International Journal of Control, Automation and Systems 14(1) (2016) 140-147
http://dx.doi.org/10.1007/s12555-014-0429-1
ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Iterative Identification Algorithms for Input Nonlinear Output Error Autoregressive Systems

Junxia Ma, Weili Xiong, and Feng Ding*

Abstract: This paper focuses on the parameter estimation problems of input nonlinear output error autoregressive systems. Based on the key variables separation technique and the auxiliary model identification idea, the output of the system is expressed as a linear combination of all the system parameters, the unknown inner variables in the information vector are replaced with the outputs of the auxiliary model and a gradient based and a least squares based iterative identification algorithms are derived. Simulation example is provided to illustrate the effectiveness of the proposed algorithms.

Keywords: Auxiliary model identification idea, input nonlinear systems, iterative identification, key variables separation technique.

1. INTRODUCTION

Block-oriented nonlinear systems have been successfully applied to modeling several physical processes, such as tractor-trailer [1], chemical processes [2], magneto rheological dampers [3] and so on. Nonlinear system modeling is also widely used in feedback control and predictive control [4–6]. The block-oriented nonlinear systems, which consist of the interaction of linear dynamic subsystems and static nonlinear elements [7], contain Hammerstein systems, Wiener systems and their combinations [8–10]. For decades, much research has been performed on the nonlinear systems, and several approaches for the parameter estimation and state estimation have been presented [11–13], such as the subspace state space approaches [14], the filtering technique [15] and the maximum likelihood methods [16]. Recently, Li *et al.* employed a kernel machine to approximate the static nonlinear function and proposed a space projection method to identify a class of nonlinear autoregressive models with exogenous inputs [17]; Paduart *et al.* applied the polynomial nonlinear state space approach to identify a nonlinear system with a Wiener Hammerstein structure [18]; Karimi and McAuley developed a Laplace approximation maximum likelihood estimation algorithm for estimating measurement noise variances and model parameters in nonlinear stochastic differential equation models [19].

Iterative methods can be used for the iterative learning control and optimal design. In the area, the iterative approaches have been widely used for finding the solutions of matrix equations [20], obtaining the parameter estimates of linear and nonlinear systems [21] and finding the optimal analysis and synthesis filters [22]. Recently, Hajarian proposed an iterative algorithm to solve the periodic Sylvester matrix equations [23]; Li and Wen proposed a normalized iterative algorithm for Hammerstein systems and proved the normalized algorithm can ensure the convergence property under arbitrary nonzero initial conditions [24]. Iterative algorithms can be combined with the gradient search, the least squares search and the Newton search to form new identification algorithms. For example, Xie and Yang proposed a least squares based iterative identification for output error moving average systems [25]; Wang and Tang presented a gradient-based iterative estimation algorithm for a class of nonlinear systems with colored noise using the decomposition technique [26].

The auxiliary model identification method is an effective approach for identifying systems containing the unknown variables in the information vector and its idea is to set up an auxiliary model by using the measurable information. For example, Ding *et al.* studied an auxiliary model based recursive extended least squares algorithm for dual-rate output error systems with colored noise based on the dual-rate noisy data [27]. Hammerstein nonlinear systems contain product terms of the parameter of the linear and nonlinear blocks. The key variables separation technique can be used to solve this problem. Vörös proposed a recursive algorithm for Hammerstein systems with discontinuous nonlinearities containing dead-zones

Manuscript received October 17, 2014; revised March 17, 2015; accepted March 30, 2015. Recommended by Associate Editor Choon Ki Ahn under the direction of Editor PooGyeon Park. This work was supported in part by the National Natural Science Foundation of China (No. 21206053) and Industry-Academia-Research Project of Jiangsu Province (BY2014023-27).

Junxia Ma, Weili Xiong, and Feng Ding are with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, P. R. China (e-mails: junxia.20@163.com, greenpre@163.com, fding@jiangnan.edu.cn).
* Corresponding author.

using the key-term separation technique [28].

On the basis of the work in [29], this paper expends the identification model from an input nonlinear finite impulse response moving average model to a Hammerstein nonlinear system with an output error autoregressive linear subsystem and studies its parameter estimation problem. Different from the work in [29], to obtain the linear regressive form of the identification model, we need to define two intermediate variables and then use the output of the auxiliary model to replace the unknown intermediate variable. The main contributions of this paper are to describe the output of the system in a linear combination of all the system parameters and to present a gradient based and a least squares based iterative identification algorithms by using the key variables separation and the auxiliary model. As a comparison, the auxiliary model based stochastic gradient and recursive generalized least squares algorithms are given. The proposed algorithms are different from the over-parameterization iterative least squares algorithm in [30] and the least squares based and the gradient based iterative algorithms using the hierarchical identification principle in [31].

The rest of this paper is organized as follows. Section 2 gives the identification model of input nonlinear systems. Section 3 presents a gradient based iterative identification algorithm and an auxiliary model based stochastic gradient algorithm. To improve the convergence rate of the iterative algorithm, a least squares based iterative identification algorithm is given in Section 4. The numerical example is provided in Section 5 to show the effectiveness of the proposed algorithms. Finally, Section 6 offers some concluding remarks.

2. SYSTEM DESCRIPTION

Let us introduce some notation.

Symbols	Meaning
$\boldsymbol{I}$	The identity matrix of appropriate sizes.
$\hat{\theta}(t)$	The estimate of θ at time t.
$\hat{\theta}_k(t)$	The estimate of θ at iteration k.
$\hat{a}_k(t)$	The estimate of a at iteration k.
$\hat{b}_k(t)$	The estimate of b at iteration k.
$\hat{c}_k(t)$	The estimate of c at iteration k.
$\hat{\gamma}_k(t)$	The estimate of γ at iteration k.
$\lambda_{\max}[X]$	The maximum eigenvalue of symmetric square matrix $\boldsymbol{X}$.
X^{T}	The transpose of the vector or matrix $\boldsymbol{X}$.
$\Vert X\Vert$	The norm of the vector $\boldsymbol{X}$.
X:=A	$\boldsymbol{X}$ is defined by $\boldsymbol{A}$.

The typical nonlinear system includes the input nonlinear system and the output nonlinear system. Here, we consider an input nonlinear output error autoregressive (IN-OEAR) system as shown in Fig. 1,

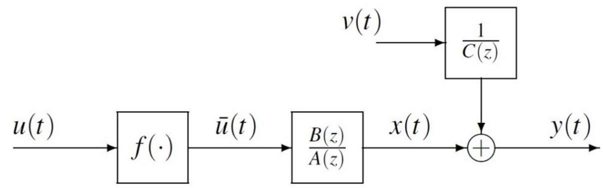


Fig. 1. The IN-OEAR system.

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

$$\bar{u}(t)=f(u(t)), \tag{2}$$

where $y(t)$ is the measured output, $v(t)$ is the white noise with zero mean and variances σ^2, $u(t)$ and $\bar{u}(t)$ are the input and output of the nonlinear block, respectively, and A(z), B(z) and C(z) are polynomials in the unit backward shift operator $z^{-1}(z^{-1}y(t)=y(t-1)$:

$$A(z):=1+a_1z^{-1}+a_2z^{-2}+\cdots+an_az^{-n_a},$$

$$B(z):=b_0+b_1z^{-1}+b_2z^{-2}+\cdots+bn_bz^{-n_b},$$

$$C(z):=1+c_1z^{-1}+c_2z^{-2}+\cdots+cn_cz^{-n_c}.$$

Assume that the orders n_a, n_b and n_c are known and $y(t)=0$, $u(t)=0$ and $v(t)=0$ for $t\leq 0$. In order to ensure the effectiveness of the identification algorithms, the input signal should motivate all characteristics of the system. The following assumptions are required.

Assumption 1: The input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance $\sigma_u^2=1.00^2$.

The nonlinear block is polynomials or trigonometric functions, more generally, $\bar{u}(t)$satisfies the following assumption:

Assumption 2: The output of the nonlinear block is a nonlinear function of the known basis:

$$\bar{u}(t):=\gamma_1f_1(u(t))+\gamma_2f_2(u(t))+\cdots+\gamma_mf_m(u(t)), \tag{3}$$

where $\alpha_i^{'}$s are the coefficients and $f_i(u(t))'$s are the base functions. Let

$$\gamma:=[\gamma_1,\gamma_2,\cdots\gamma_m]^{\mathrm{T}}\in R^{\mathrm{m}},$$

$$f(u(t)):=[f_1(u(t)),f_2(u(t))\cdots f_m(u(t))]\in R^{1\times m}.$$

From (3), we have

$$\bar{u}(t)=f(u(t))\gamma.$$

Define two intermediate variables:

$$x(t):=\frac{B(z)}{A(z)}\bar{u}(t),w(t):=\frac{1}{C(z)}v(t).$$

Then, we have

$$\begin{aligned}x(t)&=[1-A(z)]x(t)+B(z)\bar{u}(t)\\&=-\sum_{i=1}^{n_a}a_ix(t-i)+b_0\bar{u}(t)+\sum_{i=1}^{n_b}b_i\bar{u}(t-i),\end{aligned} \tag{4}$$

$$w(t) = [1 - C(z)]w(t) + v(t) = -\sum_{i=1}^{n_c} c_i w(t-i) + v(t).$$

The output $y(t)$ in (1) can be expressed as

$$y(t) = x(t) + w(t) \tag{5}$$

$$= -\sum_{i=1}^{n_a} a_i x(t-i) + b_0\bar{u}(t) + \sum_{i=1}^{n_b} b_i \bar{u}(t-i) - \sum_{i=1}^{n_c} c_i w(t-i) + v(t). \tag{6}$$

Define the parameter vectors and the information vectors:

$$\mathbf{a} := \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n_a} \end{bmatrix} \in R^{n_a}, \bar{\mathbf{b}} := \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \in R^{n_b+1},$$

$$\mathbf{c} := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n_c} \end{bmatrix} \in R^{n_c}.$$

$$\phi_{\mathrm{s}}(t) := [-x(t-1), -x(t-2), \cdots - x(t-n_a)]^{\mathrm{T}} \in R^{n_a},$$

$$\mathbf{F}(t) := [f(u(t)), f(u(t-1)), \cdots f(u(t-n_b))]^{\mathrm{T}} \in R^{(n_b+1)\times m},$$

$$\phi_{\mathrm{n}}(t) := [-w(t-1), -w(t-2), \cdots - w(t-n_c)]^{\mathrm{T}} \in R^{n_c},$$

Then, equation (6) can be written as

$$y(t) = \phi_{\mathrm{s}}^{\mathrm{T}}(t)\mathbf{a} + \bar{\mathbf{b}}^{\mathrm{T}}\mathbf{F}(t)\gamma + \phi_{\mathrm{n}}(t)\mathbf{c} + v(t). \tag{7}$$

However, difficulties arise in that the model in (7) contains the product terms of the parameter vectors $\bar{\mathbf{b}}$ and γ of the linear and nonlinear blocks. Although we can use the over-parameterization model in (7) for identification [30], the dimension of the resulting unknown parameter vector increases, so does the calculation load. In this paper, we adopt the key variables separation [32] and choose the second term $\bar{u}(t)$on the right-hand side of (6) as a separated key variable, the rests as the non-separated key variables, and let the coefficient $b_0 = 1$ [33]. Then, equation (6) can be rewritten as

$$y(t) = \phi_{\mathrm{s}}^{\mathrm{T}}(t)\mathbf{a} + \bar{u}(t) + b_1\bar{u}(t-1) + b_2\bar{u}(t-2) + \cdots + b_{n_b}\bar{u}(t-n_b) + \phi_{\mathrm{n}}(t)\mathbf{c} + v(t). \tag{8}$$

Let $n := n_a + m + n_b + n_c$, define the parameter vector θ and the information vector $\phi(t)$ as

$$\theta := \begin{bmatrix} \vartheta \\ \mathbf{c} \end{bmatrix} \in R^n, \vartheta := \begin{bmatrix} \mathbf{a} \\ \gamma \\ \mathbf{b} \end{bmatrix} \in R^{n_a+m+n_b},$$

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n_b} \end{bmatrix} \in R^{n_b},$$

$$\phi(t) := [\psi^{\mathrm{T}}(t), -w(t-1), \cdots, -w(t-n_c)]^{\mathrm{T}} \in R^n,$$

$$\psi(t) := [-x(t-1), -x(t-2), \cdots, -x(t-n_a), \mathbf{f}(u(t)), \bar{u}(t-1), \cdots, \bar{u}(t-n_b)]^{\mathrm{T}} \in R^{n_a+m+n_b}. \tag{9}$$

Substituting the separated key variable $\bar{u}(t)$ in (3) into (4) gives

$$x(t) = -a_1x(t-1) - a_2x(t-2) - \cdots - a_{n_a}x(t-n_a) + \alpha_1 f_1(u(t)) + \alpha_2 f_2(u(t)) + \cdots + \alpha_m f_m(u(t)) + b_1\bar{u}(t-1) + b_2\bar{u}(t-2) + \cdots + b_{n_b}\bar{u}(t-n_b) = \psi^{\mathrm{T}}(t)\vartheta. \tag{10}$$

Substituting (10) into (5) gives

$$y(t) = \psi^{\mathrm{T}}(t)\vartheta + w(t) = \psi^{\mathrm{T}}(t)\vartheta + \phi_{\mathrm{n}}^{\mathrm{T}}(t)\mathbf{c} + v(t) = \phi^{\mathrm{T}}(t)\theta + v(t). \tag{11}$$

Based on the key variables separation, we obtain the identification model in (11) for this IN-OEAR system in the linear regressive form.

3. THE GRADIENT BASED ITERATIVE ALGORITHM

Opt a set of data from $j = t - L + 1$ to $j = t$ (L denotes the data length) and define a quadratic cost function

$$J_1(\theta) := ||\mathbf{Y}(t) - \Phi(t)\theta||^2,$$

where

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

$$\Phi(t) := \begin{bmatrix} \phi^{\mathrm{T}}(t) \\ \phi^{\mathrm{T}}(t-1) \\ \vdots \\ \phi^{\mathrm{T}}(t-L+1) \end{bmatrix} \in R^{L\times n}.$$

Let $k = 1, 2, 3, \cdots$ be an iterative variable and $\hat{\theta}_k(t)$ be the estimate of θ at iteration k, define

$$\hat{\theta}_k(t) := \begin{bmatrix} \hat{\vartheta}_k(t) \\ \hat{\mathbf{c}}_k(t) \end{bmatrix} \in R^n,$$

$$\hat{\vartheta}_k(t) := \begin{bmatrix} \hat{\mathbf{a}}_k(t) \\ \hat{\alpha}_k(t) \\ \hat{\mathbf{b}}_k(t) \end{bmatrix} \in R^{n_a+m+n_b},$$

$$\hat{\mathbf{a}}_k(t) := [\hat{a}_{1,k}(t), \hat{a}_{2,k}(t), \cdots, \hat{a}_{n_a,k}(t)]^{\mathrm{T}} \in R^{n_a},$$

$$\begin{aligned}\hat{\alpha}_k(t) &:= [\hat{\alpha}_{1,k}(t), \hat{\alpha}_{2,k}(t), \cdots, \hat{\alpha}_{m,k}(t)]^{\mathrm{T}} \in R^m,\\ \hat{\mathbf{b}}_k(t) &:= [\hat{b}_{1,k}(t), \hat{b}_{2,k}(t), \cdots, \hat{b}_{n_b,k}(t)]^{\mathrm{T}} \in R^{n_b},\\ \hat{\mathbf{c}}_k(t) &:= [\hat{c}_{1,k}(t), \hat{c}_{2,k}(t), \cdots, \hat{c}_{n_c,k}(t)]^{\mathrm{T}} \in R^{n_c}.\end{aligned}$$

Minimizing $J_1(\theta)$ by using the negative gradient search, we have

$$\begin{aligned}\hat{\theta}_k(t) &= \hat{\theta}_{k-1}(t) - \frac{\mu_k(t)}{2}\mathrm{grad}[J_1(\hat{\theta}_{k-1}(t))]\\ &= \hat{\theta}_{k-1}(t) + \mu_k(t)\Phi^{\mathrm{T}}(t)[\mathbf{Y}(t) - \Phi(t)\hat{\theta}_{k-1}(t)].\end{aligned} \tag{12}$$

Since the information vector $\phi(t)$ contains the unknown inner variables $x(t-i)$,$\bar{u}(t-i)$and $w(t-i)$, we cannot compute the estimate $\hat{\theta}_k(t)$ from (12). The approach here is to establish an auxiliary model by using the auxiliary model identification idea [34], the unknown inner variables $x(t-i)$ are replaced with the outputs of the auxiliary model.

Define an auxiliary model

$$x_a(t) := \frac{B_a(z)}{A_a(z)}\bar{u}(t),$$

where $A_a(z)$ and $B_a(z)$ are the polynomials which have the same orders with the $A(z)$ and $B(z)$. As shown in Fig. 2, the variable $x_a(t)$ is the output of the auxiliary model.

Referring to the method in [35], we take the estimate $\frac{\hat{B}(z)}{\hat{A}(z)}$ as the transfer function of the auxiliary model. Define the estimate $\hat{\psi}_k(t)$ of $\psi(t)$:

$$\begin{aligned}\hat{\psi}_k(t) =&[-\hat{x}_{a,k-1}(t-1), \cdots, -\hat{x}_{a,k-1}(t-n_a),\\ &\mathbf{f}(u(t)), \hat{\bar{u}}_{k-1}(t-1), \cdots, \hat{\bar{u}}_{k-1}(t-n_b)]^{\mathrm{T}},\\ \hat{\bar{u}}_k(t) =&\mathbf{f}(u(t))\hat{\gamma}_k(t).\end{aligned}$$

Following (10), the unknown variables $x(t-i)$ in $\psi(t)$ in (9) are replaced with the outputs of the auxiliary model:

$$\hat{x}_{a,k}(t) = \hat{\psi}_k^{\mathrm{T}}(t)\hat{\vartheta}_k(t).$$

From (5), we have

$$w(t) = y(t) - x(t). \tag{13}$$

Substituting $x(t)$ in (13) with $\hat{x}_{a,k}(t)$, the iterative estimate $\hat{w}_k(t)$ of $w(t)$ can be computed by

$$\hat{w}_k(t) = y(t) - \hat{x}_{a,k}(t).$$

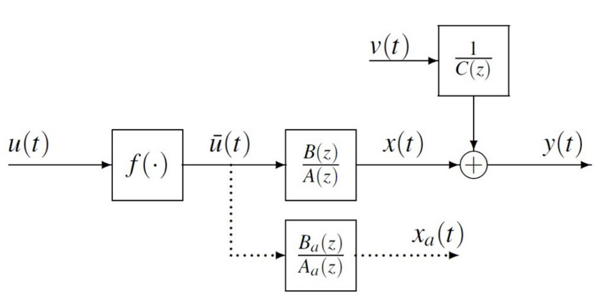


Fig. 2. The IN-OEAR system with an auxiliary model.

Replace the unknown $w(t-i)$ in $\phi(t)$ with its estimate $\hat{w}_{k-1}(t-i)$ at iteration $k-1$, and define

$$\hat{\phi}_k(t) = [\hat{\psi}_k^{\mathrm{T}}(t), -\hat{w}_{k-1}(t-1), \cdots, -\hat{w}_{k-1}(t-n_c)]^{\mathrm{T}},$$

$$\hat{\Phi}_k(t) := \begin{bmatrix} \hat{\phi}_k^{\mathrm{T}}(t) \\ \hat{\phi}_k^{\mathrm{T}}(t-1) \\ \vdots \\ \hat{\phi}_k^{\mathrm{T}}(t-L+1) \end{bmatrix} \in R^{L\times n}.$$

Replacing $\Phi(t)$in (12) with its estimate $\hat{\Phi}_k(t)$, we can obtain the auxiliary model based gradient iterative (AM-GI) algorithm for estimating the parameter vector θ of the IN-OEAR system:

$$\begin{aligned}\hat{\theta}_k(t) &= \hat{\theta}_{k-1}(t) + \mu_k(t)\hat{\Phi}_k^{\mathrm{T}}(t)\\ &\quad\times[Y(t) - \hat{\Phi}_k^{\mathrm{T}}(t)\hat{\theta}_{k-1}(t)],\end{aligned} \tag{14}$$

$$\mathbf{Y}(t) = [y(t), y(t-1), \cdots, y(t-L+1)]^{\mathrm{T}}, \tag{15}$$

$$\hat{\Phi}_k(t) = [\hat{\phi}_k(t), \hat{\phi}_k(t-1), \cdots, \hat{\phi}_k(t-L+1)]^{\mathrm{T}}, \tag{16}$$

$$\begin{aligned}\hat{\phi}_k(t) = [&\hat{\psi}_k^{\mathrm{T}}(t), -\hat{w}_{k-1}(t-1), \cdots,\\ &-\hat{w}_{k-1}(t-n_c)]^{\mathrm{T}},\end{aligned} \tag{17}$$

$$\begin{aligned}\hat{\psi}_k(t) =&[-\hat{x}_{a,k-1}(t-1), -\hat{x}_{a,k-1}(t-2), \cdots,\\ &-\hat{x}_{a,k-1}(t-n_a),\\ &\mathbf{f}(u(t)), \hat{\bar{u}}_{k-1}(t-1), \hat{\bar{u}}_{k-1}(t-2), \cdots,\\ &\hat{\bar{u}}_{k-1}(t-n_b)]^{\mathrm{T}},\end{aligned} \tag{18}$$

$$\hat{x}_{a,k}(t) = \hat{\psi}_k^{\mathrm{T}}(t)\hat{\vartheta}_k(t), \tag{19}$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t))\cdots f_m(u(t))], \tag{20}$$

$$\hat{\bar{u}}_k(t) = \mathbf{f}(u(t))\hat{\gamma}_k(t), \tag{21}$$

$$\hat{w}_k(t) = y(t) - \hat{x}_{a,k}(t), \tag{22}$$

$$0 < \mu_k(t) \le \frac{2}{\lambda_{\max}[\hat{\Phi}_k^{\mathrm{T}}(k)\hat{\Phi}_k(k)]}, \tag{23}$$

$$\hat{\theta}_k(t) = [\hat{\vartheta}_k^{\mathrm{T}}(t), \hat{\mathbf{c}}_k^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{24}$$

$$\hat{\vartheta}_k(t) = [\hat{\mathbf{a}}_k^{\mathrm{T}}(t), \hat{\gamma}_k^{\mathrm{T}}(t), \hat{\mathbf{b}}_k^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{25}$$

At each iteration, the parameter estimate $\hat{\theta}_k(t)$ is based on the estimates of the inner variables $\hat{x}_{k-1}(t-i)$, $\hat{\bar{u}}_{k-1}(t-i)$ and $\hat{w}_{k-1}(t-i)$, see (14) and (16)–(18). In turn, the inner variables are computed by the previous iterative estimate $\hat{\theta}_k(t)$, see (19), (21) and (22).

To show the advantages of the AM-GI algorithm, the following gives the stochastic gradient algorithm based on the auxiliary model identification idea. Consider the input-output data set $\{y(j), u(j), 0 \le j \le t\}$, and define a cost function

$$J_2(\theta) := \sum_{j=1}^{t}[y(j) - \phi^{\mathrm{T}}(j)\theta]^2.$$

Minimizing $J_2(\theta)$ based on the negative gradient search, we can obtain the auxiliary model based stochastic gradient (AM-SG) algorithm for estimating θ of the IN-OEAR system:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\hat{\phi}(t)}{r(t)} e(t), \tag{26}$$

$$e(t) = y(t) - \hat{\phi}^{\mathrm{T}}(t)\hat{\theta}(t-1), \tag{27}$$

$$r(t) = r(t-1) + ||\hat{\phi}(t)||^2, r(0) = 1, \tag{28}$$

$$\hat{\phi}(t) = [\hat{\psi}^{\mathrm{T}}(t), -\hat{w}(t-1), \cdots, -\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{29}$$

$$\begin{aligned}\hat{\psi}(t) = [&-\hat{x}_a(t-1), \cdots, -\hat{x}_a(t-n_a), \mathbf{f}(u(t)),\\ &\hat{\bar{u}}(t-1), \hat{\bar{u}}(t-2), \cdots, \hat{\bar{u}}(t-n_b)]^{\mathrm{T}},\end{aligned} \tag{30}$$

$$x_a(t) = \hat{\psi}^{\mathrm{T}}(t)\hat{\vartheta}(t), \tag{31}$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)) \cdots f_m(u(t))], \tag{32}$$

$$\hat{\bar{u}}(t) = \mathbf{f}(u(t))\hat{\gamma}(t), \tag{33}$$

$$\hat{w}(t) = y(t) - x_a(t), \tag{34}$$

$$\hat{\theta}(t) = [\hat{\vartheta}^{\mathrm{T}}(t), \hat{\mathbf{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{35}$$

$$\hat{\vartheta}(t) = [\hat{\mathbf{a}}^{\mathrm{T}}(t), \hat{\gamma}^{\mathrm{T}}(t), \hat{\mathbf{b}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{36}$$

4. THE LEAST SQUARES BASED ITERATIVE ALGORITHM

The convergence rate of the AM-GI algorithm is slow. To improve the convergence speed, this section derives a least squares based iterative algorithm based on the auxiliary model identification idea. Minimizing $J_1(\theta)$ and letting the derivative of $J_1(\theta)$ with respect to θ be zero gives

$$\left.\frac{\partial J_1(\theta)}{\partial \theta}\right|_{\theta=\hat{\theta}(t)} = -2\Phi^{\mathrm{T}}(t)[\mathbf{Y}(t) - \Phi(t)\hat{\theta}(t)] = 0.$$

Then we can obtain the least squares estimate of the parameter vector θ:

$$\hat{\theta}_k(t) = [\Phi^{\mathrm{T}}(t)\Phi(t)]^{-1}\Phi^{\mathrm{T}}(t)\mathbf{Y}(t). \tag{37}$$

Because the inner variables $x(t-i)$, $\hat{\bar{u}}(t-i)$ and $w(t-i)$ in $\phi(t)$ are unknown, it is impossible to obtain the least squares estimate $\hat{\theta}_k(t)$ from (37). Similarly, using the auxiliary model identification idea and the unknown inner variables are replaced with the outputs of the auxiliary models, i.e., replacing $\Phi(t)$ in (37) with $\hat{\Phi}_k(t)$, we can obtain the auxiliary model based iterative least squares (AM-LSI) algorithm for estimating the parameter vector θ:

$$\hat{\theta}_k(t) = [\hat{\Phi}_k^{\mathrm{T}}(t)\hat{\Phi}_k(t)]^{-1}\Phi_k^{\mathrm{T}}(t)\mathbf{Y}(t), \tag{38}$$

$$\mathbf{Y}(t) = [y(t), y(t-1), \cdots, y(t-L+1)]^{\mathrm{T}}, \tag{39}$$

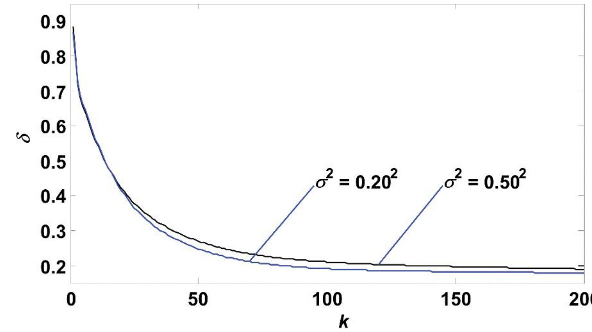


Fig. 3. The AM-GI estimation errors δ versus k.

$$\hat{\Phi}_k(t) = [\hat{\phi}_k(t), \hat{\phi}_k(t-1), \cdots, \hat{\phi}_k(t-L+1)]^{\mathrm{T}}, \tag{40}$$

$$\hat{\varphi}_k(t) = [\hat{\psi}_k^{\mathrm{T}}(t), -\hat{w}_{k-1}(t-1), \cdots, -\hat{w}_{k-1}(t-n_c)]^{\mathrm{T}}, \tag{41}$$

$$\begin{aligned}\hat{\psi}_k(t) = [&-\hat{x}_{a,k-1}(t-1), \cdots, -\hat{x}_{a,k-1}(t-n_a),\\ &\mathbf{f}(u(t)), \hat{\bar{u}}_{k-1}(t-1), \cdots, \hat{\bar{u}}_{k-1}(t-n_b)]^{\mathrm{T}},\end{aligned} \tag{42}$$

$$\hat{x}_{a,k}(t) = [\hat{\mathbf{a}}_k^{\mathrm{T}}(t), \hat{\gamma}_k^{\mathrm{T}}(t), \hat{\mathbf{b}}_k^{\mathrm{T}}(t)]\hat{\psi}_k(t), \tag{43}$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)) \cdots f_m(u(t))], \tag{44}$$

$$\hat{\bar{u}}_k(t) = \mathbf{f}(u(t))\hat{\gamma}_k(t), \tag{45}$$

$$\hat{w}_k(t) = y(t) - \hat{x}_{a,k}(t), \tag{46}$$

$$\hat{\theta}_k(t) = [\hat{\mathbf{a}}_k^{\mathrm{T}}(t), \hat{\gamma}_k^{\mathrm{T}}(t), \hat{\mathbf{b}}_k^{\mathrm{T}}(t), \hat{\mathbf{c}}_k^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{47}$$

To initialize the AM-LSI algorithm, the initial value $\hat{\theta}(0)$ is generally taken to be a small real vector, e.g., $\hat{\theta}(0) = \mathbf{1}_n/p_0$ (p_0 being normally a large positive number , e.g., $p_0 = 10^6$).

For comparisons, we simply give the recursive generalized least squares algorithm of estimating the parameter vector θ. Minimizing $J_2(\theta)$ and letting the derivative of $J_2(\theta)$ with respect to θ be zero leads to the following recursive generalized least squares algorithm of estimating θ based on the key-term separation and the auxiliary model (the AM-RGLS algorithm for short):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\phi}^{\mathrm{T}}(t)\hat{\theta}(t-1)], \tag{48}$$

$$\mathbf{L}(t) = \mathbf{P}(t-1)\hat{\phi}(t)[1 + \hat{\phi}^{\mathrm{T}}(t)\mathbf{P}(t-1)\hat{\phi}(t)]^{-1}, \tag{49}$$

$$\mathbf{P}(t) = [\mathbf{I} - \mathbf{L}(t)\hat{\phi}^{\mathrm{T}}(t)]\mathbf{P}(t-1), \tag{50}$$

$$\hat{\phi}(t) = [\hat{\psi}^{\mathrm{T}}(t), -\hat{w}(t-1), \cdots, -\hat{w}(t-n_c)]^{\mathrm{T}}, \tag{51}$$

$$\begin{aligned}\hat{\psi}(t) = [&-x_a(t-1), \cdots, -x_a(t-n_a), \mathbf{f}(u(t)),\\ &\hat{\bar{u}}(t-1), \cdots, \hat{\bar{u}}(t-n_b)]^{\mathrm{T}},\end{aligned} \tag{52}$$

$$x_a(t) = [\hat{\mathbf{a}}^{\mathrm{T}}(t), \hat{\gamma}^{\mathrm{T}}(t), \hat{\mathbf{b}}^{\mathrm{T}}(t)]\hat{\psi}(t), \tag{53}$$

$$\mathbf{f}(u(t)) = [f_1(u(t)), f_2(u(t)) \cdots f_m(u(t))], \tag{54}$$

$$\hat{\bar{u}}(t) = \mathbf{f}(u(t))\hat{\gamma}(t), \tag{55}$$

$$\hat{w}(t) = y(t) - x_a(t), \tag{56}$$

$$\hat{\theta}(t) = [\hat{\mathbf{a}}^{\mathrm{T}}(t), \hat{\gamma}^{\mathrm{T}}(t), \hat{\mathbf{b}}^{\mathrm{T}}(t), \hat{\mathbf{c}}^{\mathrm{T}}(t)]^{\mathrm{T}}. \tag{57}$$

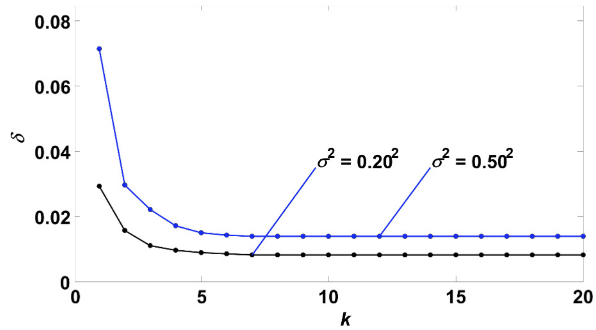


Fig. 4. The AM-LSI estimation errors δ versus k.

Compared with the AM-RGLS algorithm, AM-LSI algorithm fully uses all measured input and output data at each iteration.

5. EXAMPLE

Consider the following IN-OEAR simulation system:

$$\begin{aligned} y(t) &= \frac{B(z)}{A(z)}\bar{u}(t) + \frac{1}{C(z)}v(t), \\ \bar{u}(t) &= \alpha_1 u(t) + \alpha_2 u^2(t) = 0.80u(t) - 0.60u^2(t), \\ A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} = 1 - 1.00z^{-1} + 0.32z^{-2}, \\ B(z) &= 1 + b_1 z^{-1} + b_2 z^{-2} = 1 + 1.25z^{-1} - 0.48z^{-2}, \\ C(z) &= 1 + c_1 z^{-1} = 1 - 0.95z^{-1}, \\ \theta &= [-1.00, 0.32, 0.80, -0.60, 1.25, -0.48, -0.95]^{\mathrm{T}}. \end{aligned}$$

In simulation, the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance, and $\{v(t)\}$ as a white noise sequence with zero mean and variance σ^2. Taking the data length $L = 1000$, and applying the AM-GI algorithm in (14)-(25) and AM-LSI algorithm in (38)-(47) to estimate the parameters of this system, the parameter estimates and their estimation errors $\delta := ||\hat{\theta}_k(t) - \theta||/||\theta||$ with different noise variances are shown in Tables 1-2 and Figs. 3-4.

From Tables 1-2 and Figs. 3-4, we can draw the following conclusions:

- The estimation errors are becoming smaller (in general) as iteration k increases. Thus the proposed algorithms are effective;
- The AM-LSI algorithm has faster convergence speed and can generate more accurate parameter estimates than the AM-GI algorithm;
- The estimation errors given by the AM-GI and the AM-LSI algorithms become small as the noise variance decreases.

6. CONCLUSION

The iterative identification algorithms are presented for IN-OEAR systems by using the key variables and the auxiliary model. They are the gradient based and the least squares based iterative algorithms. Compared with the gradient based iterative algorithm, the least squares based iterative algorithm has faster convergence speed. The proposed iterative identification algorithms can be extended to Wiener nonlinear systems [36, 37] or other multivariable systems or multirate sampled-data systems [38] and applied to other areas [39–41].

REFERENCES

[1] E. Kayacan, H. Ramon, and W. Saeys, “Nonlinear modeling and identification of an autonomous tractor-trailer system,” *Computers and Electronics in Agriculture*, vol. 106, pp. 1-10, August 2014. [click]

[2] M. Matinfar, M. Saeidy, B. Gharahsuflu, and M. Eslami, “Solutions of Nonlinear Chemistry Problems by Homotopy Analysis,” *Computational Mathematics and Modeling*, vol. 25, no. 1, pp. 103-114, January 2014. [click]

[3] J. Wang, A. Sano, T. Chen, and B. Huang, “Identification of Hammerstein systems without explict parameterization of nonlinearity,” *International Journal of Control*, vol. 82, no. 5, pp. 937-952, May 2009. [click]

[4] S. Saat, S. K. Nguang, A. M. Darsono, and N. Azman, “Nonlinear H_∞ feedback control with integrator for polynomial discrete-time systems,” *Journal of the Franklin Institute*, vol. 351, no. 8, pp. 4023-4038, August 2014. [click]

[5] L. C. Guo, X. Zuo, J. W. Liu, and H. Q. Liang, “Output-feedback control of a class of stochastic nonlinear systems with power growth conditions,” *International Journal of Control, Automation, and Systems*, vol. 12, no. 2, pp. 274-282, April 2014. [click]

[6] M. H. Murillo, A. C. Limache, P. S. Rojas Fredini, and L. L. Giovanini “Generalized nonlinear optimal predictive control using iterative state-space trajectories: applications to autonomous flight of UAVs,” *International Journal of Control, Automation, and Systems*, vol. 13, no. 2, pp. 361-370, April 2015. hrefhttp://dx.doi.org/10.1007/s12555-013-0416-y[click]

[7] F. Giri and E. W. Bai, “Block-oriented nonlinear system identification,” *Lecture Notes in Control and Information Sciences*, vol. 404, 2010. [click]

[8] B. Zhang and Z. Z. Mao, “Consistent parameter estimation and convergence properties analysis of Hammerstein output-error models,” *International Journal of Control, Automation, and Systems*, vol. 13, no. 2, pp.1-9, February 2015. [click]

[9] F. Ikhouane and F. Giri, “A unified approach for the parametric identification of SISO/MIMO Wiener and Hammerstein systems,” *Journal of the Franklin Institute*, vol. 351, no. 3, pp. 1717-1727, March 2014. [click]

[10] M. Schoukens, R. Pintelon, and Y. Rolain, “Identification of Wiener-Hammerstein systems by a nonparametric separation of the best linear approximation,” *Automatica*, vol. 50, no. 2, pp. 628-634, February 2014. [click]

[11] F. Ding, *System Identification—New Theory and Methods*, Science Press, Beijing, 2013.

Table 1. The AM-GI estimates and errors versus iteration k.

σ^2	k	a_1	a_2	α_1	α_2	b_1	b_2	c_1	$\delta(\%)$
0.20^2	2	-0.16029	-0.12517	0.22927	-0.97336	0.15264	0.12803	-0.63057	79.65744
	5	-0.42777	-0.18347	0.30822	-0.95132	0.40310	0.16483	-0.76814	66.35291
	50	-0.83121	0.26399	0.69017	-0.72365	1.22699	0.00042	-0.88508	24.77675
	100	-0.83622	0.23237	0.76903	-0.63921	1.39168	-0.13705	-0.92751	19.10917
	200	-0.83758	0.21693	0.79580	-0.60556	1.44030	-0.19693	-0.93538	17.89322
0.50^2	2	-0.11526	-0.06408	0.23599	-0.97175	0.14691	0.12218	-0.55161	80.70227
	5	-0.42591	-0.15435	0.31756	-0.94824	0.43645	0.19063	-0.72421	65.70359
	50	-0.81966	0.26439	0.70509	-0.70911	1.22860	0.02561	-0.75625	27.01557
	100	-0.82078	0.22874	0.77832	-0.62787	1.39003	-0.11474	-0.81677	21.04387
	200	-0.82318	0.21573	0.79901	-0.60132	1.43629	-0.17530	-0.88904	19.01118
True values		-1.00000	0.32000	0.80000	-0.60000	1.25000	-0.48000	-0.95000	

Table 2. The AM-LSI estimates and errors versus iteration k.

σ^2	k	a_1	a_2	α_1	α_2	b_1	b_2	c_1	$\delta(\%)$
0.20^2	1	-1.03054	0.33682	0.79991	-0.60013	1.23060	-0.52797	-0.93581	2.92117
	2	-1.01544	0.32633	0.80068	-0.59909	1.24515	-0.50593	-0.93549	1.57238
	5	-1.00583	0.32018	0.79947	-0.60071	1.25573	-0.48943	-0.93586	0.86190
	10	-1.00362	0.31871	0.79924	-0.60101	1.25820	-0.48575	-0.93583	0.81331
	20	-1.00344	0.31860	0.79922	-0.60104	1.25840	-0.48546	-0.93583	0.81235
0.50^2	1	-1.07635	0.36207	0.79978	-0.60029	1.20152	-0.59994	-0.93651	7.14843
	2	-1.03160	0.33190	0.80179	-0.59760	1.24475	-0.53352	-0.93612	2.97063
	5	-1.01357	0.31998	0.79866	-0.60179	1.26536	-0.50197	-0.93658	1.50566
	10	-1.00866	0.31667	0.79814	-0.60248	1.27086	-0.49376	-0.93656	1.37023
	20	-1.00828	0.31641	0.79809	-0.60253	1.27129	-0.49313	-0.93656	1.36773
True values		-1.00000	0.32000	0.80000	-0.60000	1.25000	-0.48000	-0.95000	

[12] F. Ding, *System Identification—Performances Analysis for Identification Methods*, Science Press, Beijing, 2014.

[13] H. Z. Fang, Y. Shi, and J. G. Yi, "On stable simultaneous input and state estimation for discrete-time linear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 8, pp. 671-686, August 2011. [click]

[14] A. Savran, "Discrete state space modeling and control of nonlinear unknown systems," *ISA Transactions*, vol. 52, no. 6, pp. 795-806, November 2013. [click]

[15] F. Ding, Y. J. Wang, and J. Ding, "Recursive least squares parameter identification for systems with colored noise using the filtering technique and the auxiliary model," *Digital Signal Processing*, vol. 37, pp. 100-108, 2015. [click]

[16] A. Haryanto and K. S. Hong, "Maximum likelihood identification of Wiener-Hammerstein models," *Mechanical Systems and Signal Processing*, vol. 41, no. 1-2, pp. 54-70, December 2013. [click]

[17] G. Q. Li, C. Y. Wen, W. X. Zheng, and Y. Chen, "Identification of a class of nonlinear autoregressive models with exogenous inputs based on kernel machines," *IEEE Trans. on Signal Processing*, vol. 59, no. 5, pp. 2146-2159, May 2011. [click]

[18] J. Paduart, L. Lauwers, R. Pintelon, and J. Schoukens, "Identification of a Wiener–Hammerstein system using the polynomial nonlinear state space approach," *Control Engineering Practice*, vol. 20, no. 11, pp. 1133-1139, November 2012. [click]

[19] H. Karimi, and K. B. McAuley, "A maximum-likelihood method for estimating parameters, stochastic disturbance intensities and measurement noise variances in nonlinear dynamic models with process disturbances," *Computers & Chemical Engineering*, vol. 67, pp. 178-198, August 2014. [click]

[20] C. Q. Song, J. E. Feng, X. D. Wang, and J. L. Zhao, "Finite iterative method for solving coupled Sylvester-transpose matrix equations," *Journal of Applied Mathematics and Computing*, vol. 46, no. 1-2, pp. 351-372, October 2014. [click]

[21] M. Jafari, M. Salimifard, and M. Dehghani, "Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares algorithm," *ISA Transactions*, vol. 53, no. 4, pp. 1243-1252, July 2014. [click]

[22] Y. Shi, and T. Chen, "Optimal design of multichannel transmultiplexers with stop-band energy and pass-band magnitude constraints," *IEEE Trans. on Circuits and Systems II-Analog and Digital Signal Processing*, vol. 50, no. 9, pp. 659-662, September 2003. [click]

[23] M. Hajarian, "Matrix iterative methods for solving the Sylvester-transpose and periodic Sylvester matrix equations," *Journal of the Franklin Institute*, vol. 350, no. 10, pp. 3328-3341, December 2013. [click]

[24] G. Q. Li, and C. Y. Wen, "Convergence of normalized iterative identification of Hammerstein systems," *Systems & Control Letters*, vol. 60, no. 11, pp. 929-935, November 2011. [click]

[25] L. Xie, H. Z. Yang, "Interactive parameter estimation for output error moving average systems," *Transactions of the Institute of Measurement and Control*, vol. 35, no. 1, pp. 34-43, February 2013. [click]

[26] C. Wang, and T. Tang, "Several gradient-based iterative estimation algorithms for a class of nonlinear systems using the filtering technique," *Nonlinear Dynamics*, vol. 77, no. 3, pp. 769-780, August 2014. [click]

[27] J. Ding, C. X. Fan, and J. X. Lin, "Auxiliary model based parameter estimation for dual-rate output error systems with colored noise," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 4051-4058, March 2013. [click]

[28] J. Vörös, "Recursive identification of Hammerstein systems with discontinuous nonlinearities containing dead-zones," *IEEE Trans. on Automatic Control*, vol. 48, no. 12, pp. 2203-2206, December 2003. [click]

[29] K. P. Deng, and F. Ding, "Newton iterative identification method for an input nonlinear finite impulse response system with moving average noise using the key variables separation technique," *Nonlinear Dynamics*, vol. 76, no. 2, pp. 1195-1202, April 2014. [click]

[30] F. Ding and T. Chen, "Identification of Hammerstein nonlinear ARMAX systems," *Automatica*, vol. 41, no. 9, pp. 1479-1489, September 2005. [click]

[31] Y. J. Wang and F. Ding, "Iterative estimation for a nonlinear IIR filter with moving average noise by means of the data filtering technique," *IMA Journal of Mathematical Control and Information*, 2016. [click]

[32] J. Vörös, "Modeling and parameter identification of systems with multi-segment piecewise-linear characteristics," *IEEE Trans. on Automatic Control*, vol. 47, no. 1, pp. 184-188, January 2002. [click]

[33] J. H. Li, "Parameter estimation for Hammerstein CARARMA systems based on the Newton iteration," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 91-96, January 2013. [click]

[34] J. Ding, C. X. Fan, and J. X. Lin, "Auxiliary model based parameter estimation for dual-rate output error systems with colored noise," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 4051-4058, March 2013. [click]

[35] F. Ding, X. M. Liu, and Y. Gu, "An auxiliary model based least squares algorithm for a dual-rate state space system with time-delay using the data filtering," *Journal of the Franklin Institute*, vol. 353, no. 2, pp. 398-408, February 2016. [click]

[36] F. Ding, X. H. Wang, Q. J. Chen, and Y. S. Xiao, "Recursive least squares parameter estimation for a class of output nonlinear systems based on the model decomposition," *Circuits, Systems and Signal Processing*, vol. 35, 2016. [click]

[37] D. L. Zhang, Y. G. Tang, J. H. Ma, and X. P. Guan, "Identification of wiener model with discontinuous nonlinearities using differential evolution," *International Journal of Control, Automation, and Systems*, vol. 11, no. 3, pp. 511-518, June 2013. [click]

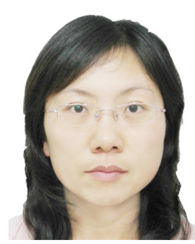
[38] J. Huang, Y. Shi, H. N. Huang, and Z. Li, "l-2–l-infinity filtering for multirate nonlinear sampled-data systems using T-S fuzzy models," *Digital Signal Processing*, vol. 23, no. 1, pp. 418-426, January 2013. [click]

[39] B. Sun, D. Q. Zhu, and S.X. Yang, "A bioinspired filtered backstepping tracking control of 7000-m manned submarine vehicle," *IEEE Trans. on Industrial Electronics*, vol. 61, no. 7, pp. 3682-3693 , July 2014. [click]

[40] B. Sun, D. Q. Zhu, and S. X. Yang, "A bio-inspired cascaded approach for three-dimensional tracking control of unmanned underwater vehicles," *International Journal of Robotics and Automation* , vol. 29, no. 4, pp. 349-358, October 2014. [click]

[41] D. Q. Zhu, X. Hua, and B. Sun, "A neurodynamics control strategy for real-time tracking control of autonomous underwater vehicles," *Journal of Navigation*, vol. 67, no. 1, pp. 113-127, January 2014. [click]

Junxia Ma received her B.Sc. degree in School of Communication and Control Engineering and M.Sc. degree in School of Internet of Things Engineering from Jiangnan University (Wuxi, China) in 2006 and 2014, respectively. She is currently a Ph.D candidate at the Jiangnan University. Her research interests include System Identification and Processing Control.



Weili Xiong received her Ph.D. degree in School of Communication and Control Engineering from Jiangnan University (Wuxi, China) in 2007. She has been a professor in the School of Internet of Things Engineering at the Jiangnan University. She was a visiting scholar in Department of Chemical and Materials Engineering, University of Alberta, Edmonton,Canada from 2013 to 2014. Her research interests include system identification, soft sensor of industry processes and optimization.

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. He is a Colleges and Universities "Blue Project" Middle-Aged Academic Leader, Jiangsu, China. His current research interests include model identification and adaptive control. He authored two books: System Identification – New Theory and Methods (Science University Press, Beijing, 2013), and System Identification – Performance Analysis for Identification Methods (Science University Press, Beijing, 2014).