bmatrix} \varphi_a^{\mathrm{T}}(t) \\ \varphi_a^{\mathrm{T}}(t-1) \\ \vdots \\ \varphi_a^{\mathrm{T}}(t-p+1) \end{bmatrix} \in \mathbb{R}^{p\times n},$$

$$\Phi_b(p,t) := \begin{bmatrix} \varphi_b^{\mathrm{T}}(t) \\ \varphi_b^{\mathrm{T}}(t-1) \\ \vdots \\ \varphi_b^{\mathrm{T}}(t-p+1) \end{bmatrix} \in \mathbb{R}^{p\times n_b}.$$

According to the two-stage identification models in (3)-(4), define two criterion functions:

$$\begin{aligned} J_5(a) &:= \frac{1}{2}\sum_{j=t-p+1}^{t}[y_1(j) - \varphi_a^{\mathrm{T}}(j)a]^2 \\ &= \frac{1}{2}\|Y_1(p,t) - \Phi_a(p,t)a\|^2, \\ J_6(b) &:= \frac{1}{2}\sum_{j=t-p+1}^{t}[y_2(j) - \psi^{\mathrm{T}}(j)b]^2 \\ &= \frac{1}{2}\|Y_2(p,t) - \Phi_b(p,t)b\|^2. \end{aligned}$$

Let $\mu_1(t) \geqslant 0$ and $\mu_2(t) \geqslant 0$ be the step-sizes. Using the negative gradient search to minimize the criterion functions $J_5(a)$ and $J_6(b)$ gives the following gradient-based iterative relations:

$$\begin{aligned} \hat{a}_k(t) =& \hat{a}_{k-1}(t) - \mu_1(t)\mathrm{grad}[J_5(\hat{a}_{k-1}(t))] \\ =& \hat{a}_{k-1}(t) + \mu_1(t)\Phi_a^{\mathrm{T}}(p,t) \\ & \times [Y_1(p,t) - \Phi_a(p,t)\hat{a}_{k-1}(t)] \\ =& \hat{a}_{k-1}(t) + \mu_1(t)\Phi_a^{\mathrm{T}}(p,t) \\ & \times [Y(p,t) - \Phi_b(p,t)b - \Phi_a(p,t)\hat{a}_{k-1}(t)] \\ =& [I_{n_a} - \mu_1(t)\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)]\hat{a}_{k-1}(t) \end{aligned} \tag{24}$$

$$+\mu_1(t)\Phi_a^{\mathrm{T}}(p,t)[Y(p,t)-\Phi_b(p,t)b], \tag{25}$$

$$\begin{aligned}\hat{b}_k(t)=&\hat{b}_{k-1}(t)-\mu_2(t)\mathrm{grad}[J_6(\hat{b}_{k-1}(t))]\\ =&\hat{b}_{k-1}(t)+\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)\\ &\times[Y_2(p,t)-\Phi_b(p,t)\hat{b}_{k-1}(t)]\\ =&\hat{b}_{k-1}(t)+\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)\\ &\times[Y(p,t)-\Phi_a(p,t)a-\Phi_b(p,t)\hat{b}_{k-1}(t)] \quad (26)\\ =&[I_{n_b}-\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)]\hat{b}_{k-1}(t)\\ &+\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)[Y(p,t)-\Phi_a(p,t)a]. \quad (27)\end{aligned}$$

Similarly to the derivation of the 2S-GI algorithm in (13)-(23), Equations (25) and (27) can be seen as two discrete-time systems of the state variables $\hat{a}_k(t)$ and $\hat{b}_k(t)$. In order to make sure of the convergence of the parameter estimation vectors $\hat{a}_k$ and $\hat{b}_k$, all the eigenvalues of matrices $[I_{n_a}-\mu_1(t)\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)]$ and $[I_{n_b}-\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)]$ must be in the unit circle, that is to say, $\mu_1(t)$ and $\mu_2(t)$ must satisfy

$$\begin{aligned}&-I_{n_a}\leqslant I_{n_a}-\mu_1(t)\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)\leqslant I_{n_a},\\ &-I_{n_b}\leqslant I_{n_b}-\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)\leqslant I_{n_b}.\end{aligned}$$

Therefore, the step-sizes can be conservatively chosen as

$$\begin{aligned}\mu_1(t)&\leqslant\frac{2}{\lambda_{\max}[\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)]}\\ &=2\lambda_{\max}^{-1}[\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)], \quad (28)\\ \mu_2(t)&\leqslant\frac{2}{\lambda_{\max}[\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)]}\\ &=2\lambda_{\max}^{-1}[\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)]. \quad (29)\end{aligned}$$

In consideration of the computational complexity, we take the step-sizes as

$$\mu_1(t)\leqslant\frac{2}{\|\Phi_a(p,t)\|^2}=2\|\Phi_a(p,t)\|^{-2}, \tag{30}$$

$$\mu_2(t)\leqslant\frac{2}{\|\Phi_b(p,t)\|^2}=2\|\Phi_b(p,t)\|^{-2}. \tag{31}$$

Remark 3: Similar to (11) and (12), we use the traces of the matrices $\Phi_a(p,t)$ and $\Phi_b(p,t)$ in (30) and (31) to replace the complicated eigenvalues of square matrices in computing the convergence factors $\mu_1(t)$ and $\mu_2(t)$, so as to reduce the computational cost of finding the convergence factors.

To obtain a realizable algorithm, we replace the unknown parameters b and a in (24) and (26) with their estimates $\hat{b}_{k-1}(t)$ and $\hat{a}_{k-1}(t)$. Combining (28) to (31), we can obtain the two-stage multi-innovation gradient-based iterative (2S-MIGI) algorithm for estimating parameter vectors a and b:

$$\begin{aligned}\hat{a}_k(t)=&\hat{a}_{k-1}(t)+\mu_1(t)\Phi_a^{\mathrm{T}}(p,t)[Y(p,t)\\ &-\Phi_a(p,t)\hat{a}_{k-1}(t)-\Phi_b(p,t)\hat{b}_{k-1}(t)], \quad (32)\end{aligned}$$

$$\begin{aligned}&\mu_1(t)=\lambda_{\max}^{-1}[\Phi_a^{\mathrm{T}}(p,t)\Phi_a(p,t)] \ \text{ or}\\ &\mu_1(t)=\|\Phi_a(p,t)\|^{-2}, \quad (33)\\ &\hat{b}_k(t)=\hat{b}_{k-1}(t)+\mu_2(t)\Phi_b^{\mathrm{T}}(p,t)[Y(p,t)\\ &\qquad-\Phi_a(p,t)\hat{a}_{k-1}(t)-\Phi_b(p,t)\hat{b}_{k-1}(t)], \quad (34)\\ &\mu_2(t)=\lambda_{\max}^{-1}[\Phi_b^{\mathrm{T}}(p,t)\Phi_b(p,t)] \ \text{ or}\\ &\mu_2(t)=\|\Phi_b(p,t)\|^{-2}, \quad (35)\\ &Y(p,t)=[y(t),y(t-1),\cdots,y(t-p+1)]^{\mathrm{T}}, \quad (36)\\ &\Phi_a(p,t)=[\varphi_a(t),\varphi_a(t-1),\cdots,\varphi_a(t-p+1)]^{\mathrm{T}}, \quad (37)\\ &\Phi_b(p,t)=[\varphi_b(t),\varphi_b(t-1),\cdots,\varphi_b(t-p+1)]^{\mathrm{T}}, \quad (38)\\ &\varphi_a(t)=[-y(t-1),\cdots,-y(t-n_a)]^{\mathrm{T}}, \quad (39)\\ &\varphi_b(t)=[u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}, \quad (40)\\ &\hat{a}_k(t)=[\hat{a}_{1,k}(t),\hat{a}_{2,k}(t),\cdots,\hat{a}_{n_a,k}(t)]^{\mathrm{T}}, \quad (41)\\ &\hat{b}_k(t)=[\hat{b}_{1,k}(t),\hat{b}_{2,k}(t),\cdots,\hat{b}_{n_b,k}(t)]^{\mathrm{T}}. \quad (42)\end{aligned}$$

If we take $p=t=L$, then the 2S-MIGI algorithm in (32)-(42) reduces to 2S-GI algorithm in (13)-(23).

The procedure of computing the parameter estimation vectors $\hat{a}_k(t)$ and $\hat{b}_k(t)$ by the 2S-MIGI algorithm in (32)-(42) is listed as follows:

1) For $t\leqslant 0$, all the variables are set to to be zero. Let $k=1$, give the moving data window length p ($p\gg n_a+n_b$) and set the initial values: $\hat{a}_0=\mathbf{1}_{n_a}/p_0$, $\hat{b}_0=\mathbf{1}_{n_b}/p_0$, $p_0=10^6$, the maximum iteration $k_{\max}$ and ε.
2) Let $k=1$, collect the input and output data $u(t)$ and $y(t)$, form the information vectors $\varphi_a(t)$ and $\varphi_b(t)$ using (39)-(40).
3) Construct the stacked output vector $Y(p,t)$ by (36) and the stacked information matrices $\Phi_a(p,t)$ and $\Phi_b(p,t)$ by using (37)-(38).
4) Compute the step-sizes $\mu_1(t)$ and $\mu_2(t)$ according to (33) and (35).
5) Update the parameter estimation vectors $\hat{a}_k(t)$ and $\hat{b}_k(t)$ by using (32) and (34). Read out the estimates $\hat{a}_{i,k}(t)$ and $\hat{b}_{j,k}(t)$ from (41) and (42).
6) If $k<k_{\max}$, increase k by 1 and go to Step 5; otherwise, proceed to the next step.
7) Compare $\hat{a}_k(t)$ and $\hat{b}_k(t)$ with $\hat{a}_{k-1}(t)$ and $\hat{b}_{k-1}(t)$: If $\|\hat{a}_k(t)-\hat{a}_{k-1}(t)\|+\|\hat{b}_k(t)-\hat{b}_{k-1}(t)\|>\varepsilon$, then set $\hat{a}_0(t+1):=\hat{a}_k(t)$ and $\hat{b}_0(t+1):=\hat{b}_k(t)$, increase t by 1 and go to Step 2; otherwise, obtain the parameter estimation vectors $\hat{a}_k(t)$ and $\hat{b}_k(t)$, terminate this computational procedure.

Remark 4: The 2S-MIGI algorithm both have the time variable t and the iterative variable k. At time t, we collect p data in a moving data window with the fixed length and compute the parameter estimates $\hat{a}_k(t)$ and $\hat{b}_k(t)$ with k increasing. If there is no significant change in the parameter estimates with increasing k, we increase t rather than k and introduce new data into the algorithm.

Remark 5: Compared with the 2S-GI algorithm in (13)-(23), the 2S-MIGI algorithm can track time-varying parameters. Because the 2S-MIGI expands the dimensions of the innovation and makes full use of the system information. For $k=1,2,\cdots,k_{\max}-1$, we use the newest p data from $j=t-p+1$ to $j=t$ to estimate the parameter vectors. For $k=k_{\max}$, we move the data window forward to the next moment, introduce new observation data and eliminate the oldest data to keep p data in the data window. Based on the above idea, we repeat the iterative process until we obtain the satisfactory parameter estimates.

At each iterative estimation, both the 2S-GI and the 2S-MIGI algorithms use the data with the fixed length to identify the unknown parameters, and the iterative computing continues with increasing iteration variable k. The difference is that, the iterative computing of the 2S-GI algorithm stops when $k=k_{\max}$, where $k_{\max}$ is the iterations for desirable parameter estimation accuracy; for the 2S-MIGI algorithm, when $k=k_{\max}$, the iterative estimation will continue with the data window moved forward to the next moment. This is also the reason why the 2S-MIGI algorithm can track time-varying parameters compared with the 2S-GI algorithm. The 2S-MIGI algorithm is the extension of the 2S-GI algorithm, or the 2S-GI algorithm is the special case of the 2S-MIGI algorithm.

5. THE GRADIENT-BASED ITERATIVE ALGORITHM

In order to show that the proposed 2S-GI algorithm can generate highly accurate parameter estimates, the following gives the gradient-based iterative (GI) algorithm for estimating the parameter vector $\vartheta:=\begin{bmatrix} a \\ b \end{bmatrix}$ of the CAR system in (1):

$$\hat{\vartheta}_k=\hat{\vartheta}_{k-1}+\mu\Phi^{\mathrm{T}}(L)[Y(L)-\Phi(L)\hat{\vartheta}_{k-1}], \tag{43}$$

$$\mu=\lambda_{\max}^{-1}[\Phi^{\mathrm{T}}(L)\Phi(L)], \quad \text{or} \quad \mu=\|\Phi(L)\|^{-2}, \tag{44}$$

$$Y(L)=[y(1),y(2),\cdots,y(L)]^{\mathrm{T}}, \tag{45}$$

$$\Phi(L)=[\varphi(1),\varphi(2),\cdots,\varphi(L)]^{\mathrm{T}}, \tag{46}$$

$$\varphi(t)=\begin{bmatrix} \varphi_a(t) \\ \varphi_b(t) \end{bmatrix}, \tag{47}$$

$$\varphi_a(t)=[-y(t-1),-y(t-2),\cdots,-y(t-n_a)]^{\mathrm{T}}, \tag{48}$$

$$\varphi_b(t)=[u(t-1),u(t-2),\cdots,u(t-n_b)]^{\mathrm{T}}. \tag{49}$$

$$\hat{\vartheta}_k=\begin{bmatrix} \hat{a}_k \\ \hat{b}_k \end{bmatrix}, \tag{50}$$

$$\hat{a}_k=[\hat{a}_{1,k},\hat{a}_{2,k},\cdots,\hat{a}_{n_a,k}]^{\mathrm{T}}, \tag{51}$$

$$\hat{b}_k=[\hat{b}_{1,k},\hat{b}_{2,k},\cdots,\hat{b}_{n_b,k}]^{\mathrm{T}}. \tag{52}$$

The 2S-GI algorithm can generate more accurate parameter estimates than the GI algorithm. The proposed 2s-GI algorithm and 2S-MIGI algorithm for controlled autoregressive systems can combine other estimation algorithms [29–31] to explore new identification methods of different systems with colored noises [32–37] and can be applied to other fields such as information processing and communication [38–44].

6. EXAMPLE

Consider the following second-order controlled autoregressive system:

$$\begin{aligned} &A(z)y(t)=B(z)u(t)+v(t),\\ &A(z)=1+a_1z^{-1}+a_2z^{-2}=1+1.35z^{-1}+0.75z^{-2},\\ &B(z)=b_1z^{-1}+b_2z^{-2}=1.68z^{-1}+2.32z^{-2}. \end{aligned}$$

The parameter vector to be identified is given by

$$\vartheta=[a_1,a_2,b_1,b_2]^{\mathrm{T}}=[1.35,0.75,1.68,2.32]^{\mathrm{T}}.$$

In simulation, the input $\{u(t)\}$ is taken as an uncorrelated stochastic signal sequence with zero mean and unit variance, $\{v(t)\}$ is taken as a white noise sequence with zero mean and variance σ^2. We use the example parameters and the input signal generates the output signal sequence $\{y(t)\}$.

Taking the noise variances $\sigma^2=0.50^2$ and $\sigma^2=1.50^2$, respectively, the corresponding noise-to-signal ratios are $\delta_{\mathrm{ns}}=33.78\%$ and $\delta_{\mathrm{ns}}=101.33\%$. Taking the data length $L=3000$, applying the GI and 2S-GI algorithms and the input-output data $\{u(t),y(t):\ t=1,2,\cdots,L\}$ to estimate the parameters of this example system, the GI and 2S-GI parameter estimates and errors versus k are shown in Tables 1 to 4, the GI and 2S-GI estimation errors $\delta:=\|\hat{\vartheta}(t)-\vartheta\|/\|\vartheta\|$ versus k are shown in Figs. 1 and 2 with $\sigma^2=0.50^2$ and $\sigma^2=1.50^2$, and the GI and 2S-GI estimates of the parameters a_1, a_2, b_1 and b_2 versus k are shown in Figs. 3 and 4 for $\sigma^2=1.50^2$.

From Tables 1-4 and Figs. 1-4, we draw the following conclusions:

- As the noise levels decrease, the GI and 2S-GI algorithms can give more accurate parameter estimates - see and compare the parameter estimation errors in the last columns in Tables 1 and 2, and in Tables 3 and 4.
- The 2S-GI parameter estimation errors tend to zero faster than the GI parameter estimation errors - see the GI and 2S-GI estimation error curves in Figs. 1 and 2.
- Under the same noise levels, the 2S-GI algorithm can give more accurate parameter estimates than the GI algorithm - see the parameter estimation errors in the last columns in Tables 1 and 3 and in Tables 2 and 4, and compare the parameter estimation error curves in Figs. 1 and 2.

Table 1. The GI estimates and their errors with $\sigma^2 = 0.50^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.37465	-0.04311	0.07643	0.00366	94.80048022
2	0.50048	0.08091	0.15063	0.03563	90.77808869
5	0.73119	0.30193	0.35429	0.17235	81.03294080
10	0.89670	0.44666	0.63673	0.43991	68.16606805
20	1.03276	0.54489	1.03527	0.91685	48.85155929
50	1.22057	0.66760	1.53018	1.74466	18.86833601
100	1.31967	0.73236	1.67003	2.18910	4.17567060
150	1.34189	0.74693	1.68257	2.28945	0.97914391
200	1.34690	0.75022	1.68366	2.31211	0.28382179
True values	1.35000	0.75000	1.68000	2.32000	

Table 2. The GI estimates and their errors with $\sigma^2 = 1.50^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.40419	-0.09774	0.03871	0.00127	95.62556223
2	0.52849	0.02605	0.07719	0.01802	92.53172774
5	0.77934	0.27426	0.18776	0.09185	85.51106304
10	0.97476	0.46353	0.35621	0.25142	76.85043759
20	1.09838	0.57410	0.63948	0.58087	62.98815992
50	1.20883	0.65557	1.17798	1.31009	35.04771968
100	1.29146	0.71446	1.53679	1.91203	13.45227896
150	1.32465	0.73817	1.64518	2.15478	5.25923540
200	1.33802	0.74773	1.67770	2.25281	2.09961774
True values	1.35000	0.75000	1.68000	2.32000	

Table 3. The 2S-GI estimates and their errors with $\sigma^2 = 0.50^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.37572	-0.04323	1.61535	0.07731	79.01908360
2	0.51210	0.07990	1.66516	0.66077	60.71998715
5	0.85747	0.34711	1.67069	1.32367	36.32947056
10	1.14245	0.58174	1.67830	1.90179	15.25053891
20	1.31214	0.72137	1.68279	2.24539	2.71856140
50	1.34816	0.75101	1.68374	2.31832	0.14149581
100	1.34836	0.75117	1.68374	2.31872	0.13649796
150	1.34836	0.75117	1.68374	2.31872	0.13649723
200	1.34836	0.75117	1.68374	2.31872	0.13649723
True values	1.35000	0.75000	1.68000	2.32000	

Table 4. The 2S-GI estimates and their errors with $\sigma^2 = 1.50^2$.

k	a_1	a_2	b_1	b_2	δ (%)
1	0.40448	-0.09781	1.64023	0.05364	79.84208003
2	0.53307	0.02506	1.69099	0.68996	60.29629113
5	0.84810	0.29961	1.68855	1.30870	37.35311552
10	1.12344	0.55047	1.69009	1.86655	16.74281499
20	1.30215	0.71329	1.69107	2.22842	3.38681701
50	1.34670	0.75388	1.69132	2.31863	0.38362505
100	1.34707	0.75421	1.69132	2.31937	0.38237428
150	1.34707	0.75421	1.69132	2.31937	0.38237419
200	1.34707	0.75421	1.69132	2.31937	0.38237419
True values	1.35000	0.75000	1.68000	2.32000	

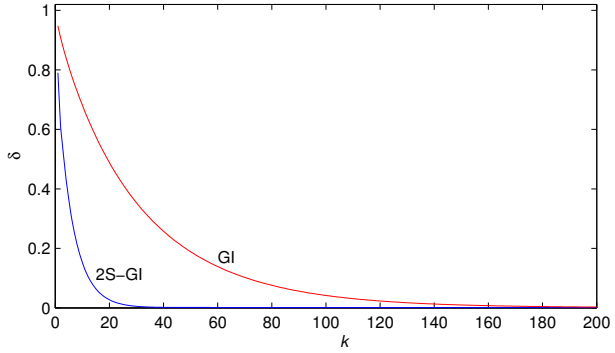


Fig. 1. The GI and 2S-GI estimation errors δ versus k with $\sigma^2 = 0.50^2$.

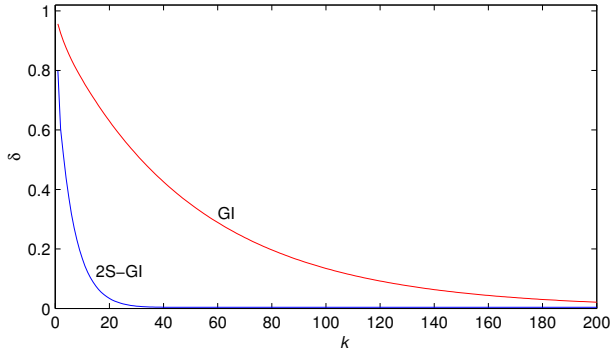


Fig. 2. The GI and 2S-GI estimation errors δ versus k with $\sigma^2 = 1.50^2$.

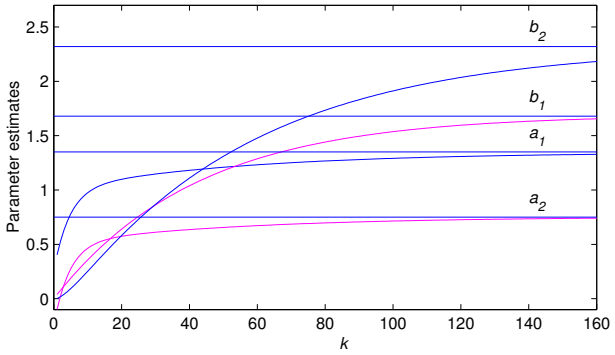


Fig. 3. The GI estimates versus k with $\sigma^2 = 1.50^2$.

- The 2S-GI parameter estimates approach to their true values faster the GI parameter estimates - see and compare the parameter estimates versus k in Figs. 3 and 4.

7. CONCLUSIONS

In this paper, the identification problems of controlled autoregressive systems have been studied. By adopting the hierarchical identification principle, we define two fictitious output variables and decompose the original system into two subsystems. Then a two-stage gradient-based iterative algorithm is proposed based on the gradient search principle. By means of the multi-innovation identification theory, we propose a two-stage multi-innovation gradient-based iterative algorithm. The simulation results indicate that the proposed algorithms are effective. In addition, the 2S-MIGI algorithm can track time-varying parameters. The proposed methods proposed in this paper can combine other tools and strategies [45–54] to study the parameter estimation problems of time-varying systems, nonlinear systems and multivariable systems, and can be applied to other literatures [55–65] such as system identification [66-69].

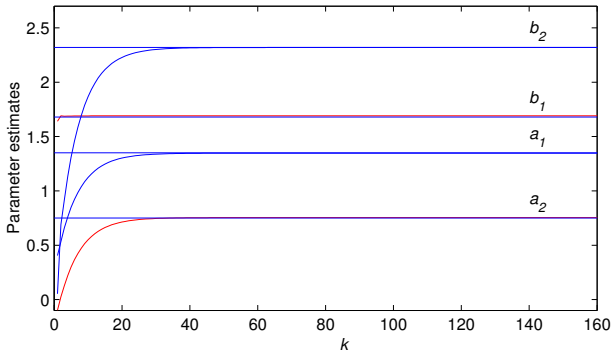


Fig. 4. The 2S-GI estimates versus k with $\sigma^2 = 1.50^2$.

REFERENCES

[1] J. Ding, J. Z. Chen, J. X. Lin, and L. J. Wan, "Particle filtering based parameter estimation for systems with output-error type model structures," *Journal of the Franklin Institute*, vol. 356, no. 10, pp. 5521-5540, July 2019.

[2] J. Ding, J. Z. Chen, J. X. Lin, and G. P. Jiang, "Particle filtering-based recursive identification for controlled autoregressive systems with quantised output," *IET Control Theory and Applications*, vol. 13, no. 14, pp. 2181-2187, 2019.

[3] G. Y. Chen, M. Gan, C. L. P. Chen, and H. X. Li, "A regularized variable projection algorithm for separable nonlinear least-squares problems," *IEEE Transactions on Automatic Control*, vol. 64, no. 2, pp. 526-537, February 2019.

[4] M. Gan, X. X. Chen, F. Ding, G. Y. Chen, and C. L. P. Chen, "Adaptive RBF-AR models based on multi-innovation least squares method," *IEEE Signal Processing Letters*, vol. 26, no. 8, pp. 1182-1186, August 2019.

[5] L. Xu, "A proportional differential control method for a time-delay system using the Taylor expansion approximation," *Applied Mathematics and Computation*, vol. 236, pp. 391-399, June 2014.

[6] L. Xu, "Application of the Newton iteration algorithm to the parameter estimation for dynamical systems," *Journal of Computational and Applied Mathematics*, vol. 288, pp. 33-43, November 2015.

[7] L. Xu, L. Chen, and W. L. Xiong, "Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration," *Nonlinear Dynamics*, vol. 79, no. 3, pp. 2155-2163, February 2015.

[8] L. Xu, "The damping iterative parameter identification method for dynamical systems based on the sine signal measurement," *Signal Processing*, vol. 120, pp. 660-667, March 2016.

[9] M. Gan, C. L. P. Chen, G. Y. Chen, and L. Chen, "On some separated algorithms for separable nonlinear squares problems," *IEEE Transactions on Cybernetics*, vol. 48, no. 10, pp. 2866-2874, October 2018.

[10] L. Xu, "The parameter estimation algorithms based on the dynamical response measurement data," *Advances in Mechanical Engineering*, vol. 9, no. 11, pp. 1-12, November 2017.

[11] L. Xu, W. L. Xiong, A. Alsaedi, and T. Hayat, "Hierarchical parameter estimation for the frequency response based on the dynamical window data," *International Journal of Control Automation and Systems*, vol. 16, no. 4, pp. 1756-1764, August 2018.

[12] S. Y. Liu, F. Ding, L, Xu, and T. Hayat, "Hierarchical principle-based iterative parameter estimation algorithm for dual-frequency signals," *Circuits Systems and Signal Processing*, vol. 38, no. 7, pp. 3251-3268, July 2019.

[13] H. Xu, F. Ding, and E. F. Yang, "Modeling a nonlinear process using the exponential autoregressive time series model," *Nonlinear Dynamics*, vol. 95, no. 3, pp. 2079-2092, February 2019.

[14] J. J. Rubio, E. Lughofer, J. A. Meda-Campaña, L. A. Páramo, J. F. Novoa, and J. Pacheco, "Neural network updating via argument Kalman filter for modeling of Takagi-Sugeno fuzzy models," *Journal of Intelligent and Fuzzy Systems*, vol. 35, no. 2, pp. 2585-2596, August 2018.

[15] Y. P. Pan, T. R. Sun, and H. Y. Yu, "On parameter convergence in least squares identification and adaptive control," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 10, pp 2898-2911, 2019.

[16] J. J. Rubio, "SOFMLS: Online self-organizing fuzzy modified least square network," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 6, pp. 1296-1309, December 2009.

[17] J. J. Rubio, "Stable Kalman filter and neural network for the chaotic systems identification," *Journal of the Franklin Institute*, vol. 354, no. 16, pp. 7444-7462, November 2017.

[18] M. H. Li, X. M. Liu, and F. Ding, "The filtering-based maximum likelihood iterative estimation algorithms for a special class of nonlinear systems with autoregressive moving average noise using the hierarchical identification principle," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1189-1211, July 2019.

[19] M. H. Li and X. M. Liu, "Auxiliary model based least squares iterative algorithms for parameter estimation of bilinear systems using interval-varying measurements," *IEEE Access*, vol. 6, pp. 21518-21529, January 2018.

[20] L. Xu, F. Ding, and Q. M. Zhu, "Hierarchical Newton and least squares iterative estimation algorithm for dynamic systems by transfer functions based on the impulse responses," *International Journal of Systems Science*, vol. 50, no. 1, pp. 141-151, January 2019.

[21] M. H. Li and X. M. Liu, "The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique," *Signal Processing*, vol. 147, pp. 23-34, June 2018.

[22] L. Xu and F. Ding, "The parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle," *IET Signal Processing*, vol. 11, no. 2, pp. 228-237, April 2017.

[23] L. Xu, F. Ding, Y. Gu, A. Alsaedi, and T. Hayat, "A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay," *Signal Processing*, vol. 140, pp. 97-103, November 2017.

[24] L. Xu and F. Ding, "Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling," *Circuits Systems and Signal Processing*, vol. 36, no. 4, pp. 1735-1753, April 2017.

[25] J. Chen, J. Li, and Y. J. Liu, "Gradient iterative algorithm for dual-rate nonlinear systems based on a novel particle filter," *Journal of the Franklin Institute*, vol. 354, no. 11, pp. 4425-4437, July 2017.

[26] L. Xu and F. Ding, "Iterative parameter estimation for signal models based on measured data," *Circuits Systems and Signal Processing*, vol. 37, no. 7, pp. 3046-3069, July 2018.

[27] F. Ding, *System Identification – Iterative Search Principle and Identification Methods*, Science Press, Beijing, 2018.

[28] F. Ding, J. Pan, A. Alsaedi, and T. Hayat, "Gradient-based iterative parameter estimation algorithms for dynamical systems from observation data," *Mathematics*, vol. 7, no. 5, Article Number: 428, May 2019.

[29] Y. J. Wang, F. Ding, and M. H. Wu, "Recursive parameter estimation algorithm for multivariate output-error systems," *Journal of the Franklin Institute*, vol. 355, no. 12, pp. 5163-5181, August 2018.

[30] H. Ma, J. Pan, L. Lv, G. H. Xu, F. Ding, A. Alsaedi, and T. Hayat, "Recursive algorithms for multivariable output-error-like ARMA systems," *Mathematics*, vol. 7, no. 6, Article Number: 558, June 2019.

[31] L. J. Wan and F. Ding, "Decomposition- and gradient-based iterative identification algorithms for multivariable systems using the multi-innovation theory," *Circuits Systems and Signal Processing*, vol. 38, no. 7, pp. 2971-2991, July 2019.

[32] Q. Y. Liu, F. Ding, L. Xu, and E. F. Yang, "Partially coupled gradient estimation algorithm for multivariable equation-error autoregressive moving average systems using the data filtering technique," *IET Control Theory and Applications*, vol. 13, no. 5, pp. 642-650, March 2019.

[33] X. Zhang, F. Ding, L. Xu, and E. F. Yang, "Highly computationally efficient state filter based on the delta operator," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 6, pp. 875-889, June 2019.

[34] X. Zhang, F. Ding, and E. F. Yang, "State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1157-1173, July 2019.

[35] 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 and Applications*, vol. 11, no. 3, pp. 390-400, February 2017.

[36] F. Ding, L. Xu, F. E. Alsaadi, and T. Hayat, "Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique," *IET Control Theory and Applications*, vol. 12, no. 7, pp. 892-899, May 2018.

[37] L. Xu and F. Ding, "Parameter estimation for control systems based on impulse responses," *International Journal of Control Automation and Systems*, vol. 15, no. 6, pp. 2471-2479, December 2017.

[38] M. H. Wu, X. Li, C. Liu, M. Liu, N. Zhao, J. Wang, X. K. Wan, Z. H. Rao, and L. Zhu, "Robust global motion estimation for video security based on improved k-means clustering," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 2, pp. 439-448, February 2019.

[39] L. Feng, Q. X. Li, and Y. F. Li, "Imaging with 3-D aperture synthesis radiometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 2395-2406, April 2019.

[40] B. Fu, C. X. Ouyang, C. S. Li, J. W. Wang, and E. Gul, "An improved mixed integer linear programming approach based on symmetry diminishing for unit commitment of hybrid power system," *Energies*, vol. 12, no. 5, Article Number: 833, March 2019.

[41] W. X. Shi, N. Liu, Y. M. Zhou, and X. A. Cao, "Effects of postannealing on the characteristics and reliability of polyfluorene organic light-emitting diodes," *IEEE Transactions on Electron Devices*, vol. 66, no. 2, pp. 1057-1062, February 2019.

[42] N. Liu, S. Mei, D. Sun, W. Shi, J. Feng, Y. M. Zhou, F. Mei, J. Xu, Y. Jiang, and X. A. Cao, "Effects of charge transport materials on blue fluorescent organic light-emitting diodes with a host-dopant system," *Micromachines*, vol. 10, no. 5, Article Number: 344, May 2019.

[43] T. Z. Wu, X. Shi, L. Liao, C. J. Zhou, H. Zhou, and Y. H. Su, "A capacity configuration control strategy to alleviate power fluctuation of hybrid energy storage system based on improved particle swarm optimization," *Energies*, vol. 12, no. 4, Article Number: 642, February 2019.

[44] X. L. Zhao, Z. Y. Lin, B. Fu, L. He, and C. S. Li, "Research on the predictive optimal PID plus second order derivative method for AGC of power system with high penetration of photovoltaic and wind power," *Journal of Electrical Engineering & Technology*, vol. 14, no. 3, pp. 1075-1086, May 2019.

[45] X. S. Zhan, L. L. Cheng, J. Wu, Q. S. Yang, and T. Han, "Optimal modified performance of MIMO networked control systems with multi-parameter constraints," *ISA Transactions*, vol. 84, no. 1, pp. 111-117, January 2019.

[46] X. S. Zhan, L. L. Cheng, J. Wu, and H. C. Yan, "Modified tracking performance limitation of networked time-delay systems with two-channel constraints," *Journal of the Franklin Institute*, vol. 356, no. 12, pp. 6401-6418, August 2019.

[47] X. D. Liu, H. S. Yu, J. P. Yu, and L. Zhao, "Combined speed and current terminal sliding mode control with nonlinear disturbance observer for PMSM drive," *IEEE Access*, vol. 6, pp. 29594-29601, 2018.

[48] X. D. Liu, H. S. Yu, J. P. Yu, and Y. Zhao, "A novel speed control method based on port-controlled Hamiltonian and disturbance observer for PMSM drives," *IEEE Access*, vol. 7, pp. 111115-111123, 2019.

[49] W. Wei, W. C. Xue, and D. H. Li, "On disturbance rejection in magnetic levitation," *Control Engineering Practice*, vol. 82, pp. 24-35, January 2019.

[50] Z. Y. Sun, D. Zhang, Q. Meng, and C. C. Chen, "Feedback stabilization of time-delay nonlinear systems with continuous time-varying output function," *International Journal of Systems Science*, vol. 50, no. 2, pp. 244-255, January 2019.

[51] J. X. Ma, W. L. Xiong, J. Chen, and F. Ding, "Hierarchical identification for multivariate Hammerstein systems by using the modified Kalman filter," *IET Control Theory and Applications*, vol. 11, no. 6, pp. 857-869, April 2017.

[52] F. Yang, Y. R. Sun, X. X. Li, and C. Y. Huang, "The quasi-boundary value method for identifying the initial value of heat equation on a columnar symmetric domain," *Numerical Algorithms*, vol. 82, no. 2, pp. 623-639, October 2019.

[53] F. Yang, N. Wang, X. X. Li, and C. Y. Huang, "A quasi-boundary regularization method for identifying the initial value of time-fractional diffusion equation on spherically symmetric domain," *Journal of Inverse and Ill-posed Problems*, 2019. DOI: 10.1515/jiip-2018-0050

[54] F. Yang, Y. Zhang, and X. X. Li, "Landweber iterative method for identifying the initial value problem of the time-space fractional diffusion-wave equation," *Numerical Algorithms*, 2020. DOI: 10.1007/s11075-019-00734-6

[55] X. K. Wan, Y. Li, C. Xia, M. H. Wu, J. Liang, and N. Wang, "A T-wave alternans assessment method based on least squares curve fitting technique," *Measurement*, vol. 86, pp. 93-100, May 2016.

[56] N. Zhao, "Joint optimization of cooperative spectrum sensing and resource allocation in multi-channel cognitive radio sensor networks," *Circuits Systems and Signal Processing*, vol. 35, no. 7, pp. 2563-2583, July 2016.

[57] X. L. Zhao, F. Liu, B. Fu, and F. Na, "Reliability analysis of hybrid multi-carrier energy systems based on entropy-based Markov model," *Proceedings of the Institution of Mechanical Engineers Part O-Journal of Risk and Reliability*, vol. 230, no. 6, pp. 561-569, December 2016.

[58] N. Zhao, M. H. Wu, and J. J. Chen, "Android-based mobile educational platform for speech signal processing," *International Journal of Electrical Engineering Education*, vol. 54, no. 1, pp. 3-16, January 2017.

[59] J. Pan, X. Jiang, X. K. Wan, and W. Ding, "A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems," *International Journal of Control Automation and Systems*, vol. 15, no. 3, pp. 1189-1197, June 2017.

[60] J. Pan, W. Li, and H. P. Zhang, "Control algorithms of magnetic suspension systems based on the improved double exponential reaching law of sliding mode control," *International Journal of Control Automation and Systems*, vol. 16, no. 6, pp. 2878-2887, December 2018.

[61] P. C. Gong, W. Q. Wang, and X. R. Wan, "Adaptive weight matrix design and parameter estimation via sparse modeling for MIMO radar," *Signal Processing*, vol. 139, pp. 1-11, October 2017.

[62] P. C. Gong, W. Q. Wang, F. C. Li, and H. Cheung, "Sparsity-aware transmit beamspace design for FDA-MIMO radar," *Signal Processing*, vol. 144, pp. 99-103, March 2018.

[63] N. Zhao, Y. Liang, and Y. Pei, "Dynamic contract incentive mechanism for cooperative wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10970-10982, November 2018.

[64] X. L. Zhao, Z. Y. Lin, B. Fu, L. He, and F. Na, "Research on automatic generation control with wind power participation based on predictive optimal 2-degree-of-freedom PID strategy for multi-area interconnected power system," *Energies*, vol. 11, no. 12, Article Number: 3325, December 2018.

[65] L. Wang, H. Liu, L. V. Dai, and Y. W. Liu, "Novel method for identifying fault location of mixed lines," *Energies*, vol. 11, no. 6, Article Number: 1520, June 2018.

[66] J. X. Ma and F. Ding, "Filtering-based multistage recursive identification algorithm for an input nonlinear output-error autoregressive system by using key the term separation technique," *Circuits Systems and Signal Processing*, vol. 36, no. 2, pp. 577-599, February 2017.

[67] P. Ma and F. Ding, "New gradient based identification methods for multivariate pseudo-linear systems using the multi-innovation and the data filtering," *Journal of the Franklin Institute*, vol. 354, no. 3, pp. 1568-1583, February 2017.

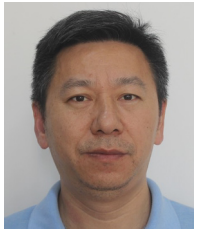
[68] F. Ding, F. F. Wang, L. Xu, and M. H. Wu, "Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering," *Journal of the Franklin Institute*, vol. 354, no. 3, pp. 1321-1339, February 2017.

[69] H. Ma, J. Pan, F. Ding, L. Xu, and W. Ding, "Partially-coupled least squares based iterative parameter estimation for multi-variable output-error-like autoregressive moving average systems," *IET Control Theory and Applications*, vol. 13, 2019. DOI: 10.1049/iet-cta.2019.0112

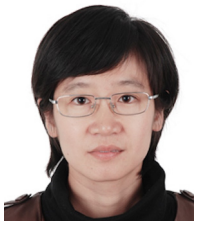
Feng Ding received his B.Sc. degree from the Hubei University of Technology (Wuhan, China) in 1984, and his M.Sc. and Ph.D. degrees both from the Tsinghua University, in 1991 and 1994, respectively. He has been a professor in the School of Internet of Things Engineering at the Jiangnan University (Wuxi, China) since 2004. His current research interests include model identification and adaptive control. He authored four books on System Identification.

Lei Lv was born in Jiaocheng, Shanxi Province, China. He received his B.Sc. degree from the Jiangsu Ocean University (Lianyungang, China) in 2017. He is now a master student at the Hubei University of Technology, Wuhan, China. His research interests include system identification and control theory.

Jian Pan was born in Wuhan, China. He received his B.Sc. degree from the Hubei University of Technology (Wuhan, China) in 1984. From 1984 to 2002, he was with Department of Electrical Engineering and Computer Science, Hubei University of Technology. He has been a Professor in School of Electrical and Electronic Engineering, Hubei University of Technology. His currently research interests include control science and engineering, computer control systems, power electronics. He respectively received Hubei Province Science and Technology Progress Awards, won second prize in 2003 and 2017, third prize in 2005 and 2012. He also received Wuhan City Science and Technology Progress Awards, won third prize in 2012 and second prize in 2014. He is currently a director of Hubei Association of Automation and a director of Wuhan Power Supply Society.



Xiangkui Wan received his M.S. and Ph.D. degrees in the Mechatronic Engineering from Chongqing University, in 2002 and 2005, respectively. He is currently a professor in the School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan, China. His research interests include digital signal processing, biomedical signal processing and analysis, biomedical modeling and simulation, cardiovascular system.



Xue-Bo Jin received her B.E. degree in industrial electrical and automation and a Master degree in industrial automation from Jilin University, Jilin, China, in 1994 and 1997, and her Ph.D. degree in control theory and control engineering from Zhejiang University, Zhejiang, China, in 2004. From 1997 to 2012 she was with College of Informatics and Electronics, Zhejiang Sci-Tech University. Since 2012 she has been with College of Computer and Information Engineering, Beijing Technology and Business University as a Professor. Her research interests include multisensor fusion, statistical signal processing, video/image processing, robust filtering, Bayesian theory and Time series analysis.