Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 596141, 7 pages
http://dx.doi.org/10.1155/2013/596141

Research Article

Multi-Innovation Stochastic Gradient Identification Algorithm for Hammerstein Controlled Autoregressive Autoregressive Systems Based on the Key Term Separation Principle and on the Model Decomposition

Huiyi Hu,[1] Xiao Yongsong,[1] and Rui Ding[2]

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

Correspondence should be addressed to Rui Ding; rding12@126.com

Received 10 June 2013; Revised 22 August 2013; Accepted 6 September 2013

Academic Editor: Reinaldo Martinez Palhares

An input nonlinear system is decomposed into two subsystems, one including the parameters of the system model and the other including the parameters of the noise model, and a multi-innovation stochastic gradient algorithm is presented for Hammerstein controlled autoregressive autoregressive (H-CARAR) systems based on the key term separation principle and on the model decomposition, in order to improve the convergence speed of the stochastic gradient algorithm. The key term separation principle can simplify the identification model of the input nonlinear system, and the decomposition technique can enhance computational efficiencies of identification algorithms. The simulation results show that the proposed algorithm is effective for estimating the parameters of IN-CARAR systems.

1. Introduction

There exist many nonlinear systems in process control [1–3]. A nonlinear system can be modeled by input nonlinear systems [4] and output nonlinear systems [5], input-output nonlinear systems [6], feedback nonlinear systems [7], and so on. Input nonlinear systems, which are called Hammerstein systems [8], include input nonlinear equation error type systems and input nonlinear output error type systems. Recently, many identification algorithms have been developed for input nonlinear systems, such as the iterative methods [9–11], the separable least squares methods [12, 13], the blind methods [14], the subspace methods [15], and the overparameterization methods [16, 17]. Some methods require paying much extra computation.

The stochastic gradient (SG) algorithm is widely applied to parameter estimation. For example, Wang and Ding presented an extended SG identification algorithm for Hammerstein-Wiener ARMAX systems [18], but it is well known that the SG algorithm has slower convergence rates. In order to improve the convergence rate of the SG algorithm, Xiao et al. presented a multi-innovation stochastic gradient parameter estimation algorithm for input nonlinear controlled autoregressive (IN-CAR) models using the over-parameterization method [19]; Chen et al. proposed a modified stochastic gradient algorithm by introducing a convergence index in order to improve the convergence rate of the parameter estimation [20]; Han and Ding developed a multi-innovation stochastic gradient algorithm for multi-input single-output systems [21]; Liu et al. studied the performance of the stochastic gradient algorithm for multivariable systems [22].

The decomposition identification techniques include matrix decomposition and model decomposition. Hu and Ding presented a least squares based iterative identification algorithm for controlled moving average systems using the matrix decomposition [23]; Ding derived an iterative least squares algorithm to estimate the parameters of output error systems, and the matrix decomposition can enhance computational efficiencies [24]. Ding also divided a Hammerstein nonlinear system into two subsystems based on the model decomposition and presented a hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems [25].

This paper discusses identification problems of input nonlinear controlled autoregressive autoregressive (IN-CARAR) systems or Hammerstein controlled autoregressive autoregressive (H-CARAR) systems, which is one kind of input nonlinear equation error type systems. The basic idea is using the key term separation principle [26] and the decomposition technique [24] to derive a multi-innovation stochastic gradient identification algorithm, which is different from the work in [19, 21, 25].

The rest of this paper is organized as follows. Section 2 gives the identification model for the IN-CARAR systems. Section 3 introduces the SG algorithm for the IN-CARAR system. Section 4 deduces a multi-innovation SG algorithm for IN-CARAR system using the decomposition technique. Section 5 provides a numerical example to show the effectiveness of the proposed algorithm. Finally, Section 6 offers some concluding remarks.

2. The System Identification Model

The paper focuses on the parameter estimation of a Hammerstein nonlinear controlled autoregressive autoregressive (H-CARAR) system, that is, an input nonlinear controlled autoregressive autoregressive (IN-CARAR) system, which consists of a nonlinear block and a linear dynamic subsystem. It is worth noting that Xiao and Yue discussed a data filtering based recursive least squares algorithm for H-CARAR systems [27] and a multi-innovation stochastic gradient parameter estimation algorithm for input nonlinear controlled autoregressive (IN-CAR) systems.

An IN-CARAR system shown as Figure 1 is expressed as [27]

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

$$w(t) = \frac{1}{C(z)}v(t), \tag{2}$$

$$\bar{u}(t) = \mathbf{f}(u(t))\,\gamma, \tag{3}$$

where $u(t)$ and $y(t)$ are the system input and output, $\bar{u}(t)$ is the output of the nonlinear part, and $v(t)$ is an uncorrelated stochastic noise with zero mean. Here, $\bar{u}(t)$ is expressed as [28]

$$\begin{aligned}\bar{u}(t) &= \mathbf{f}(u(t))\,\gamma = \sum_{i=1}^{n_\gamma}\gamma_i f_i(u(t)) \\ &= \gamma_1 f_1(u(t)) + \gamma_2 f_2(u(t)) + \cdots + \gamma_{n_\gamma} f_{n_\gamma}(u(t)),\end{aligned} \tag{4}$$

where $\mathbf{f}(u(t)) := [f_1(u(t)), f_2(u(t)), \ldots, f_{n_\gamma}(u(t))] \in \mathbb{R}^{1\times n_\gamma}$ and $\gamma := [\gamma_1, \gamma_2, \ldots, \gamma_{n_\gamma}]^T \in \mathbb{R}^{n_\gamma}$ is the parameters vector of the nonlinear part.

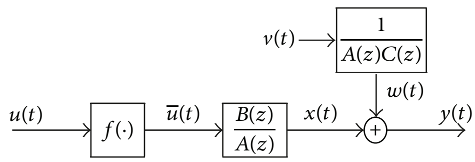

Figure 1: The IN-CARAR system.

In (1), $A(z)$, $B(z)$, and $C(z)$ are the polynomials, of known orders n_a, n_b, and n_c, in the unit backward shift operator z^{-1} $[z^{-1}y(t) = y(t-1)]$, defined by

$$\begin{aligned}A(z) &:= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{n_a} z^{-n_a}, \\ B(z) &:= 1 + b_1 z^{-1} + b_2 z^{-2} + \cdots + b_{n_b} z^{-n_b}, \\ C(z) &:= 1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_{n_c} z^{-n_c}.\end{aligned} \tag{5}$$

Assume $y(t) = 0$, $u(t) = 0$, and $v(t) = 0$ for $t \leqslant 0$. a_i, b_i, and c_i are the parameters to be estimated from measured input–output data $\{u(t), y(t)\}$.

Define the parameter vectors:

$$\begin{aligned}\boldsymbol{\theta} &:= \left[\boldsymbol{\theta}_s^T, \boldsymbol{\theta}_n^T\right]^T \in \mathbb{R}^{n_a+n_b+n_\gamma+n_c}, \\ \boldsymbol{\theta}_s &:= \left[\mathbf{a}^T, \mathbf{b}^T, \gamma^T\right]^T \in \mathbb{R}^{n_a+n_b+n_\gamma}, \\ \mathbf{a} &:= \left[a_1, a_2, \ldots, a_{n_a}\right]^T \in \mathbb{R}^{n_a}, \\ \mathbf{b} &:= \left[b_1, b_2, \ldots, b_{n_b}\right]^T \in \mathbb{R}^{n_b}, \\ \boldsymbol{\theta}_n &:= \left[c_1, c_2, \ldots, c_{n_c}\right]^T \in \mathbb{R}^{n_c},\end{aligned} \tag{6}$$

and the information vectors:

$$\begin{aligned}\boldsymbol{\varphi}(t) &:= \begin{bmatrix}\boldsymbol{\varphi}_s(t) \\ \boldsymbol{\varphi}_n(t)\end{bmatrix} \in \mathbb{R}^{n_a+n_b+n_\gamma+n_c}, \\ \boldsymbol{\varphi}_s(t) &:= \left[-y(t-1), -y(t-2), \ldots, -y(t-n_a),\right. \\ &\qquad \left.\bar{u}(t-1), \ldots, \bar{u}(t-n_b), \mathbf{f}(u(t))\right]^T \in \mathbb{R}^{n_a+n_b+n_\gamma}, \\ \boldsymbol{\varphi}_n(t) &:= \left[-w(t-1), -w(t-2), \ldots, -w(t-n_c)\right]^T \in \mathbb{R}^{n_c}.\end{aligned} \tag{7}$$

Equation (2) can be written as

$$\begin{aligned}w(t) &= [1 - C(z)]\,w(t) + v(t) \\ &= -\sum_{i=1}^{n_c} c_i w(t-i) + v(t) \\ &= \boldsymbol{\varphi}_n^T(t)\,\boldsymbol{\theta}_n + v(t).\end{aligned} \tag{8}$$

Using the key term separation principle [26], (1) can be written as

$$\begin{aligned} y(t) &= [1 - A(z)]\, y(t) + [B(z) - 1]\, \bar{u}(t) + \bar{u}(t) + w(t) \\ &= -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{i=1}^{n_b} b_i \bar{u}(t-i) + \sum_{i=1}^{n_\gamma} \gamma_i \mathbf{f}_i(u(t)) + w(t) \\ &= \boldsymbol{\varphi}_s^T(t)\boldsymbol{\theta}_s + w(t) \end{aligned} \tag{9}$$

$$= \boldsymbol{\varphi}_s^T(t)\boldsymbol{\theta}_s + \boldsymbol{\varphi}_n^T(t)\boldsymbol{\theta}_n + v(t) \tag{10}$$

$$= \boldsymbol{\varphi}^T(t)\boldsymbol{\theta}. \tag{11}$$

This is the identification model of the IN-CARAR system.

3. The Stochastic Gradient Algorithm

According to [1] and based on the identification model in (11), we can obtain the stochastic gradient (SG) algorithm:

$$\begin{aligned} \widehat{\boldsymbol{\theta}}(t) &= \widehat{\boldsymbol{\theta}}(t-1) + \frac{\widehat{\boldsymbol{\varphi}}(t)}{r(t)} e(t), \\ e(t) &= y(t) - \widehat{\boldsymbol{\varphi}}^T(t)\widehat{\boldsymbol{\theta}}(t-1), \\ r(t) &= r(t-1) + \|\widehat{\boldsymbol{\varphi}}(t)\|^2, \qquad r(0) = 1, \\ \widehat{\boldsymbol{\varphi}}(t) &= \begin{bmatrix} \widehat{\boldsymbol{\varphi}}_s(t) \\ \widehat{\boldsymbol{\varphi}}_n(t) \end{bmatrix}, \\ \widehat{\boldsymbol{\varphi}}_s(t) &= \big[-y(t-1), -y(t-2), \ldots, -y(t-n_a), \\ &\qquad \widehat{\bar{u}}(t-1), \widehat{\bar{u}}(t-2), \ldots, \widehat{\bar{u}}(t-n_b), \mathbf{f}(u(t)) \big]^T, \\ \widehat{\boldsymbol{\varphi}}_n(t) &= [-\widehat{w}(t-1), -\widehat{w}(t-2), \ldots, -\widehat{w}(t-n_c)]^T, \\ \widehat{\bar{u}}(t) &= \mathbf{f}(u(t))\,\widehat{\gamma}(t), \\ \widehat{w}(t) &= y(t) - \widehat{\boldsymbol{\varphi}}_s^T(t)\widehat{\boldsymbol{\theta}}_s(t), \\ \mathbf{f}(u(t)) &= \left[f_1(u(t)), f_2(u(t)), \ldots, f_{n_r}(u(t)) \right], \\ \widehat{\boldsymbol{\theta}}(t) &= \begin{bmatrix} \widehat{\boldsymbol{\theta}}_n(t) \\ \widehat{\boldsymbol{\theta}}_s(t) \end{bmatrix}, \\ \widehat{\boldsymbol{\theta}}_s(t) &= \left[\widehat{\mathbf{a}}^T(t), \widehat{\mathbf{b}}^T(t), \widehat{\gamma}^T(t) \right]^T, \\ \widehat{\mathbf{a}}(t) &= \left[\widehat{a}_1(t), \widehat{a}_2(t), \ldots, \widehat{a}_{n_a}(t) \right]^T, \\ \widehat{\mathbf{b}}(t) &= \left[\widehat{b}_1(t), \widehat{b}_2(t), \ldots, \widehat{b}_{n_b}(t) \right]^T, \\ \widehat{\gamma}(t) &= \left[\widehat{\gamma}_1(t), \widehat{\gamma}_2(t), \ldots, \widehat{\gamma}_{n_\gamma}(t) \right]^T, \\ \widehat{\boldsymbol{\theta}}_n(t) &= \left[\widehat{c}_1(t), \widehat{c}_2(t), \ldots, \widehat{c}_{n_c}(t) \right]^T, \end{aligned} \tag{12}$$

where $\widehat{X}(t)$ represents the estimate of X at time t; for example, $\widehat{\boldsymbol{\theta}}(t) = \left[\begin{smallmatrix} \widehat{\boldsymbol{\theta}}_n(t) \\ \widehat{\boldsymbol{\theta}}_s(t) \end{smallmatrix}\right] \in \mathbb{R}^{n_a+n_b+n_\gamma+n_c}$ is the estimate of $\boldsymbol{\theta} = \left[\begin{smallmatrix} \boldsymbol{\theta}_s \\ \boldsymbol{\theta}_n \end{smallmatrix}\right]$ at time t.

4. The Multi-Innovation Stochastic Gradient Algorithm

This section deduces the multi-innovation stochastic gradient identification algorithm for the IN-CARAR system using the decomposition technique [1].

Define two intermediate variables,

$$\begin{aligned} y_1(t) &:= y(t) - \boldsymbol{\varphi}_n^T(t)\boldsymbol{\theta}_n, \\ y_2(t) &:= y(t) - \boldsymbol{\varphi}_s^T(t)\boldsymbol{\theta}_s. \end{aligned} \tag{13}$$

From (10), we have

$$\begin{aligned} y_1(t) &:= \boldsymbol{\varphi}_s^T(t)\boldsymbol{\theta}_s + v(t), \\ y_2(t) &:= \boldsymbol{\varphi}_n^T(t)\boldsymbol{\theta}_n + v(t). \end{aligned} \tag{14}$$

These two subsystems include the parameter vectors $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$, respectively. $\boldsymbol{\theta}_s$ contains the parameters od the system model and $\boldsymbol{\theta}_n$ contains the parameters od the noise model.

Define the stacked information matrices and the stacked output vectors:

$$\begin{aligned} \mathbf{Y}(p,t) &:= [y(t), y(t-1), \ldots, y(t-p+1)]^T \in \mathbb{R}^p, \\ \mathbf{Y}_1(p,t) &:= [y_1(t), y_1(t-1), \ldots, y_1(t-p+1)]^T \in \mathbb{R}^p, \\ \mathbf{Y}_2(p,t) &:= [y_2(t), y_2(t-1), \ldots, y_2(t-p+1)]^T \in \mathbb{R}^p, \\ \boldsymbol{\Phi}_s(p,t) & \\ &:= [\boldsymbol{\varphi}_s(t), \boldsymbol{\varphi}_s(t-1), \ldots, \boldsymbol{\varphi}_s(t-p+1)]^T \in \mathbb{R}^{p\times(n_a+n_b+n_c)}, \\ \widehat{\boldsymbol{\Phi}}_n(p,t) &:= [\widehat{\boldsymbol{\varphi}}_n(t), \widehat{\boldsymbol{\varphi}}_n(t-1), \ldots, \widehat{\boldsymbol{\varphi}}_n(t-p+1)]^T \in \mathbb{R}^{p\times n_\gamma}, \\ \mathbf{E}_s(p,t) &:= [e_s(t), e_s(t-1), \ldots, e_s(t-p+1)]^T \in \mathbb{R}^p, \\ \mathbf{E}_n(p,t) &:= [e_n(t), e_n(t-1), \ldots, e_n(t-p+1)]^T \in \mathbb{R}^p. \end{aligned} \tag{15}$$

According to the multi-innovation identification theory [29–41], we expand the scalar innovations:

$$\begin{aligned} e_s(t) &= y_1(t) - \boldsymbol{\varphi}_s^T(t)\widehat{\boldsymbol{\theta}}_s(t-1), \\ e_n(t) &= y_2(t) - \widehat{\boldsymbol{\varphi}}_n^T(t)\widehat{\boldsymbol{\theta}}_n(t-1), \end{aligned} \tag{16}$$

to the innovation vectors,

$$\begin{aligned} \mathbf{E}_s(p,t) &= \widehat{\mathbf{Y}}_1(p,t) - \boldsymbol{\Phi}_s^T(p,t)\widehat{\boldsymbol{\theta}}_s(t-1), \\ \mathbf{E}_n(p,t) &= \widehat{\mathbf{Y}}_2(p,t) - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\widehat{\boldsymbol{\theta}}_n(t-1). \end{aligned} \tag{17}$$

Define two criterion functions,

$$\begin{aligned} J_1(\boldsymbol{\theta}_s) &:= \left\| \mathbf{Y}_1(p,t) - \boldsymbol{\Phi}_s^T(p,t)\boldsymbol{\theta}_s \right\|^2, \\ J_2(\boldsymbol{\theta}_n) &:= \left\| \mathbf{Y}_2(p,t) - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\boldsymbol{\theta}_n \right\|^2. \end{aligned} \tag{18}$$

Table 1: The SG parameter estimates and errors.

t	a_1	a_2	b_1	b_2	γ_1	γ_2	γ_3	c_1	c_2	δ (%)
100	0.49020	−0.49380	0.01083	0.01505	0.00905	0.01482	0.02428	−0.00163	−0.00411	95.19650
200	0.49438	−0.49611	0.01052	0.01544	0.00908	0.01488	0.02437	−0.00159	−0.00415	95.15025
500	0.50037	−0.50067	0.01024	0.01579	0.00910	0.01491	0.02443	−0.00156	−0.00418	95.11481
1000	0.50136	−0.50130	0.01017	0.01588	0.00911	0.01492	0.02444	−0.00156	−0.00418	95.10595
2000	0.50017	−0.50007	0.01016	0.01589	0.00911	0.01492	0.02444	−0.00156	−0.00418	95.10507
3000	0.50042	−0.50031	0.01016	0.01589	0.00911	0.01492	0.02444	−0.00156	−0.00418	95.10474
True values	1.80000	0.80000	0.50000	0.65000	1.00000	0.50000	0.25000	0.30000	0.20000	

Table 2: The MISG parameter estimates and errors.

t	a_1	a_2	b_1	b_2	γ_1	γ_2	γ_3	c_1	c_2	δ (%)
100	1.79885	0.80074	0.49998	0.65013	0.99968	0.50004	0.24953	0.21240	0.36032	7.46185
200	1.79813	0.80153	0.49999	0.65014	0.99955	0.50000	0.24927	0.37423	0.25902	3.87441
500	1.79820	0.80154	0.50001	0.65020	0.99927	0.49994	0.24876	0.32965	0.06371	5.69751
1000	1.79998	0.79973	0.50000	0.65023	0.99910	0.49990	0.24844	0.32849	0.19041	1.23012
2000	1.79979	0.79992	0.50000	0.65023	0.99909	0.49990	0.24842	0.28734	0.20472	0.55706
3000	1.79989	0.79982	0.50000	0.65023	0.99909	0.49990	0.24842	0.28556	0.20496	0.62823
True values	1.80000	0.80000	0.50000	0.65000	1.00000	0.50000	0.25000	0.30000	0.20000	

The gradients of J_1 and J_2 with respect to $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$, respectively, are

$$\begin{aligned}
\operatorname{grad}\left[J_1(\boldsymbol{\theta}_s)\right] &= \frac{\partial J_1(\boldsymbol{\theta}_s)}{\partial \boldsymbol{\theta}_n} \\
&= -2\boldsymbol{\Phi}_s(p,t)\left[\mathbf{Y}_1(p,t) - \boldsymbol{\Phi}_s^T(p,t)\boldsymbol{\theta}_s\right], \\
\operatorname{grad}\left[J_2(\boldsymbol{\theta}_n)\right] &= \frac{\partial J_2(\boldsymbol{\theta}_n)}{\partial \boldsymbol{\theta}_n} \\
&= -2\widehat{\boldsymbol{\Phi}}_n(p,t)\left[\mathbf{Y}_2(p,t) - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\boldsymbol{\theta}_n\right].
\end{aligned} \tag{19}$$

Minimizing $J_1(\boldsymbol{\theta}_s)$ and $J_2(\boldsymbol{\theta}_n)$ using the negative gradient search, we can obtain the multi-innovation stochastic gradient algorithm (MISG) for the IN-CARAR system:

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_s(t) &= \widehat{\boldsymbol{\theta}}_s(t-1) + \frac{\widehat{\boldsymbol{\Phi}}_s(p,t)}{r_1(t)}\mathbf{E}_s(p,t), \\
\mathbf{E}_s(p,t) &= \widehat{\mathbf{Y}}_1(p,t) - \widehat{\boldsymbol{\Phi}}_s^T(p,t)\widehat{\boldsymbol{\theta}}_s(t-1) \\
&= \mathbf{Y}(p,t) - \widehat{\boldsymbol{\Phi}}_s^T(p,t)\widehat{\boldsymbol{\theta}}_s(t-1) \\
&\quad - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\widehat{\boldsymbol{\theta}}_n(t-1), \\
r_1(t) &= r_1(t-1) + \left\|\widehat{\boldsymbol{\Phi}}_s(p,t)\right\|^2, \quad r_1(0) = 1, \\
\mathbf{Y}(p,t) &= \left[y(t), y(t-1), \ldots, y(t-p+1)\right]^T, \\
\widehat{\boldsymbol{\Phi}}_s(p,t) &= \left[\widehat{\boldsymbol{\varphi}}_s(t), \widehat{\boldsymbol{\varphi}}_s(t-1), \ldots, \widehat{\boldsymbol{\varphi}}_s(t-p+1)\right]^T, \\
\widehat{\boldsymbol{\theta}}_n(t) &= \widehat{\boldsymbol{\theta}}_n(t-1) + \frac{\widehat{\boldsymbol{\Phi}}_n(p,t)}{r_2(t)}\mathbf{E}_n(p,t), \\
\mathbf{E}_n(p,t) &= \widehat{\mathbf{Y}}_2(p,t) - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\widehat{\boldsymbol{\theta}}_n(t-1) \\
&= \mathbf{Y}(p,t) - \widehat{\boldsymbol{\Phi}}_s^T(p,t)\widehat{\boldsymbol{\theta}}_s(t-1) \\
&\quad - \widehat{\boldsymbol{\Phi}}_n^T(p,t)\widehat{\boldsymbol{\theta}}_n(t-1), \\
r_2(t) &= r_2(t-1) + \left\|\widehat{\boldsymbol{\Phi}}_n(p,t)\right\|^2, r_2(0) = 1, \\
\widehat{\boldsymbol{\Phi}}_n(p,t) &= \left[\widehat{\boldsymbol{\varphi}}_n(t), \widehat{\boldsymbol{\varphi}}_n(t-1), \ldots, \widehat{\boldsymbol{\varphi}}_n(t-p+1)\right]^T, \\
\widehat{\boldsymbol{\varphi}}_s(t) &= \Big[-y(t-1), -y(t-2), \ldots, -y(t-n_a), \\
&\qquad \widehat{\bar{u}}(t-1), \widehat{\bar{u}}(t-2), \ldots, \widehat{\bar{u}}(t-n_b), \mathbf{f}(u(t))\Big]^T, \\
\widehat{\boldsymbol{\varphi}}_n(t) &= \left[-\widehat{w}(t-1), -\widehat{w}(t-2), \ldots, -\widehat{w}(t-n_c)\right]^T, \\
\widehat{\bar{u}}(t) &= \mathbf{f}(u(t))\,\widehat{\gamma}(t), \\
\widehat{w}(t) &= y(t) - \widehat{\boldsymbol{\varphi}}_s^T(t)\widehat{\boldsymbol{\theta}}_s(t), \\
\mathbf{f}(u(t)) &= \left[f_1(u(t)), f_2(u(t)), \ldots, f_{n_\gamma}(u(t))\right], \\
\widehat{\boldsymbol{\theta}}(t) &= \begin{bmatrix} \widehat{\boldsymbol{\theta}}_n(t) \\ \widehat{\boldsymbol{\theta}}_s(t) \end{bmatrix}, \\
\widehat{\boldsymbol{\theta}}_s(t) &= \left[\widehat{\mathbf{a}}^T(t), \widehat{\mathbf{b}}^T(t), \widehat{\gamma}^T(t)\right]^T, \\
\widehat{\mathbf{a}}(t) &= \left[\widehat{a}_1(t), \widehat{a}_2(t), \ldots, \widehat{a}_{n_a}(t)\right]^T, \\
\widehat{\mathbf{b}}(t) &= \left[\widehat{b}_1(t), \widehat{b}_2(t), \ldots, \widehat{b}_{n_b}(t)\right]^T, \\
\widehat{\gamma}(t) &= \left[\widehat{\gamma}_1(t), \widehat{\gamma}_2(t), \ldots, \widehat{\gamma}_{n_\gamma}(t)\right]^T, \\
\widehat{\boldsymbol{\theta}}_n(t) &= \left[\widehat{c}_1(t), \widehat{c}_2(t), \ldots, \widehat{c}_{n_c}(t)\right]^T.
\end{aligned} \tag{20}$$

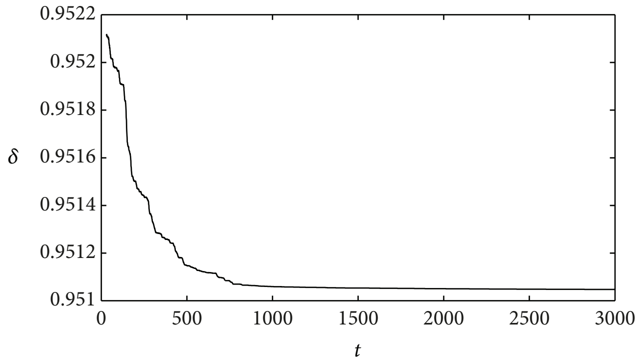

Figure 2: The SG estimation error δ versus t.

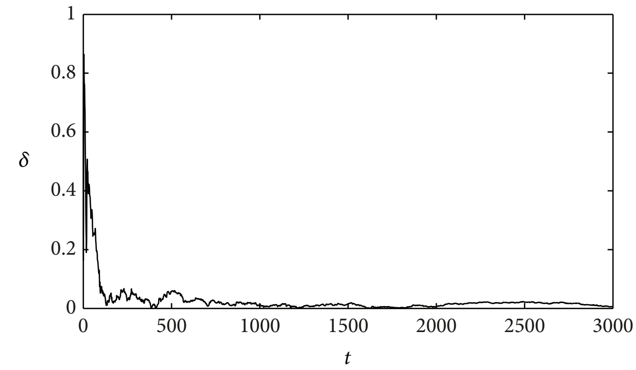

Figure 3: The MISG estimation error δ versus.

The initial values can be taken to be $\widehat{\boldsymbol{\theta}}(0) = \mathbf{1}_{n_a+n_b+n_\gamma+n_c}/p_0$, $\widehat{w}(i) = 1/p_0$, $i \leqslant 0$, and $p_0 = 10^6$.

5. Numerical Examples

Consider the following IN-CARAR system:

$$\begin{aligned}
A(z)\,y(t) &= B(z)\,\bar{u}(t) + \frac{1}{C(z)}v(t),\\
A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} = 1 + 1.80z^{-1} + 0.80z^{-2},\\
B(z) &= 1 + b_1 z^{-1} + b_2 z^{-2} = 1 + 0.50z^{-1} + 0.65z^{-2},\\
C(z) &= 1 + c_1 z^{-1} + c_2 z^{-2} = 1 + 0.30z^{-1} + 0.20z^{-2},\\
\bar{u}(t) &= \mathbf{f}(u(t))\,\gamma = \gamma_1 f_1(u(t)) + \gamma_2 f_2(u(t)) + \gamma_3 f_3(u(t))\\
&= 1.00u(t) + 0.50u^2(t) + 0.25u^3(t),\\
\boldsymbol{\theta} &= [a_1, a_2, b_1, b_2, \gamma_1, \gamma_2, \gamma_3, c_1, c_2]^T\\
&= [1.80, 0.80, 0.5, 0.65, 1.00, 0.50, 0.25, 0.30, 0.20]^T.
\end{aligned} \tag{21}$$

In this example, 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 $\sigma^2 = 0.50^2$. Applying the SG algorithm and the MISG algorithm to estimate the parameters of this IN-CARAR system, the parameter estimates and their estimation errors are shown in Tables 1 and 2 with the data length $L = 3000$ and the estimation error $\delta := \|\widehat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t being shown in Figures 2 and 3.

From Tables 1 and 2 and Figures 2 and 3, we can draw the conclusions. The parameters estimation errors become smaller with the data length t increasing. The estimation errors given by the MISG algorithm are much smaller than that of the SG algorithm. The convergence speed of the multi-innovation SG algorithm is faster than those of the SG algorithm. These indicate that the MISG algorithm has better performance than the SG algorithm.

6. Conclusions

The gradient and least squares algorithms are two different kinds of important identification methods. It is well known that the gradient algorithm has poor convergence rates. This paper studies the multi-innovation SG identification methods for IN-CARAR systems. The numerical examples show that the proposed MISG algorithm can estimate effectively the parameters of input nonlinear systems and indicate that increasing the innovation length can improve parameter estimation accuracy of the multi-innovation identification algorithm because the algorithm uses more information in each recursion for a large innovation length. The proposed method can be applied to nonlinear output error systems. Although the algorithm is presented for the IN-CARAR systems, the basic idea can be extended to other linear or nonlinear systems with colored noises [42–61].

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (nos. 61273194, 61203111), the University Graduate Scientific Research Innovation Program of Jiangsu Province (CXLX13-738), the Fundamental Research Funds for the Central Universities (JUDCF13035), the PAPD of Jiangsu Higher Education Institutions.

References

[1] F. Ding, *System Identification: New Theory and Methods*, Science Press, Beijing, China, 2013.

[2] Y. Shi and H. Fang, "Kalman filter-based identification for systems with randomly missing measurements in a network environment," *International Journal of Control*, vol. 83, no. 3, pp. 538–551, 2010.

[3] B. Yu, Y. Shi, and H. Huang, "l-2 and l-infnity filtering for multirate systems using lifted models," *Circuits, Systems, and Signal Processing*, vol. 27, no. 5, pp. 699–711, 2008.

[4] D. Wang, F. Ding, and Y. Chu, "Data filtering based recursive least squares algorithm for Hammerstein systems using the key-term separation principle," *Information Sciences*, vol. 222, pp. 203–212, 2013.

[5] D. Q. Wang and F. Ding, "Least squares based and gradient based iterative identification for Wiener nonlinear systems," *Signal Processing*, vol. 91, no. 5, pp. 1182–1189, 2011.

[6] D. Q. Wang and F. Ding, "Hierarchical least squares estimation algorithm for Hammerstein-Wiener systems," *IEEE Signal Processing Letters*, vol. 19, no. 12, pp. 825–828, 2012.

[7] P. P. Hu and F. Ding, "Multistage least squares based iterative estimation for feedback nonlinear systems with moving average noises using the hierarchical identification principle," *Nonlinear Dynamics*, vol. 73, no. 1-2, pp. 583–592, 2013.

[8] F. Ding, X. P. Liu, and G. Liu, "Identification methods for Hammerstein nonlinear systems," *Digital Signal Processing*, vol. 21, no. 2, pp. 215–238, 2011.

[9] M. Dehghan and M. Hajarian, "An iterative method for solving the generalized coupled Sylvester matrix equations over generalized bisymmetric matrices," *Applied Mathematical Modelling*, vol. 34, no. 3, pp. 639–654, 2010.

[10] M. Dehghan and M. Hajarian, "Two algorithms for finding the Hermitian reflexive and skew-Hermitian solutions of Sylvester matrix equations," *Applied Mathematics Letters*, vol. 24, no. 4, pp. 444–449, 2011.

[11] H. M. Zhang and F. Ding, "On the Kronecker products and their applications," *Journal of Applied Mathematics*, vol. 2013, Article ID 296185, 8 pages, 2013.

[12] E.-W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," *Automatica*, vol. 38, no. 5, pp. 853–860, 2002.

[13] F. Ding, X. Liu, and J. Chu, "Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle," *IET Control Theory & Applications*, vol. 7, no. 2, pp. 176–184, 2013.

[14] E.-W. Bai, "A blind approach to the Hammerstein-Wiener model identification," *Automatica*, vol. 38, no. 6, pp. 967–979, 2002.

[15] M. Verhaegen and D. Westwick, "Identifying MIMO Hammerstein systems in the context of subspace model identification methods," *International Journal of Control*, vol. 63, no. 2, pp. 331–349, 1996.

[16] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems & Control Letters*, vol. 56, no. 5, pp. 373–380, 2007.

[17] F. Ding, Y. Shi, and T. Chen, "Gradient-based identification methods for Hammerstein nonlinear ARMAX models," *Nonlinear Dynamics*, vol. 45, no. 1-2, pp. 31–43, 2006.

[18] D. Wang and F. Ding, "Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems," *Computers & Mathematics with Applications*, vol. 56, no. 12, pp. 3157–3164, 2008.

[19] Y. S. Xiao, G. L. Song, Y. W. Liao, and R. F. Ding, "Multi-innovation stochastic gradient parameter estimation for input nonlinear controlled autoregressive models," *International Journal of Control, Automation, and Systems*, vol. 10, no. 3, pp. 639–643, 2012.

[20] J. Chen, L. Lv, and R. Ding, "Multi-innovation stochastic gradient algorithms for dual-rate sampled systems with preload nonlinearity," *Applied Mathematics Letters*, vol. 26, no. 1, pp. 124–129, 2013.

[21] L. L. Han and F. Ding, "Multi-innovation stochastic gradient algorithms for multi-input multi-output systems," *Digital Signal Processing*, vol. 19, no. 4, pp. 545–554, 2009.

[22] Y. Liu, J. Sheng, and R. Ding, "Convergence of stochastic gradient estimation algorithm for multivariable ARX-like systems," *Computers & Mathematics with Applications*, vol. 59, no. 8, pp. 2615–2627, 2010.

[23] H. Hu and F. Ding, "An iterative least squares estimation algorithm for controlled moving average systems based on matrix decomposition," *Applied Mathematics Letters*, vol. 25, no. 12, pp. 2332–2338, 2012.

[24] F. Ding, "Decomposition based fast least squares algorithm for output error systems," *Signal Processing*, vol. 93, no. 5, pp. 1235–1242, 2013.

[25] F. Ding, "Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling," *Applied Mathematical Modelling*, vol. 37, no. 4, pp. 1694–1704, 2013.

[26] J. Vörös, "Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2145–2149, 1999.

[27] Y. Xiao and N. Yue, "Parameter estimation for nonlinear dynamical adjustment models," *Mathematical and Computer Modelling*, vol. 54, no. 5-6, pp. 1561–1568, 2011.

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

[29] F. Ding and T. Chen, "Performance analysis of multi-innovation gradient type identification methods," *Automatica*, vol. 43, no. 1, pp. 1–14, 2007.

[30] F. Ding, X. P. Liu, and G. Liu, "Multi-innovation least squares identification for system modeling," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 40, no. 3, pp. 767–778, 2010.

[31] F. Ding, X. P. Liu, and G. Liu, "Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises," *Signal Processing*, vol. 89, no. 10, pp. 1883–1890, 2009.

[32] F. Ding, "Several multi-innovation identification methods," *Digital Signal Processing*, vol. 20, no. 4, pp. 1027–1039, 2010.

[33] F. Ding, H. Chen, and M. Li, "Multi-innovation least squares identification methods based on the auxiliary model for MISO systems," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 658–668, 2007.

[34] D. Q. Wang and F. Ding, "Performance analysis of the auxiliary models based multi-innovation stochastic gradient estimation algorithm for output error systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 750–762, 2010.

[35] L. Xie, Y. J. Liu, H. Z. Yang, and F. Ding, "Modelling and identification for non-uniformly periodically sampled-data systems," *IET Control Theory & Applications*, vol. 4, no. 5, pp. 784–794, 2010.

[36] Y. Liu, L. Yu, and F. Ding, "Multi-innovation extended stochastic gradient algorithm and its performance analysis," *Circuits, Systems, and Signal Processing*, vol. 29, no. 4, pp. 649–667, 2010.

[37] Y. Liu, Y. Xiao, and X. Zhao, "Multi-innovation stochastic gradient algorithm for multiple-input single-output systems using the auxiliary model," *Applied Mathematics and Computation*, vol. 215, no. 4, pp. 1477–1483, 2009.

[38] L. L. Han and F. Ding, "Parameter estimation for multirate multi-input systems using auxiliary model and multiinnovation," *Journal of Systems Engineering and Electronics*, vol. 21, no. 6, pp. 1079–1083, 2010.

[39] L. Han and F. Ding, "Identification for multirate multi-input systems using the multi-innovation identification theory," *Computers & Mathematics with Applications*, vol. 57, no. 9, pp. 1438–1449, 2009.

[40] J. Zhang, F. Ding, and Y. Shi, “Self-tuning control based on multi-innovation stochastic gradient parameter estimation,” *Systems & Control Letters*, vol. 58, no. 1, pp. 69–75, 2009.

[41] F. Ding, G. Liu, and X. P. Liu, “Parameter estimation with scarce measurements,” *Automatica*, vol. 47, no. 8, pp. 1646–1655, 2011.

[42] W. Xiong, W. Fan, and R. Ding, “Least-squares parameter estimation algorithm for a class of input nonlinear systems,” *Journal of Applied Mathematics*, vol. 2012, Article ID 684074, 14 pages, 2012.

[43] F. Ding, “Coupled-least-squares identification for multivariable systems,” *IET Control Theory & Applications*, vol. 7, no. 1, pp. 68–79, 2013.

[44] F. Ding, “Two-stage least squares based iterative estimation algorithm for CARARMA system modeling,” *Applied Mathematical Modelling*, vol. 37, no. 7, pp. 4798–4808, 2013.

[45] J. Ding, F. Ding, X. P. Liu, and G. Liu, “Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data,” *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2677–2683, 2011.

[46] J. Ding and F. Ding, “Bias compensation-based parameter estimation for output error moving average systems,” *International Journal of Adaptive Control and Signal Processing*, vol. 25, no. 12, pp. 1100–1111, 2011.

[47] J. Ding, L. Han, and X. Chen, “Time series AR modeling with missing observations based on the polynomial transformation,” *Mathematical and Computer Modelling*, vol. 51, no. 5-6, pp. 527–536, 2010.

[48] F. Ding, Y. J. Liu, and B. Bao, “Gradient based and least squares based iterative estimation algorithms for multiinput multi-output systems,” *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 226, no. 1, pp. 43–55, 2012.

[49] F. Ding and Y. Gu, “Performance analysis of the auxiliary model-based least-squares identification algorithm for one-step state-delay systems,” *International Journal of Computer Mathematics*, vol. 89, no. 15, pp. 2019–2028, 2012.

[50] F. Ding and Y. Gu, “Performance analysis of the auxiliary model-based stochastic gradient parameter estimation algorithm for state-space systems with one-step state delay,” *Circuits, Systems, and Signal Processing*, vol. 32, no. 2, pp. 585–599, 2013.

[51] W. Wang, F. Ding, and J. Dai, “Maximum likelihood least squares identification for systems with autoregressive moving average noise,” *Applied Mathematical Modelling*, vol. 36, no. 5, pp. 1842–1853, 2012.

[52] J. Li, F. Ding, and G. Yang, “Maximum likelihood least squares identification method for input nonlinear finite impulse response moving average systems,” *Mathematical and Computer Modelling*, vol. 55, no. 3-4, pp. 442–450, 2012.

[53] F. Ding, G. Liu, and X. P. Liu, “Partially coupled stochastic gradient identification methods for non-uniformly sampled systems,” *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1976–1981, 2010.

[54] F. Ding and T. Chen, “Hierarchical gradient-based identification of multivariable discrete-time systems,” *Automatica*, vol. 41, no. 2, pp. 315–325, 2005.

[55] F. Ding and T. Chen, “Hierarchical least squares identification methods for multivariable systems,” *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 397–402, 2005.

[56] F. Ding and T. Chen, “Hierarchical identification of lifted state-space models for general dual-rate systems,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, no. 6, pp. 1179–1187, 2005.

[57] F. Ding, L. Qiu, and T. Chen, “Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems,” *Automatica*, vol. 45, no. 2, pp. 324–332, 2009.

[58] H. Q. Han, L. Xie, F. Ding, and X. G. Liu, “Hierarchical least squares based iterative identification for multivariable systems with moving average noises,” *Mathematical and Computer Modelling*, vol. 51, no. 9-10, pp. 1213–1220, 2010.

[59] Y. Liu, F. Ding, and Y. Shi, “Least squares estimation for a class of non-uniformly sampled systems based on the hierarchical identification principle,” *Circuits, Systems, and Signal Processing*, vol. 31, no. 6, pp. 1985–2000, 2012.

[60] Z. Zhang, F. Ding, and X. Liu, “Hierarchical gradient based iterative parameter estimation algorithm for multivariable output error moving average systems,” *Computers & Mathematics with Applications*, vol. 61, no. 3, pp. 672–682, 2011.

[61] D. Wang, R. Ding, and X. Dong, “Iterative parameter estimation for a class of multivariable systems based on the hierarchical identification principle and the gradient search,” *Circuits, Systems, and Signal Processing*, vol. 31, no. 6, pp. 2167–2177, 2012.