

Research Article

Least-Squares Parameter Estimation Algorithm for a Class of Input Nonlinear Systems

Weili Xiong,¹ Wei Fan,² and Rui Ding²

¹ Key Laboratory of Advanced Process Control for Light Industry of 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, rding2003@yahoo.cn

Received 18 March 2012; Accepted 26 April 2012

Academic Editor: Morteza Rafei

Copyright © 2012 Weili Xiong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies least-squares parameter estimation algorithms for input nonlinear systems, including the input nonlinear controlled autoregressive (IN-CAR) model and the input nonlinear controlled autoregressive autoregressive moving average (IN-CARARMA) model. The basic idea is to obtain linear-in-parameters models by overparameterizing such nonlinear systems and to use the least-squares algorithm to estimate the unknown parameter vectors. It is proved that the parameter estimates consistently converge to their true values under the persistent excitation condition. A simulation example is provided.

1. Introduction

Parameter estimation has received much attention in many areas such as linear and nonlinear system identification and signal processing [1–9]. Nonlinear systems can be simply divided into the input nonlinear systems, the output nonlinear systems, the feedback nonlinear systems, and the input and output nonlinear systems, and so forth. The Hammerstein models can describe a class of input nonlinear systems which consist of static nonlinear blocks followed by linear dynamical subsystems [10, 11].

Nonlinear systems are common in industrial processes, for example, the dead-zone nonlinearities and the valve saturation nonlinearities. Many estimation methods have been developed to identify the parameters of nonlinear systems, especially for Hammerstein nonlinear systems [12, 13]. For example, Ding et al. presented a least-squares-based iterative algorithm and a recursive extended least squares algorithm for Hammerstein ARMAX systems [14] and an auxiliary model-based recursive least squares algorithm for Hammerstein output error systems [15]. Wang and Ding proposed an extended stochastic gradient identification algorithm for Hammerstein-Wiener ARMAX Systems [16].

Recently, Wang et al. derived an auxiliary model-based recursive generalized least-squares parameter estimation algorithm for Hammerstein output error autoregressive systems and auxiliary model-based RELS and MI-ELS algorithms for Hammerstein output error moving average systems using the key term separation principle [17, 18]. Ding et al. presented a projection estimation algorithm and a stochastic gradient (SG) estimation algorithm for Hammerstein nonlinear systems by using the gradient search and further derived a Newton recursive estimation algorithm and a Newton iterative estimation algorithm by using the Newton method (Newton-Raphson method) [19]. Wang and Ding studied least-squares-based and gradient-based iterative identification methods for Wiener nonlinear systems [20].

Fan et al. discussed the parameter estimation problem for Hammerstein nonlinear ARX models [21]. On the basis of the work in [14, 15, 21], this paper studies the identification problems and their convergence for input nonlinear controlled autoregressive (IN-CAR) models using the martingale convergence theorem and gives the recursive generalized extended least-squares algorithm for input nonlinear controlled autoregressive autoregressive moving average (IN-CARARMA) models.

Briefly, the paper is organized as follows. Section 2 derives a linear-in-parameters identification model and gives a recursive least squares identification algorithm for input nonlinear CAR systems and analyzes the properties of the proposed algorithm. Section 4 gives the recursive generalized extended least squares algorithm for input nonlinear CARARMA systems. Section 5 provides an illustrative example to show the effectiveness of the proposed algorithms. Finally, we offer some concluding remarks in Section 6.

2. The Input Nonlinear CAR Model and Estimation Algorithm

Let us introduce some notations first. The symbol \mathbf{I} (\mathbf{I}_n) stands for an identity matrix of appropriate sizes ($n \times n$); the superscript T denotes the matrix transpose; $\mathbf{1}_n$ represents an n -dimensional column vector whose elements are 1; $|\mathbf{X}| = \det[\mathbf{X}]$ represents the determinant of the matrix \mathbf{X} ; the norm of a matrix \mathbf{X} is defined by $\|\mathbf{X}\|^2 = \text{tr}[\mathbf{X}\mathbf{X}^T]$; $\lambda_{\max}[\mathbf{X}]$ and $\lambda_{\min}[\mathbf{X}]$ represent the maximum and minimum eigenvalues of the square matrix \mathbf{X} , respectively; $f(t) = o(g(t))$ represents $f(t)/g(t) \rightarrow 0$ as $t \rightarrow \infty$; for $g(t) \geq 0$, we write $f(t) = O(g(t))$ if there exists a positive constant δ_1 such that $|f(t)| \leq \delta_1 g(t)$.

2.1. The Input Nonlinear CAR Model

Consider the following input nonlinear controlled autoregressive (IN-CAR) systems [14, 21]:

$$A(z)y(t) = B(z)\bar{u}(t) + v(t), \quad (2.1)$$

where $y(t)$ is the system output, $v(t)$ is a disturbance noise, the output of the nonlinear block $\bar{u}(t)$ is a nonlinear function of a known basis (f_1, f_2, \dots, f_m) of the system input $u(t)$ [19],

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)), \quad (2.2)$$

$A(z)$ and $B(z)$ are polynomials in the unit backward shift operator z^{-1} [$z^{-1}\mathbf{y}(t) = \mathbf{y}(t-1)$], defined as

$$\begin{aligned} A(z) &:= 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_nz^{-n}, \\ B(z) &:= b_1z^{-1} + b_2z^{-2} + b_3z^{-3} + \cdots + b_nz^{-n}. \end{aligned} \quad (2.3)$$

In order to obtain the identifiability of parameters b_i and c_i , without loss of generality, we suppose that $c_1 = 1$ or $b_1 = 1$ [14, 21].

Define the parameter vector $\boldsymbol{\vartheta}$ and information vector $\boldsymbol{\psi}(t)$ as

$$\begin{aligned} \boldsymbol{\vartheta} &:= [\mathbf{a}^T, c_1\mathbf{b}^T, c_2\mathbf{b}^T, \dots, c_m\mathbf{b}^T]^T \in \mathbb{R}^{n_0}, \quad n_0 := n + mn, \\ \mathbf{a} &:= [a_1, a_2, \dots, a_n]^T \in \mathbb{R}^n, \\ \mathbf{b} &:= [b_1, b_2, \dots, b_n]^T \in \mathbb{R}^n, \\ \boldsymbol{\psi}(t) &:= [\boldsymbol{\psi}_0^T(t), \boldsymbol{\psi}_1^T(t), \boldsymbol{\psi}_2^T(t), \dots, \boldsymbol{\psi}_m^T(t)]^T \in \mathbb{R}^{n_0}, \\ \boldsymbol{\psi}_0(t) &:= [-\mathbf{y}(t-1), -\mathbf{y}(t-2), \dots, -\mathbf{y}(t-n)]^T \in \mathbb{R}^n, \\ \boldsymbol{\psi}_j(t) &:= [f_j(u(t-1)), f_j(u(t-2)), \dots, f_j(u(t-n))]^T \in \mathbb{R}^n, \quad j = 1, 2, \dots, m. \end{aligned} \quad (2.4)$$

From (2.1), we have

$$\begin{aligned} \mathbf{y}(t) &= [1 - A(z)]\mathbf{y}(t) + B(z)\bar{\mathbf{u}}(t) + \mathbf{v}(t) \\ &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + \sum_{i=1}^n b_i \sum_{j=1}^m c_j f_j(u(t-i)) + \mathbf{v}(t) \\ &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + \sum_{j=1}^m \sum_{i=1}^n c_j b_i f_j(u(t-i)) + \mathbf{v}(t) \\ &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + c_1 b_1 f_1(u(t-1)) + c_1 b_2 f_1(u(t-2)) + \cdots + c_1 b_n f_1(u(t-n)) \\ &\quad + c_2 b_1 f_2(u(t-1)) + c_2 b_2 f_2(u(t-2)) + \cdots + c_2 b_n f_2(u(t-n)) + \cdots \\ &\quad + c_m b_1 f_m(u(t-1)) + c_m b_2 f_m(u(t-2)) + \cdots + c_m b_n f_m(u(t-n)) + \mathbf{v}(t) \\ &= \boldsymbol{\psi}^T(t) \boldsymbol{\vartheta} + \mathbf{v}(t). \end{aligned} \quad (2.5)$$

$$\begin{aligned} &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + c_1 b_1 f_1(u(t-1)) + c_1 b_2 f_1(u(t-2)) + \cdots + c_1 b_n f_1(u(t-n)) \\ &\quad + c_2 b_1 f_2(u(t-1)) + c_2 b_2 f_2(u(t-2)) + \cdots + c_2 b_n f_2(u(t-n)) + \cdots \\ &\quad + c_m b_1 f_m(u(t-1)) + c_m b_2 f_m(u(t-2)) + \cdots + c_m b_n f_m(u(t-n)) + \mathbf{v}(t) \end{aligned} \quad (2.6)$$

An alternative way is to define the parameter vector θ and information vector $\varphi(t)$ as

$$\begin{aligned}
 \theta &:= [\mathbf{a}^T, b_1 \mathbf{c}^T, b_2 \mathbf{c}^T, \dots, b_n \mathbf{c}^T]^T \in \mathbb{R}^{n_0}, \\
 \mathbf{a} &:= [a_1, a_2, \dots, a_n]^T \in \mathbb{R}^n, \\
 \mathbf{c} &:= [c_1, c_2, \dots, c_m]^T \in \mathbb{R}^m, \\
 \varphi(t) &:= [\varphi_0^T(t), \varphi_1^T(t), \varphi_2^T(t), \dots, \varphi_n^T(t)]^T \in \mathbb{R}^{n_0}, \\
 \varphi_0(t) &:= [-y(t-1), -y(t-2), \dots, -y(t-n)]^T \in \mathbb{R}^n, \\
 \varphi_j(t) &:= [f_1(u(t-j)), f_2(u(t-j)), \dots, f_m(u(t-j))]^T \in \mathbb{R}^m, \quad j = 1, 2, \dots, n.
 \end{aligned} \tag{2.7}$$

Then (2.5) can be written as

$$\begin{aligned}
 y(t) &= -\sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n \sum_{j=1}^m b_i c_j f_j(u(t-i)) \\
 &= -\sum_{i=1}^n a_i y(t-i) + b_1 c_1 f_1 u(t-1) + b_1 c_2 f_2 u(t-1) + \dots + b_1 c_m f_m u(t-1) \\
 &\quad + b_2 c_1 f_1 u(t-2) + b_2 c_2 f_2 u(t-2) + \dots + b_2 c_m f_m u(t-2) + \dots \\
 &\quad + b_n c_1 f_1 u(t-n) + b_n c_2 f_2 u(t-n) + \dots + b_n c_m f_m u(t-n) + v(t) \\
 &= \varphi^T(t) \theta + v(t).
 \end{aligned} \tag{2.8}$$

Equations (2.6) and (2.8) are both linear-in-parameters identification model for Hammerstein CAR systems by using parametrization.

2.2. The Recursive Least Squares Algorithm

Minimizing the cost function

$$J(\theta) := \sum_{j=1}^t [y(j) - \varphi^T(j) \theta]^2 \tag{2.9}$$

gives the following recursive least squares algorithm for computing the estimate $\hat{\theta}(t)$ of θ in (2.8):

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \mathbf{P}(t) \varphi(t) [y(t) - \varphi^T(t) \hat{\theta}(t-1)], \tag{2.10}$$

$$\mathbf{P}^{-1}(t) = \mathbf{P}^{-1}(t-1) + \varphi(t) \varphi^T(t), \quad \mathbf{P}(0) = p_0 \mathbf{I}. \tag{2.11}$$

Applying the matrix inversion formula [22]

$$(\mathbf{A} + \mathbf{BC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1} \quad (2.12)$$

to (2.11) and defining the gain vector $\mathbf{L}(t) = \mathbf{P}(t)\boldsymbol{\varphi}(t) \in \mathbb{R}^{n_0}$, the algorithm in (2.10)-(2.11) can be equivalently expressed as

$$\begin{aligned} \hat{\boldsymbol{\theta}}(t) &= \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t) \left[\mathbf{y}(t) - \boldsymbol{\varphi}(t)\hat{\boldsymbol{\theta}}(t-1) \right], \\ \mathbf{L}(t) &= \mathbf{P}(t)\boldsymbol{\varphi}(t) = \frac{\mathbf{P}(t-1)\boldsymbol{\varphi}(t)}{1 + \boldsymbol{\varphi}^T(t)\mathbf{P}(t-1)\boldsymbol{\varphi}(t)}, \\ \mathbf{P}(t) &= \mathbf{P}(t-1) - \frac{\mathbf{P}(t)\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^T(t)\mathbf{P}(t)}{1 + \boldsymbol{\varphi}^T(t)\mathbf{P}(t-1)\boldsymbol{\varphi}(t)} \\ &= [\mathbf{I} - \mathbf{L}(t)\boldsymbol{\varphi}^T(t)]\mathbf{P}(t-1), \quad \mathbf{P}(0) = p_0\mathbf{I}. \end{aligned} \quad (2.13)$$

To initialize the algorithm, we take p_0 to be a large positive real number, for example, $p_0 = 10^6$, and $\hat{\boldsymbol{\theta}}(0)$ to be some small real vector, for example, $\hat{\boldsymbol{\theta}}(0) = 10^{-6}\mathbf{1}_{n_0}$.

3. The Main Convergence Theorem

The following lemmas are required to establish the main convergence results.

Lemma 3.1 (Martingale convergence theorem: Lemma D.5.3 in [23, 24]). *If T_t , α_t , β_t are nonnegative random variables, measurable with respect to a nondecreasing sequence of σ algebra \mathcal{F}_{t-1} , and satisfy*

$$E[T_t | \mathcal{F}_{t-1}] \leq T_{t-1} + \alpha_t - \beta_t, \quad a.s., \quad (3.1)$$

then when $\sum_{t=1}^{\infty} \alpha_t < \infty$, one has $\sum_{t=1}^{\infty} \beta_t < \infty$, a.s. and $T_t \rightarrow T$, a.s. (a.s.: almost surely) a finite nonnegative random variable.

Lemma 3.2 (see [14, 21, 25]). *For the algorithm in (2.10)-(2.11), for any $\gamma > 1$, the covariance matrix $\mathbf{P}(t)$ in (2.11) satisfies the following inequality:*

$$\sum_{t=1}^{\infty} \frac{\boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)}{[\ln|\mathbf{P}^{-1}(t)|]^{\gamma}} < \infty, \quad a.s. \quad (3.2)$$

Theorem 3.3. *For the system in (2.8) and the algorithm in (2.10)-(2.11), assume that $\{v(t), \mathcal{F}_t\}$ is a martingale difference sequence defined on a probability space $\{\Omega, \mathcal{F}, P\}$, where $\{\mathcal{F}_t\}$ is the σ algebra sequence generated by the observations $\{\mathbf{y}(t), \mathbf{y}(t-1), \dots, \mathbf{u}(t), \mathbf{u}(t-1), \dots\}$ and the noise sequence $\{v(t)\}$ satisfies $E[v(t) | \mathcal{F}_{t-1}] = 0$, and $E[v^2(t) | \mathcal{F}_{t-1}] \leq \sigma^2 < \infty$, a.s [23], and $[\ln|\mathbf{P}^{-1}(t)|]^{\gamma} = o(\lambda_{\min}[\mathbf{P}^{-1}(t)])$, $\gamma > 1$. Then the parameter estimation error $\hat{\boldsymbol{\theta}}(t)$ converges to zero.*

Proof. Define the parameter estimation error vector $\tilde{\boldsymbol{\theta}}(t) := \hat{\boldsymbol{\theta}}(t) - \boldsymbol{\theta}$ and the stochastic Lyapunov function $T(t) := \tilde{\boldsymbol{\theta}}^T(t)\mathbf{P}^{-1}(t)\tilde{\boldsymbol{\theta}}(t)$. Let $\tilde{\mathbf{y}}(t) := \boldsymbol{\varphi}^T(t)\hat{\boldsymbol{\theta}}(t-1) - \boldsymbol{\varphi}^T(t)\boldsymbol{\theta} = \boldsymbol{\varphi}^T(t)\tilde{\boldsymbol{\theta}}(t-1)$. According to the definitions of $\tilde{\boldsymbol{\theta}}(t)$ and $T(t)$ and using (2.10) and (2.11), we have

$$\begin{aligned} \tilde{\boldsymbol{\theta}}(t) &= \tilde{\boldsymbol{\theta}}(t-1) + \mathbf{P}(t)\boldsymbol{\varphi}(t)[- \tilde{\mathbf{y}}(t) + v(t)], \\ T(t) &= T(t-1) - \left[1 - \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)\right]\tilde{\mathbf{y}}^2(t) + \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)v^2(t) \\ &\quad + 2\left[1 - \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)\right]\tilde{\mathbf{y}}(t)v(t) \\ &\leq T(t-1) + \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)v^2(t) + 2\left[1 - \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)\right]\tilde{\mathbf{y}}(t)v(t). \end{aligned} \quad (3.3)$$

Here, we have used the inequality $1 - \boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t) = [1 + \boldsymbol{\varphi}^T(t)\mathbf{P}(t-1)\boldsymbol{\varphi}(t)]^{-1} \geq 0$. Because $\tilde{\mathbf{y}}(t)$ and $\boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)$ are uncorrelated with $v(t)$ and are \mathcal{F}_{t-1} measurable, taking the conditional expectation with respect to \mathcal{F}_{t-1} , we have

$$E[T(t) \mid F_{t-1}] \leq T(t-1) + 2\boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)\sigma^2. \quad (3.4)$$

Since $\ln|\mathbf{P}^{-1}(t)|$ is nondecreasing, letting

$$V(t) := \frac{T(t)}{[\ln|\mathbf{P}^{-1}(t)|]^\gamma}, \quad \gamma > 1, \quad (3.5)$$

we have

$$\begin{aligned} E[V(t) \mid F_{t-1}] &\leq \frac{T(t-1)}{[\ln|\mathbf{P}^{-1}(t)|]^\gamma} + \frac{2\boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)}{[\ln|\mathbf{P}^{-1}(t)|]^\gamma}\sigma^2 \\ &\leq V(t-1) + \frac{2\boldsymbol{\varphi}^T(t)\mathbf{P}(t)\boldsymbol{\varphi}(t)}{[\ln|\mathbf{P}^{-1}(t)|]^\gamma}\sigma^2, \quad \text{a.s.} \end{aligned} \quad (3.6)$$

Using Lemma 3.2, the sum of the last term in the right-hand side for t from 1 to ∞ is finite. Applying Lemma 3.1 to the previous inequality, we conclude that $V(t)$ converges a.s. to a finite random variable, say V_0 , that is:

$$V(t) = \frac{T(t)}{[\ln|\mathbf{P}^{-1}(t)|]^\gamma} \rightarrow V_0 < \infty, \quad \text{a.s.,} \quad \text{or} \quad T(t) = O\left([\ln|\mathbf{P}^{-1}(t)|]^\gamma\right), \quad \text{a.s.} \quad (3.7)$$

Thus, according to the definition of $T(t)$, we have

$$\|\tilde{\boldsymbol{\theta}}(t)\|^2 \leq \frac{\text{tr}\left[\tilde{\boldsymbol{\theta}}^T(t)\mathbf{P}^{-1}(t)\tilde{\boldsymbol{\theta}}(t)\right]}{\lambda_{\min}[\mathbf{P}^{-1}(t)]} = O\left(\frac{[\ln|\mathbf{P}^{-1}(t)|]^\gamma}{\lambda_{\min}[\mathbf{P}^{-1}(t)]}\right) = O\left(\frac{o(\lambda_{\min}[\mathbf{P}^{-1}(t)])}{\lambda_{\min}[\mathbf{P}^{-1}(t)]}\right) \rightarrow 0, \quad \text{a.s.} \quad (3.8)$$

This completes the proof of Theorem 3.3. \square

According to the definition of θ and the assumption $b_1 = 1$, the estimates $\hat{\mathbf{a}}(t) = [\hat{a}_1(t), \hat{a}_2(t), \dots, \hat{a}_n(t)]^T$ and $\hat{\mathbf{c}}(t) = [\hat{c}_1(t), \hat{c}_2(t), \dots, \hat{c}_m(t)]^T$ of \mathbf{a} and \mathbf{c} can be read from the first n and second m entries of $\hat{\theta}$, respectively. Let $\hat{\theta}_i$ be the i th element of $\hat{\theta}$. Referring to the definition of θ , the estimates $\hat{b}_j(t)$ of b_j , $j = 2, 3, \dots, n$, may be computed by

$$\hat{b}_j(t) = \frac{\hat{\theta}_{n+(j-1)m+i}(t)}{\hat{c}_i(t)}, \quad j = 2, 3, \dots, n; \quad i = 1, 2, \dots, m. \quad (3.9)$$

Notice that there is a large amount of redundancy about $\hat{b}_j(t)$ for each $i = 1, 2, \dots, m$. Since we do not need such m estimates $\hat{b}_j(t)$, one way is to take their average as the estimate of b_j [14], that is:

$$\hat{b}_j(t) = \frac{1}{m} \sum_{i=1}^m \frac{\hat{\theta}_{n+(j-1)m+i}(t)}{\hat{c}_i(t)}, \quad j = 2, 3, \dots, n. \quad (3.10)$$

4. The Input Nonlinear CARARMA System and Estimation Algorithm

Consider the following input nonlinear controlled autoregressive autoregressive moving average (IN-CARARMA) systems:

$$A(z)y(t) = B(z)\bar{u}(t) + \frac{D(z)}{\gamma(z)}v(t), \quad (4.1)$$

$$\bar{u}(t) = f(u(t)) = c_1 f_1(u(t)) + c_2 f_2(u(t)) + \dots + c_m f_m(u(t)),$$

$$\gamma(z) := 1 + \gamma_1 z^{-1} + \gamma_2 z^{-2} + \dots + \gamma_{n_\gamma} z^{-n_\gamma}, \quad (4.2)$$

$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \dots + d_{n_d} z^{-n_d}.$$

Let

$$w(t) := \frac{D(z)}{\gamma(z)}v(t), \quad (4.3)$$

or

$$\begin{aligned} w(t) &= [1 - \gamma(z)]w(t) + D(z)v(t) \\ &= - \sum_{i=1}^{n_\gamma} \gamma_i w(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t). \end{aligned} \quad (4.4)$$

Define the parameter vector θ and information vector $\varphi(t)$ as

$$\begin{aligned} \theta &:= [\theta_1^T, \gamma_1, \gamma_2, \dots, \gamma_{n_\gamma}, d_1, d_2, \dots, d_{n_d}]^T \in \mathbb{R}^{n+mn+n_\gamma+n_d}, \\ \varphi(t) &:= [\varphi_1^T(t), -w(t-1), -w(t-2), \dots, -w(t-n_\gamma), v(t-1), v(t-2), \dots, v(t-n_d)]^T \\ &\in \mathbb{R}^{n+mn+n_\gamma+n_d}, \\ \theta_1 &:= \begin{bmatrix} \mathbf{a} \\ b_1 \mathbf{c} \\ b_2 \mathbf{c} \\ \vdots \\ b_n \mathbf{c} \end{bmatrix} \in \mathbb{R}^{n+nm}, \quad \varphi_1(t) := \begin{bmatrix} \varphi_0(t) \\ \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{bmatrix} \in \mathbb{R}^{n+nm}, \\ \mathbf{a} &:= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n, \quad \mathbf{c} := \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \in \mathbb{R}^m, \\ \varphi_0(t) &:= \begin{bmatrix} -y(t-1) \\ -y(t-2) \\ \vdots \\ -y(t-n) \end{bmatrix} \in \mathbb{R}^n, \quad \varphi_j(t) := \begin{bmatrix} f_1(u(t-j)) \\ f_2(u(t-j)) \\ \vdots \\ f_m(u(t-j)) \end{bmatrix} \in \mathbb{R}^m, \quad j = 1, 2, \dots, n. \end{aligned} \tag{4.5}$$

Then (4.1) can be written as

$$\begin{aligned} \mathbf{y}(t) &= [1 - A(z)]\mathbf{y}(t) + B(z)\bar{\mathbf{u}}(t) + \mathbf{w}(t) \\ &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + \sum_{i=1}^n b_i \sum_{j=1}^m c_j f_j(u(t-i)) + \mathbf{w}(t) \\ &= -\sum_{i=1}^n a_i \mathbf{y}(t-i) + \sum_{i=1}^n \sum_{j=1}^m b_i c_j f_j(u(t-i)) + \mathbf{w}(t) \\ &= \varphi_1^T(t) \theta_1 + \mathbf{w}(t) \\ &= \varphi_1^T(t) \theta_1 - \sum_{i=1}^{n_\gamma} \gamma_i \mathbf{w}(t-i) + \sum_{i=1}^{n_d} d_i v(t-i) + v(t) \\ &= \varphi^T(t) \theta + v(t). \end{aligned} \tag{4.6}$$

This is a linear-in-parameter identification model for IN-CARARMA systems.

The unknown $w(t-i)$ and $v(t-i)$ in the information vector $\boldsymbol{\varphi}(t)$ are replaced with their estimates $\hat{w}(t-i)$ and $\hat{v}(t-i)$, and then we can obtain the following recursive generalized extended least squares algorithm for estimating $\boldsymbol{\theta}$ in (4.6):

$$\begin{aligned}
\hat{\boldsymbol{\theta}}(t) &= \hat{\boldsymbol{\theta}}(t-1) + \mathbf{L}(t) \left[y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t-1) \right], \\
\mathbf{L}(t) &= \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) \left[1 + \hat{\boldsymbol{\varphi}}^T(t) \mathbf{P}(t-1) \hat{\boldsymbol{\varphi}}(t) \right]^{-1}, \\
\mathbf{P}(t) &= \left[\mathbf{I} - \mathbf{L}(t) \hat{\boldsymbol{\varphi}}^T(t) \right] \mathbf{P}(t-1), \quad \mathbf{P}(0) = p_0 \mathbf{I}, \\
\hat{\boldsymbol{\varphi}}(t) &= [\boldsymbol{\varphi}_1^T(t), -\hat{w}(t-1), -\hat{w}(t-2), \dots, -\hat{w}(t-n_y), \hat{v}(t-1), \hat{v}(t-2), \dots, \hat{v}(t-n_d)]^T, \\
\boldsymbol{\varphi}_1(t) &= \begin{bmatrix} \boldsymbol{\varphi}_0(t) \\ \boldsymbol{\varphi}_1(t) \\ \boldsymbol{\varphi}_2(t) \\ \vdots \\ \boldsymbol{\varphi}_n(t) \end{bmatrix}, \quad \boldsymbol{\varphi}_0(t) = \begin{bmatrix} -y(t-1) \\ -y(t-2) \\ \vdots \\ -y(t-n) \end{bmatrix}, \quad \boldsymbol{\varphi}_j(t) = \begin{bmatrix} f_1(u(t-j)) \\ f_2(u(t-j)) \\ \vdots \\ f_m(u(t-j)) \end{bmatrix}, \\
\hat{w}(t) &= y(t) - \hat{\boldsymbol{\varphi}}_1^T(t) \hat{\boldsymbol{\theta}}_1(t), \\
\hat{v}(t) &= y(t) - \hat{\boldsymbol{\varphi}}^T(t) \hat{\boldsymbol{\theta}}(t), \\
\hat{\boldsymbol{\theta}}(t) &= [\hat{\boldsymbol{\theta}}_1^T(t), \hat{\gamma}_1(t), \hat{\gamma}_2(t), \dots, \hat{\gamma}_{n_y}(t), \hat{d}_1(t), \hat{d}_2(t), \hat{d}_{n_d}(t)]^T.
\end{aligned} \tag{4.7}$$

This paper presents a recursive least squares algorithm for IN-CAR systems and a recursive generalized extended least squares algorithm for IN-CARARMA systems with ARMA noise disturbances, which differ not only from the input nonlinear controlled autoregressive moving average (IN-CARMA) systems in [14] but also from the input nonlinear output error systems in [15].

5. Example

Consider the following IN-CAR system:

$$\begin{aligned}
A(z)y(t) &= B(z)\bar{u}(t) + v(t), \\
A(z) &= 1 + a_1z^{-1} + a_2z^{-2} = 1 - 1.35z^{-1} + 0.75z^{-2}, \\
B(z) &= b_1z^{-1} + b_2z^{-2} = z^{-1} + 1.68z^{-2}, \\
\bar{u}(t) &= f(u(t)) = c_1u(t) + c_2u^2(t) + c_3u^3(t) \\
&= u(t) + 0.50u^2(t) + 0.20u^3(t), \\
\boldsymbol{\theta} &= [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8]^T \\
&= [a_1, a_2, c_1, c_2, c_3, b_2c_1, b_2c_2, b_2c_3]^T \\
&= [-1.350, 0.75, 1.00, 0.50, 0.20, 1.68, 0.84, 0.336]^T, \\
\boldsymbol{\theta}_s &= [a_1, a_2, b_2, c_1, c_2, c_3]^T = [-1.35, 0.75, 1.68, 1.00, 0.50, 0.20]^T.
\end{aligned} \tag{5.1}$$

Table 1: The parameter estimates (θ) ($\sigma^2 = 0.50^2$, $\delta_{ns} = 10.96\%$).

t	a_1	a_2	c_1	c_2	c_3	b_2c_1	b_2c_2	b_2c_3	δ (%)
100	-1.35989	0.76938	0.94139	0.49861	0.18862	1.69875	0.86164	0.32773	2.59527
200	-1.35622	0.76001	0.96720	0.50101	0.19076	1.67233	0.84941	0.34369	1.43552
500	-1.35239	0.75452	1.00256	0.50137	0.19363	1.66468	0.84394	0.34485	0.74281
1000	-1.35034	0.75193	1.00570	0.50128	0.19460	1.65482	0.85095	0.33765	1.06112
2000	-1.34844	0.74940	0.99224	0.50089	0.20143	1.69169	0.85148	0.33583	0.67584
3000	-1.34776	0.74847	0.99012	0.49943	0.20333	1.68675	0.85321	0.33416	0.68173
True values	-1.35000	0.75000	1.00000	0.50000	0.20000	1.68000	0.84000	0.33600	

Table 2: The parameter estimates (θ_s) ($\sigma^2 = 0.50^2$, $\delta_{ns} = 10.96\%$).

t	a_1	a_2	b_2	c_1	c_2	c_3	δ (%)
100	-1.35989	0.76938	1.75670	0.94139	0.49861	0.18862	3.90775
200	-1.35622	0.76001	1.74205	0.96720	0.50101	0.19076	2.81582
500	-1.35239	0.75452	1.70821	1.00256	0.50137	0.19363	1.15783
1000	-1.35034	0.75193	1.69268	1.00570	0.50128	0.19460	0.59233
2000	-1.34844	0.74940	1.69068	0.99224	0.50089	0.20143	0.52605
3000	-1.34776	0.74847	1.68512	0.99012	0.49943	0.20333	0.46851
True values	-1.35000	0.75000	1.68000	1.00000	0.50000	0.20000	

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 constant variance σ^2 . Applying the proposed algorithm in (2.10)-(2.11) to estimate the parameters of this system, the parameter estimates θ and θ_s and their errors with different noise variances are shown in Tables 1, 2, 3, and 4, and the parameter estimation errors $\delta := \|\hat{\theta}(t) - \theta\| / \|\theta\|$ and $\delta_s := \|\hat{\theta}_s(t) - \theta_s\| / \|\theta_s\|$ versus t are shown in Figures 1 and 2. When $\sigma^2 = 0.50^2$ and $\sigma^2 = 1.50^2$, the corresponding noise-to-signal ratios are $\delta_{ns} = 10.96\%$ and $\delta_{ns} = 32.87\%$, respectively.

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

- (i) The larger the data length is, the smaller the parameter estimation errors become.
- (ii) A lower noise level leads to smaller parameter estimation errors for the same data length.
- (iii) The estimation errors δ and δ_s become smaller (in general) as t increases. This confirms the proposed theorem.

6. Conclusions

The recursive least-squares identification is used to estimate the unknown parameters for input nonlinear CAR and CARARMA systems. The analysis using the martingale convergence theorem indicates that the proposed recursive least squares algorithm can give consistent parameter estimation. It is worth pointing out that the multi-innovation identification theory [26–33], the gradient-based or least-squares-based identification methods [34–41], and other identification methods [42–49] can be used to study identification problem of this class of nonlinear systems with colored noises.

Table 3: The parameter estimates (θ) ($\sigma^2 = 1.50^2$, $\delta_{ns} = 32.87\%$).

t	a_1	a_2	c_1	c_2	c_3	b_2c_1	b_2c_2	b_2c_3	δ (%)
100	-1.37143	0.79561	0.81731	0.49688	0.16665	1.75280	0.90056	0.30999	7.98804
200	-1.36256	0.77335	0.89403	0.50353	0.17365	1.65999	0.87041	0.36082	4.46419
500	-1.35374	0.76009	1.00710	0.50417	0.18108	1.63863	0.85315	0.36321	2.08028
1000	-1.35074	0.75537	1.01710	0.50372	0.18381	1.60488	0.87297	0.34101	3.17034
2000	-1.34587	0.74895	0.97678	0.50296	0.20424	1.71432	0.87455	0.33540	2.01031
3000	-1.34448	0.74649	0.97053	0.49834	0.20990	1.69896	0.87916	0.33025	2.00404
True values	-1.35000	0.75000	1.00000	0.50000	0.20000	1.68000	0.84000	0.33600	

Table 4: The parameter estimates (θ_s) ($\sigma^2 = 1.50^2$, $\delta_{ns} = 32.87\%$).

t	a_1	a_2	b_2	c_1	c_2	c_3	δ (%)
100	-1.37143	0.79561	1.93906	0.81731	0.49688	0.16665	12.66088
200	-1.36256	0.77335	1.88773	0.89403	0.50353	0.17365	9.26624
500	-1.35374	0.76009	1.77500	1.00710	0.50417	0.18108	3.83703
1000	-1.35074	0.75537	1.72205	1.01710	0.50372	0.18381	1.90836
2000	-1.34587	0.74895	1.71204	0.97678	0.50296	0.20424	1.57452
3000	-1.34448	0.74649	1.69603	0.97053	0.49834	0.20990	1.39744
True values	-1.35000	0.75000	1.68000	1.00000	0.50000	0.20000	

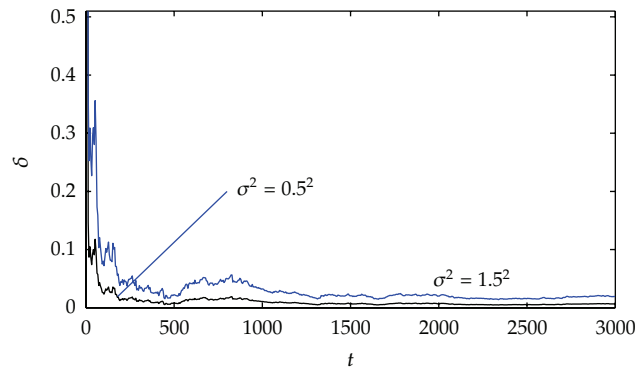


Figure 1: The parameter estimation errors δ versus t .

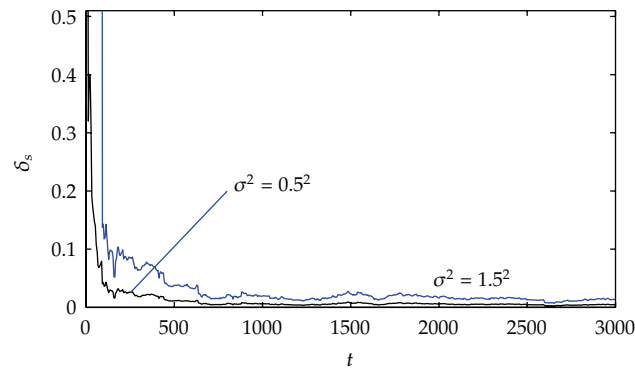


Figure 2: The parameter estimation errors δ_s versus t .

Acknowledgment

This work was supported by the 111 Project (B12018).

References

- [1] M. R. Zakerzadeh, M. Firouzi, H. Sayyaadi, and S. B. Shouraki, "Hysteresis nonlinearity identification using new Preisach model-based artificial neural network approach," *Journal of Applied Mathematics*, Article ID 458768, 22 pages, 2011.
- [2] X.-X. Li, H. Z. Guo, S. M. Wan, and F. Yang, "Inverse source identification by the modified regularization method on poisson equation," *Journal of Applied Mathematics*, vol. 2012, Article ID 971952, 13 pages, 2012.
- [3] 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.
- [4] 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.
- [5] F. Ding, G. Liu, and X. P. Liu, "Parameter estimation with scarce measurements," *Automatica*, vol. 47, no. 8, pp. 1646–1655, 2011.
- [6] 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.
- [7] 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.
- [8] J. Ding and F. Ding, "The residual based extended least squares identification method for dual-rate systems," *Computers & Mathematics with Applications*, vol. 56, no. 6, pp. 1479–1487, 2008.
- [9] 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.
- [10] 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.
- [11] F. Ding, T. Chen, and Z. Iwai, "Adaptive digital control of Hammerstein nonlinear systems with limited output sampling," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2257–2276, 2007.
- [12] J. Li and F. Ding, "Maximum likelihood stochastic gradient estimation for Hammerstein systems with colored noise based on the key term separation technique," *Computers & Mathematics with Applications*, vol. 62, no. 11, pp. 4170–4177, 2011.
- [13] 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.
- [14] F. Ding and T. Chen, "Identification of Hammerstein nonlinear ARMAX systems," *Automatica*, vol. 41, no. 9, pp. 1479–1489, 2005.
- [15] 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.
- [16] 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.
- [17] D. Wang, Y. Chu, G. Yang, and F. Ding, "Auxiliary model based recursive generalized least squares parameter estimation for Hammerstein OEAR systems," *Mathematical and Computer Modelling*, vol. 52, no. 1-2, pp. 309–317, 2010.
- [18] D. Wang, Y. Chu, and F. Ding, "Auxiliary model-based RELS and MI-ELS algorithm for Hammerstein OEMA systems," *Computers & Mathematics with Applications*, vol. 59, no. 9, pp. 3092–3098, 2010.
- [19] 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.
- [20] D. 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.

- [21] W. Fan, F. Ding, and Y. Shi, "Parameter estimation for Hammerstein nonlinear controlled auto-regression models," in *Proceedings of the IEEE International Conference on Automation and Logistics*, pp. 1007–1012, Jinan, China, August 2007.
- [22] L. Wang, F. Ding, and P. X. Liu, "Convergence of HLS estimation algorithms for multivariable ARX-like systems," *Applied Mathematics and Computation*, vol. 190, no. 2, pp. 1081–1093, 2007.
- [23] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1984.
- [24] 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.
- [25] F. Ding and T. Chen, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, p. 17391748S, 2004.
- [26] F. Ding and T. Chen, "Performance analysis of multi-innovation gradient type identification methods," *Automatica*, vol. 43, no. 1, pp. 1–14, 2007.
- [27] 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.
- [28] F. Ding, "Several multi-innovation identification methods," *Digital Signal Processing*, vol. 20, no. 4, pp. 1027–1039, 2010.
- [29] D. 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.
- [30] 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.
- [31] 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.
- [32] 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.
- [33] F. Ding, P. X. Liu, and G. Liu, "Multiinnovation least-squares identification for system modeling," *IEEE Transactions on Systems, Man, and Cybernetics B*, vol. 40, no. 3, Article ID 5299173, pp. 767–778, 2010.
- [34] J. Ding, Y. Shi, H. Wang, and F. Ding, "A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems," *Digital Signal Processing*, vol. 20, no. 4, pp. 1238–1247, 2010.
- [35] F. Ding, P. X. Liu, and H. Yang, "Parameter identification and intersample output estimation for dual-rate systems," *IEEE Transactions on Systems, Man, and Cybernetics A*, vol. 38, no. 4, pp. 966–975, 2008.
- [36] Y. Liu, D. Wang, and F. Ding, "Least squares based iterative algorithms for identifying Box-Jenkins models with finite measurement data," *Digital Signal Processing*, vol. 20, no. 5, pp. 1458–1467, 2010.
- [37] D. Wang and F. Ding, "Input-output data filtering based recursive least squares identification for CARARMA systems," *Digital Signal Processing*, vol. 20, no. 4, pp. 991–999, 2010.
- [38] F. Ding, P. X. Liu, and G. Liu, "Gradient based and least-squares based iterative identification methods for OE and OEMA systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 664–677, 2010.
- [39] D. Wang, G. Yang, and R. Ding, "Gradient-based iterative parameter estimation for Box-Jenkins systems," *Computers & Mathematics with Applications*, vol. 60, no. 5, pp. 1200–1208, 2010.
- [40] L. Xie, H. Yang, and F. Ding, "Recursive least squares parameter estimation for non-uniformly sampled systems based on the data filtering," *Mathematical and Computer Modelling*, vol. 54, no. 1-2, pp. 315–324, 2011.
- [41] F. Ding, Y. Liu, and B. Bao, "Gradient-based and least-squares-based iterative estimation algorithms for multi-input 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.
- [42] F. Ding, "Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling," *Applied Mathematical Modelling*. In press.
- [43] F. Ding and J. Ding, "Least-squares parameter estimation for systems with irregularly missing data," *International Journal of Adaptive Control and Signal Processing*, vol. 24, no. 7, pp. 540–553, 2010.
- [44] Y. Liu, L. Xie, and F. Ding, "An auxiliary model based on a recursive least-squares parameter estimation algorithm for non-uniformly sampled multirate systems," *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, vol. 223, no. 4, pp. 445–454, 2009.
- [45] 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.

- [46] 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.
- [47] 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.
- [48] F. Ding and T. Chen, "Performance bounds of forgetting factor least-squares algorithms for time-varying systems with finite measurement data," *IEEE Transactions on Circuits and Systems. I. Regular Papers*, vol. 52, no. 3, pp. 555–566, 2005.
- [49] 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.