Bias Compensation Based Recursive Least-Squares Identification Algorithm for MISO Systems

Feng Ding, Tongwen Chen, and Li Qiu

***Abstract*—For multi-input single-output output-error systems, the least-squares (LS) estimates are biased. In order to obtain the unbiased estimates, we present a recursive LS identification algorithm based on a bias compensation technique. The basic idea is to eliminate the estimation bias by adding a correction term in the LS estimates, and further to derive a bias compensation based recursive LS algorithm. Finally, we test the proposed algorithms by simulation and show their effectiveness.**

***Index Terms*—Bias compensation, bias correction, least squares (LS), multivariable systems, parameter estimation, recursive identification.**

I. Problem Formulation

Since a multi-input multi-output system may be decomposed into several multi-input single-output (MISO) subsystems, we here consider a MISO system described by the *output-error state-space model* [1]

$$\begin{cases} x(t+1) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t) + v(t). \end{cases} \tag{1}$$

Here, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) = [u_1(t), u_2(t), \ldots, u_r(t)]^{\mathrm{T}} \in \mathbb{R}^r$ the system input vector (the superscript T denoting the matrix transpose), $y(t) \in \mathbb{R}^1$ the system output, $v(t) \in \mathbb{R}^1$ the observation white noise with zero mean, and $A \in \mathbb{R}^{n\times n}, B = [b_1, b_2, \ldots, b_r] \in \mathbb{R}^{n\times r}, C \in \mathbb{R}^{1\times n}$ and $D = [d_1, d_2, \ldots, d_r] \in \mathbb{R}^{1\times r}$ the system matrices.

Taking the z transforms in (1) gives

$$y(t) = [C(zI-A)^{-1}B + D]u(t) + v(t) \tag{2}$$

$$= \left[\frac{C\mathrm{adj}[I-Az^{-1}]B}{z^{-n}\det[zI-A]} + D\right]u(t) + v(t) \tag{3}$$

$$=: \frac{1}{\alpha(z)}\sum_{i=1}^{r} B_i(z)u_i(t) + v(t) \tag{4}$$

with $\alpha(z)$ being the characteristic polynomial in the unit delay operator $z^{-1}[z^{-1}y(t) = y(t-1)]$ of degree n and $B_i(z)$ a polynomial in z^{-1}, and both represented as

$$\alpha(z) = z^{-n}\det[zI-A] \tag{5}$$

$$=: 1+\alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_n z^{-n}, \ \alpha_i \in \mathbb{R}^1 \tag{6}$$

$$B_i(z) = C\mathrm{adj}[I-Az^{-1}]b_i + d_i z^{-n}\det[zI-A] \tag{7}$$

$$=: \beta_{i0} + \beta_{i1}z^{-1} + \cdots + \beta_{\mathrm{in}}z^{-n}, \quad \beta_{ij} \in \mathbb{R}^1. \tag{8}$$

The output-error model in (2) cannot be identified by standard least-squares (LS) algorithms since it differs from the multi-input autoregressive with exogenous input (ARX) model [2]

$$\alpha(z)y(t) = \sum_{i=1}^{r} B_i(z)u_i(t) + v(t).$$

The bias correction or bias elimination identification method is an effective way of obtaining unbiased parameter estimates of stochastic systems. It has been used to study the identification problem of various system models, e.g., output-error systems [3]–[5], ARX models with correlated noise [6], autoregressive moving average (ARMA) models [7], MISO systems [8], autoregressive models [9], errors-in-variables models [10]–[12], feedback or closed-loop systems [13]–[17]. However, most correlation analysis based contributions mentioned above require the assumption that the system input is ergodic, which is very difficult to satisfy in practice, and few address *recursive* identification methods based on the bias correction technique, which is the focus of this work. This paper uses the bias compensation or bias correction technique to study the recursive identification problem of MISO systems in (2). The basic idea is to use a correction term to compensate the biased LS estimates, and then to derive a bias compensation based recursive LS (BCRLS) algorithm to estimate the unknown parameters (α_i, β_{ij}) in (2) from the given input–output measurement data $\{u_i(t), y(t) : t = 1, 2, \ldots\}$, and further, to study the numerical convergence of the algorithm presented by simulation. The approach here differs not only from the ones mentioned above because we do not assume that the system input is ergodic, but also from the ones in [18]–[20] which used the auxiliary model technique to identify and estimate the parameters and missing outputs of dual-rate sampled-data systems. The proposed approach is also different from the method in [21] which used a hierarchical identification principle to study the identification problem for multi-input, multi-output systems.

Briefly, the paper is organized as follows. Section II derives a basic identification algorithm for MISO systems based on a bias compensation technique. Section III presents an illustrative example for the results in this paper. Finally, concluding remarks are given in Section IV.

Manuscript received January 31, 2005; revised July 28, 2005. This work was supported by the Natural Sciences and Engineering Research Council of Canada, the Hong Kong Research Grants Council (HKUST6171/02E), and the National Natural Science Foundation of China (60474039, 60574051, 60528007). This paper was recommended by Associate Editor F. C. M. Lau.

F. Ding is with the Control Science and Engineering Research Center, Southern Yangtze University, Wuxi 214122, China (e-mail: fding@sytu.edu.cn; fding@ece.ualberta.ca).

T. Chen is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: tchen@ece.ualberta.ca).

L. Qiu is with the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: eeqiu@ust.hk).

Digital Object Identifier 10.1109/TCSII.2005.862281

II. Algorithm Derivation

Let

$$e(t) := \alpha(z)v(t). \tag{9}$$

Equation (2) can be written as

$$\alpha(z)y(t) = \sum_{i=1}^{r} B_i(z)u_i(t) + e(t). \tag{10}$$

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

$$\begin{aligned}
\theta &= [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n \quad \beta_{10} \quad \beta_{11} \quad \cdots \quad \beta_{1n} \quad \cdots \\
&\quad \beta_{r0} \quad \beta_{r1} \quad \cdots \quad \beta_{rn}]^{\mathrm{T}} \in \mathbb{R}^{n_0}, \\
\varphi(t) &= [-y(t-1) \quad -y(t-2) \quad \cdots \quad -y(t-n) \\
&\quad u_1(t) \quad u_1(t-1) \quad \cdots \quad u_1(t-n) \\
&\quad \cdots \quad u_r(t) \quad u_r(t-1) \quad \cdots \quad u_r(t-n)]^{\mathrm{T}} \in \mathbb{R}^{n_0} \\
\psi(t) &= [v(t-1) \quad \cdots \quad v(t-n) \quad 0 \quad \cdots \quad 0]^{\mathrm{T}} \in \mathbb{R}^{n_0} \\
n_0 &:= n + r(n+1).
\end{aligned}$$

From (9) and (10), we have

$$e(t) = \psi^{\mathrm{T}}(t)\theta + v(t) \tag{11}$$

$$y(t) = \varphi^{\mathrm{T}}(t)\theta + e(t) \tag{12}$$

$$y(t) = \varphi^{\mathrm{T}}(t)\theta + \psi^{\mathrm{T}}(t)\theta + v(t). \tag{13}$$

Further, let

$$Y(t) = [y(1) \quad y(2) \quad \cdots \quad y(t)]^{\mathrm{T}} \in \mathbb{R}^{t}, \tag{14}$$

$$\Phi(t) = [\varphi(1) \quad \varphi(2) \quad \cdots \quad \varphi(t)]^{\mathrm{T}} \in \mathbb{R}^{t \times n_0} \tag{15}$$

$$E(t) = [e(1) \quad e(2) \quad \cdots \quad e(t)]^{\mathrm{T}} \in \mathbb{R}^{t}. \tag{16}$$

It is easy to get

$$Y(t) = \Phi(t)\theta + E(t). \tag{17}$$

Form a cost function [2]

$$J(\theta) = \|Y(t) - \Phi(t)\theta\|^2$$

where $\|X\|^2 := \mathrm{tr}[XX^{\mathrm{T}}]$. According to the LS principle, we can obtain the LS estimate of θ as follows:

$$\hat{\theta}_{\mathrm{LS}}(t) = [\Phi^{\mathrm{T}}(t)\Phi(t)]^{-1}\Phi^{\mathrm{T}}(t)Y(t).$$

Because $e(t)(E(t))$ is a correlated noise (vector), this LS estimate $\hat{\theta}_{\mathrm{LS}}$ is a biased one of the parameter vector θ. In fact, using (17), we

$$\hat{\theta}_{\mathrm{LS}}(t) = [\Phi^{\mathrm{T}}(t)\Phi(t)]^{-1}\Phi^{\mathrm{T}}(t)[\Phi(t)\theta + E(t)] \tag{18}$$

$$= \theta + [\Phi^{\mathrm{T}}(t)\Phi(t)]^{-1}\Phi^{\mathrm{T}}(t)E(t) \tag{19}$$

$$= \theta + \left[\sum_{i=1}^{t} \varphi(i)\varphi^{\mathrm{T}}(i)\right]^{-1} \left[\sum_{i=1}^{t} \varphi(i)e(i)\right]. \tag{20}$$

Using (11), it follows that

$$\begin{aligned}
\left[\sum_{i=1}^{t} \varphi(i)\varphi^{\mathrm{T}}(i)\right](\hat{\theta}_{\mathrm{LS}}(t) - \theta) &= \sum_{i=1}^{t} \varphi(i)e(i) \\
&= \sum_{i=1}^{t} \varphi(i)[\psi^{\mathrm{T}}(i)\theta + v(i)].
\end{aligned} \tag{21}$$

Dividing by t and taking limit yield

$$\begin{aligned}
&\lim_{t\to\infty} \frac{1}{t}\left\{\left[\sum_{i=1}^{t} \varphi(i)\varphi^{\mathrm{T}}(i)\right](\hat{\theta}_{\mathrm{LS}}(t) - \theta)\right\} \\
&= \lim_{t\to\infty}\left[\frac{1}{t}\sum_{i=1}^{t} \varphi(i)\psi^{\mathrm{T}}(i)\right]\theta + \lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}\varphi(i)v(i).
\end{aligned} \tag{22}$$

Since $v(t)$ is a white noise with zero mean and variance σ^2 and is independent of the inputs, i.e., $\mathrm{E}[v(t)] = 0, \mathrm{E}[v(t)v(t+j)] = 0, j \neq 0, \mathrm{E}[v^2(t)] = \sigma^2$, and $\mathrm{E}[v(t)u_i(j)] = 0$, the second term on the right-hand side of the above equation converges to zero, and the first term converges to $[-\sigma^2\Lambda\theta]$, which means

$$\lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t}\varphi(i)\psi^{\mathrm{T}}(i)\theta = -\sigma^2\Lambda\theta \tag{23}$$

where

$$\Lambda := \begin{bmatrix} I_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0}.$$

Define the data product moment matrix

$$R(t) := \Phi^{\mathrm{T}}(t)\Phi(t).$$

Provided that the inputs are persistent excitation signals, for large t, the following persistent excitation (PE) condition holds:

$$\frac{R(t)}{t} > 0.$$

This includes the generalized PE condition [19], the weak PE condition [19] and the strong PE condition [22]. Hence, from (22), we have

$$\lim_{t\to\infty} \hat{\theta}_{\mathrm{LS}}(t) = \theta - \sigma^2 \lim_{t\to\infty} tR^{-1}(t)\Lambda\theta. \tag{24}$$

The following is to indicate that the results in [4] require the stationarity and ergodicity assumptions. Define the correlation function $\mathbf{R}_{\varphi}(t)$ of $\boldsymbol{\varphi}(t)$ as follows:

$$\mathbf{R}_{\varphi}(t) = \mathrm{E}[\varphi(t)\varphi^{\mathrm{T}}(t)].$$

If the inputs are stationary and ergodic, i.e., $\varphi(t)$ is 2nd-moment-ergodic and $\lim_{t\to\infty}(R(t)/t)$ exists, then $\mathbf{R}_{\varphi}(t)$ does not depend on t (denoted by $\mathbf{R}_{\varphi}$) and $\mathbf{R}_{\varphi} = \lim_{t\to\infty} R(t)/t$ according to the definition of ergodicity. Under such assumptions, (24) can be written as

$$\lim_{t\to\infty} \hat{\theta}_{\mathrm{LS}}(t) = \theta - \sigma^2\mathbf{R}_{\varphi}^{-1}\Lambda\theta. \tag{25}$$

Equation (25) is the basic equation of the bias compensation methods for stationary cases, see, e.g., [4], and obviously requires the assumption that the sample average $R(t)/t$ has limit. For nonstationary data, $R(t)/t$ is time-varying and has no limit; thus $\mathbf{R}_{\varphi}$ does not exist—see the example later.

If the noise variance σ^2 and correlation function $\mathbf{R}_{\varphi}$ are known or obtained by estimation, then from (25), an alternate way to get the unbiased estimate $\bar{\theta}(t)$ of θ can be simply expressed as

$$\bar{\theta}(t) = \left[I - \sigma^2\mathbf{R}_{\varphi}^{-1}\Lambda\right]^{-1}\hat{\theta}_{\mathrm{LS}}(t).$$

Equation (24) shows that the LS estimate $\hat{\theta}_{\mathrm{LS}}(t)$ is biased, and is a basic equation for bias compensation menthods, without assuming stationarity and ergodicity of input data. If we introduce a compensation term $\sigma^2 tR^{-1}(t)\Lambda\theta$ in the LS estimate $\hat{\theta}_{\rm LS}(t)$, then we can obtain the unbiased estimate $\hat{\theta}_{\rm LS}(t)+\sigma^2 tR^{-1}(t)\Lambda\theta =: \hat{\theta}_{\rm c}(t)$ of θ, i.e., $\hat{\theta}_{\rm c}(t)\to\theta$. This is the basic idea of the bias compensation LS method. Define a covariance matrix $P(t) := [\Phi^{\rm T}(t)\Phi(t)]^{-1}$. Let $\hat{\sigma}^2(t)$ be the estimate of the noise variance σ^2. We can write $\hat{\theta}_{\rm c}$ in a recursive form

$$\hat{\theta}_{\rm c}(t)=\hat{\theta}_{\rm LS}(t)+\hat{\sigma}^2(t)tP(t)\hat{\theta}_{\rm c}(t-1) \tag{26}$$

where $\hat{\theta}_{\rm c}(t)$ and $\hat{\theta}_{\rm LS}(t)$ are the bias compensation LS estimate and LS estimate of θ at time t, respectively.

Now, the problem is changed into how to compute the variance estimate $\hat{\sigma}^2(t)$. The details are as follows. Let

$$\varepsilon_{\rm LS}(t) := y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}. \tag{27}$$

Using (11)–(13) and the relation

$$\sum_{i=1}^{t}\varepsilon_{\rm LS}(i)\varphi^{\rm T}(i)=\mathbf{0} \tag{28}$$

it is not difficult to get

$$\begin{aligned}\sum_{i=1}^{t}\varepsilon_{\rm LS}^2(i)&=\sum_{i=1}^{t}\varepsilon_{\rm LS}(i)\left[y(i)-\varphi^{\rm T}(i)\hat{\theta}_{\rm LS}\right] &(29)\\ &=\sum_{i=1}^{t}\varepsilon_{\rm LS}(i)y(i)\\ &=\sum_{i=1}^{t}\varepsilon_{\rm LS}(i)[\varphi^{\rm T}(i)\theta+\psi^{\rm T}(i)\theta+v(i)] &(30)\\ &=\sum_{i=1}^{t}\varepsilon_{\rm LS}(i)[\psi^{\rm T}(i)\theta+v(i)] &(31)\\ &=\sum_{i=1}^{t}\left[y(i)-\varphi^{\rm T}(i)\hat{\theta}_{\rm LS}][\psi^{\rm T}(i)\theta+v(i)\right] &(32)\\ &=\sum_{i=1}^{t}\left[\varphi^{\rm T}(i)\theta+\psi^{\rm T}(i)\theta+v(i)-\varphi^{\rm T}(i)\hat{\theta}_{\rm LS}\right]\\ &\quad\times[\psi^{\rm T}(i)\theta+v(i)] &(33)\\ &=\sum_{i=1}^{t}\varphi^{\rm T}(i)[\theta-\hat{\theta}_{\rm LS}][\psi^{\rm T}(i)\theta+v(i)] &(34)\\ &\quad+\sum_{i=1}^{t}[\psi^{\rm T}(i)\theta+v(i)]^2. &(35)\end{aligned}$$

Notice that $v(t)$ is a white noise, dividing by t and taking limit give

$$\begin{aligned}&\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}\varepsilon_{\rm LS}^2(i)\\ &=\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}\varphi^{\rm T}(i)[\theta-\hat{\theta}_{\rm LS}][\psi^{\rm T}(i)\theta+v(i)] &(36)\\ &\quad+\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}[\psi^{\rm T}(i)\theta+v(i)]^2 &(37)\\ &=-\sigma^2\theta^{\rm T}\Lambda\lim_{t\to\infty}[\theta-\hat{\theta}_{\rm LS}]+\sigma^2+\sigma^2\theta^{\rm T}\Lambda\theta &(38)\\ &=\sigma^2[1+\theta^{\rm T}\Lambda\lim_{t\to\infty}\hat{\theta}_{\rm LS}] &(39)\end{aligned}$$

or

$$\sigma^2 = \frac{\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}\varepsilon_{\rm LS}^2(i)}{1+\theta^{\rm T}\Lambda\lim_{t\to\infty}\hat{\theta}_{\rm LS}}.$$

The estimate $\hat{\sigma}^2(t)$ of σ^2 may be computed by

$$\hat{\sigma}^2(t)=\frac{\frac{1}{t}J(t)}{1+\hat{\theta}_{\rm c}^{\rm T}(t)\Lambda\hat{\theta}_{\rm LS}(t)}$$

where

$$J(t)=\sum_{i=1}^{t}\varepsilon_{\rm LS}^2(i)=\sum_{i=1}^{t}\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t)\right]^2. \tag{40}$$

From the definition of $P(t)$ and (18), we easily get the recursive relation of $\hat{\theta}_{\rm LS}$ as follows:

$$\begin{aligned}\hat{\theta}_{\rm LS}(t)&=\hat{\theta}_{\rm LS}(t-1)+P(t)\varphi(t) &(41)\\ &\quad\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t-1)\right] &(42)\\ P^{-1}(t)&=P^{-1}(t-1)+\varphi(t)\varphi^{\rm T}(t) &(43)\end{aligned}$$

or

$$P(t)=P(t-1)-\frac{P(t-1)\varphi(t)\varphi^{\rm T}(t)P(t-1)}{1+\varphi^{\rm T}(t)P(t-1)\varphi(t)}.$$

Thus, we have

$$\begin{aligned}J(t)&=\sum_{i=1}^{t}\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t)\right]^2 &(44)\\ &=J(t-1)+\frac{\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t-1)\right]^2}{1+\varphi^{\rm T}(t)P(t-1)\varphi(t)}. &(45)\end{aligned}$$

From the above equations, we can summarize the BCRLS algorithm as follows:

$$\begin{aligned}\hat{\theta}_{\rm c}(t)&=\hat{\theta}_{\rm LS}(t)+t\,\hat{\sigma}^2(t)P(t)\hat{\theta}_{\rm c}(t-1) &(46)\\ \hat{\theta}_{\rm LS}(t)&=\hat{\theta}_{\rm LS}(t-1)+L(t)\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t-1)\right] &(47)\\ L(t)&=P(t-1)\varphi(t)[1+\varphi^{\rm T}(t)P(t-1)\varphi(t)]^{-1} &(48)\\ P(t)&=[I-L(t)\varphi^{\rm T}(t)]P(t-1),\ P(0)=p_0 I &(49)\\ J(t)&=J(t-1)+\frac{\left[y(t)-\varphi^{\rm T}(t)\hat{\theta}_{\rm LS}(t-1)\right]^2}{1+\varphi^{\rm T}(t)P(t-1)\varphi(t)} &(50)\\ \hat{\sigma}^2(t)&=\frac{J(t)}{t\left[1+\hat{\theta}_{\rm c}^{\rm T}(t-1)\Lambda\hat{\theta}_{\rm LS}(t)\right]} &(51)\\ &=\frac{J(t)}{t\left[1+\sum_{i=1}^{n}\hat{\theta}_{{\rm c}i}(t-1)\hat{\theta}_{{\rm LS}i}(t)\right]} &(52)\\ \varphi(t)&=[-y(t-1)\ -y(t-2)\ \cdots\ -y(t-n) &(53)\\ &\quad u_1(t)\ u_1(t-1)\cdots\ u_1(t-n)\cdots &(54)\\ &\quad u_r(t)\ u_r(t-1)\ \cdots u_r(t-n)]^{\rm T} &(55)\end{aligned}$$

where $\hat{\theta}_{{\rm c}i}(t)$ and $\hat{\theta}_{{\rm LS}i}(t)$ represent the ith element of $\hat{\theta}_{\rm c}(t)$ and $\hat{\theta}_{\rm LS}(t)$, respectively. To initialize this BCRLS algorithm, we take p_0 to be a large positive number, e.g., $p_0=10^6$, and take both $\hat{\theta}_{\rm LS}(0)$ and $\hat{\theta}_{\rm c}(0)$ to be zero vectors or some small real vectors, e.g., $\hat{\theta}_{\rm LS}(0)=\hat{\theta}_{\rm c}(0)=10^{-6}\mathbf{1}_{n_0}$ with $\mathbf{1}_{n_0}$ being an n_0-dimensional column vector whose elements are 1.

To summarize, we list the steps involved in the BCRLS algorithm to recursively compute the parameter estimation vector $\hat{\theta}_{\rm c}(t)$ as t increases:

TABLE I
ESTIMATES AND ERRORS FOR STATIONARY CASES ($\delta_{\rm ns} = 11.67\%$)

The LS estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.76131	0.09559	0.13614	0.33696	0.23580	0.21105	11.06308
200	-0.76471	0.09516	0.13093	0.32972	0.24768	0.20010	10.38983
300	-0.77813	0.10698	0.12342	0.32762	0.24087	0.20527	8.66702
500	-0.77729	0.10571	0.12474	0.32696	0.24470	0.20565	8.79196
1000	-0.79208	0.11766	0.11525	0.32903	0.25082	0.19651	6.69538
1500	-0.78688	0.11198	0.11481	0.32706	0.24718	0.19765	7.48387
2000	-0.78585	0.11194	0.11628	0.32746	0.24987	0.19700	7.54072
2500	-0.78548	0.11261	0.11448	0.32909	0.24941	0.19769	7.55706
3000	-0.78708	0.11393	0.11558	0.32802	0.24989	0.19685	7.31889
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	
The bias compensation recursive LS (BCRLS) estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.79515	0.12488	0.13634	0.33474	0.23842	0.20587	6.62243
200	-0.81031	0.13452	0.13163	0.32642	0.25036	0.19123	4.23439
300	-0.82854	0.15067	0.12441	0.32381	0.24248	0.19563	2.02125
500	-0.83067	0.15192	0.12519	0.32287	0.24662	0.19553	1.72718
1000	-0.84286	0.16141	0.11645	0.32540	0.25241	0.18624	0.56855
1500	-0.83887	0.15644	0.11631	0.32340	0.24913	0.18750	0.57109
2000	-0.83827	0.15682	0.11771	0.32373	0.25224	0.18688	0.50561
2500	-0.83885	0.15834	0.11595	0.32569	0.25193	0.18713	0.54633
3000	-0.84147	0.16067	0.11701	0.32489	0.25261	0.18599	0.44986
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	
The bias compensation LS (BCLS) estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.77866	0.12949	0.13651	0.33526	0.23645	0.20784	7.79134
200	-0.80189	0.13905	0.13099	0.32491	0.24793	0.19169	4.65091
300	-0.81980	0.14086	0.12338	0.32151	0.24013	0.19594	3.22973
500	-0.81641	0.12951	0.12399	0.32085	0.24420	0.19675	4.18709
1000	-0.83606	0.15336	0.11541	0.32306	0.24991	0.18595	0.92779
1500	-0.84112	0.15571	0.11530	0.31976	0.24643	0.18503	0.87976
2000	-0.84138	0.15504	0.11673	0.31992	0.24948	0.18426	0.77649
2500	-0.84390	0.16004	0.11489	0.32152	0.24915	0.18399	0.74193
3000	-0.84845	0.16682	0.11584	0.32055	0.24977	0.18232	1.30282
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	

1) Collect the input–output data $\{u(t), y(t)\}$, and data length L.
2) To initialize, let $t = 1 : p_0 = 10^6, P(0) = p_0 I_6, \hat{\theta}_{\rm LS}(0) = \hat{\theta}_{\rm c}(0) = \mathbf{1}_6/p_0, J(0) = 0$.
3) Form $\varphi(t)$ by (53), compute $J(t)$ by (50), $P(t)$ by (49), $L(t)$ by (48) and $\hat{\theta}_{\rm LS}(t)$ by (47).
4) Compute $\hat{\sigma}^2(t)$ by (51) and $\hat{\theta}_{\rm c}(t)$ by (46).
5) If $t = L + 1$, then terminate the procedure and obtain the estimate $\hat{\theta}_{\rm c}(L)$ of the parameter vector θ; otherwise, increment t by 1 and go to step 3.

III. EXAMPLE

An example is given to show the effectiveness of the proposed algorithms. Consider the 2-input and 1-output system

$$
\begin{aligned}
y(t) &= \frac{1}{\alpha(z)}[B_1(z)u_1(t) + B_2(z)u_2(t)] + v(t) \\
\alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} = 1 - 0.84z^{-1} + 0.16z^{-2} \\
B_1(z) &= \beta_{11} z^{-1} + \beta_{12} z^{-2} = 0.12z^{-1} + 0.324z^{-2}, \\
B_2(z) &= \beta_{21} z^{-1} + \beta_{22} z^{-2} = 0.25z^{-1} + 0.186z^{-2} \\
\theta &= [\alpha_1\ \alpha_2\ \beta_{11}\ \beta_{12}\ \beta_{21}\ \beta_{12}]^{\rm T}.
\end{aligned}
$$

$\{u_1(t), u_2(t)\}$ is taken as an uncorrelated persistent excitation vector sequence with zero mean and unit variance $\sigma_{u_1}^2 = \sigma_{u_2}^2 = 1.00^2$, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$. Under such conditions, this example gives rise to a stationary problem. Apply the LS and BCRLS algorithms and a comparable BCLS algorithm in [4] to estimate the parameters of this system, the LS estimates $\hat{\theta}_{\rm LS}(t)$, and bias compensation recursive LS (BCRLS) estimates $\hat{\theta}_{\rm c}(t)$, bias compensation LS (BCLS) estimates $\hat{\theta}_{\rm b}(t)$ and their errors are shown in Table I, and the parameter estimation errors $\delta = \|\hat{\theta}(t) - \theta\|/\|\theta\|$ versus t are shown in Fig. 1, where $\hat{\theta}(t)$ represents $\hat{\theta}_{\rm LS}(t)$ or $\hat{\theta}_{\rm c}(t)$ or $\hat{\theta}_{\rm b}(t)$, and $\delta_{\rm ns}$ represents the noise-to-signal ratio of the system and is defined by the square root of the ratio of the variance of the output of the system driven by the noise $v(t)$ and the noise-free output $x(t)$ (namely, the output $y(t)$ when $v(t) \equiv 0$). For the output error system in (2), $\delta_{\rm ns}$ is computed by the following:

$$
\begin{aligned}
\delta_{\rm ns} &= \sqrt{\frac{{\rm var}[v(t)]}{{\rm var}[x(t)]}} \times 100\% = \frac{\sigma_v}{\sigma_x} \times 100\% \\
x(t) &:= \frac{1}{\alpha(z)} \sum_{i=1}^{r} B_i(z) u_i(t).
\end{aligned}
$$

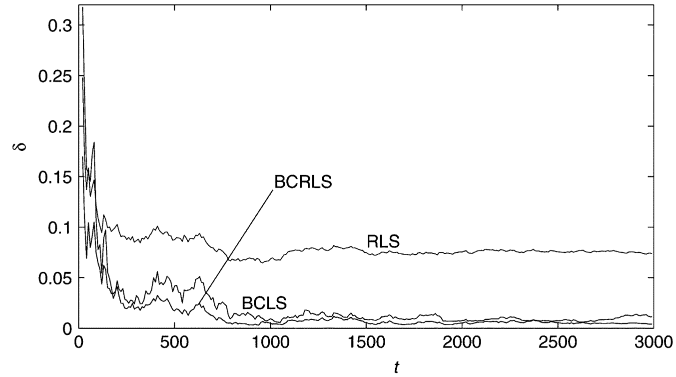

Fig. 1. The parameter estimation errors δ versus t (stationary case).

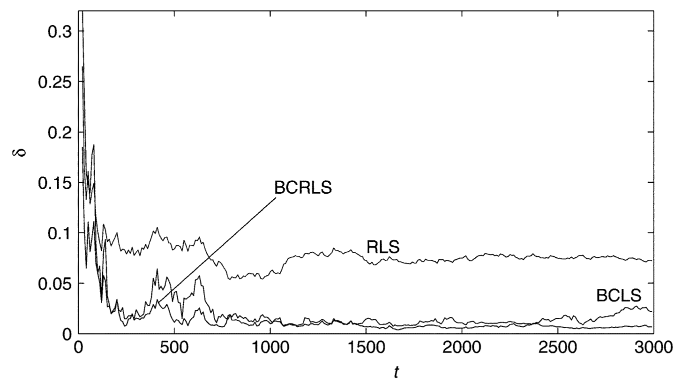

Fig. 2. The parameter estimation errors δ versus t (nonstationary case).

The simulation results with nonstationary cases are shown in Table II and Fig. 2, where the inputs $\{u_1'(t), u_2'(t)\}$ and noise $\{v'(t)\}$ are taken as

$$
\begin{aligned}
u_1'(t) &= (1 + t^{0.4})u_1(t) \\
u_2'(t) &= (1 + t^{0.4})u_2(t) \\
v'(t) &= (1 + t^{0.4})v(t).
\end{aligned}
$$

Under such cases, $R(t)/t$ is time varying and has no limit as $t \to \infty$ even if $v'(t) = v(t)$.

From Tables I and II and Figs. 1 and 2, we can see that the BCRLS and BCLS algorithms have obvious advantages over the LS algorithm, and the BCRLS estimates have higher accuracy and are more stationary than the comparable BCLS estimates in [4], especially in the nonstationary cases.

TABLE II
ESTIMATES AND ERRORS FOR NON-STATIONARY CASES

The LS estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.76806	0.10261	0.13862	0.33845	0.23329	0.21199	10.26160
200	-0.76637	0.09556	0.12962	0.32865	0.24955	0.19783	10.18384
300	-0.78328	0.11165	0.11993	0.32632	0.23926	0.20687	8.03036
500	-0.78024	0.10816	0.12337	0.32689	0.24667	0.20625	8.40823
1000	-0.79855	0.12337	0.11090	0.33033	0.25329	0.19273	5.84587
1500	-0.78872	0.11319	0.11200	0.32687	0.24659	0.19569	7.26272
2000	-0.78629	0.11249	0.11540	0.32729	0.25110	0.19557	7.45497
2500	-0.78575	0.11353	0.11292	0.32984	0.25005	0.19697	7.48303
3000	-0.78852	0.11560	0.11520	0.32786	0.25077	0.19555	7.08090
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	
The bias compensation recursive LS (BCRLS) estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.80074	0.13112	0.13885	0.33667	0.23568	0.20641	6.08875
200	-0.81541	0.13796	0.13070	0.32536	0.25224	0.18758	3.57969
300	-0.83718	0.15851	0.12137	0.32229	0.24026	0.19628	1.50784
500	-0.83697	0.15732	0.12385	0.32256	0.24838	0.19532	1.13787
1000	-0.84924	0.16692	0.11237	0.32679	0.25469	0.18221	1.57723
1500	-0.84135	0.15794	0.11380	0.32329	0.24858	0.18542	0.70677
2000	-0.83956	0.15796	0.11701	0.32360	0.25376	0.18535	0.54386
2500	-0.84034	0.16026	0.11453	0.32667	0.25287	0.18606	0.69233
3000	-0.84452	0.16378	0.11670	0.32505	0.25385	0.18422	0.82658
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	
The bias compensation LS (BCLS) estimates							
t	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	δ (%)
100	-0.79429	0.15341	0.14009	0.33648	0.23333	0.20624	5.97219
200	-0.81123	0.14967	0.13005	0.32312	0.24955	0.18697	3.30823
300	-0.83113	0.14631	0.12041	0.31944	0.23765	0.19552	2.36476
500	-0.82128	0.12719	0.12246	0.32026	0.24609	0.19676	4.08117
1000	-0.84372	0.16294	0.11135	0.32438	0.25199	0.18148	1.13434
1500	-0.84851	0.16361	0.11283	0.31896	0.24570	0.18174	1.44846
2000	-0.84709	0.16011	0.11609	0.31919	0.25092	0.18170	1.06745
2500	-0.84979	0.16750	0.11345	0.32189	0.25006	0.18184	1.51210
3000	-0.85557	0.17656	0.11550	0.32029	0.25104	0.17953	2.50296
True	-0.84000	0.16000	0.12000	0.32400	0.25000	0.18600	

IV. CONCLUSION

According to the bias correction technique, a bias compensation recursive LS identification algorithm is developed for multi-input, single-output systems without assuming that the system is stationary. The simulation results show that the proposed algorithm can give higher parameter estimation accuracy than the LS algorithm and bias compensation LS algorithm.

REFERENCES

[1] M. Verhaegen and P. Dewilde, "Subspace model identification Part 1: The output-error state space model identification class of algorithms," *Int. J. Contr.*, vol. 56, no. 5, pp. 1187–1210, 1992.

[2] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[3] P. Stoica and T. Söderström, "Bias correction in least-squares identification," *Int. J. Contr.*, vol. 35, no. 3, pp. 449–457, 1982.

[4] W. X. Zheng, "On a least-squares-based algorithm for identification of stochastic linear systems," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1631–1638, Dec. 1998.

[5] C. Z. Fang and D. Y. Xiao, *Process Identification*. Beijing, China: Tsinghua University Press, 1988.

[6] C. B. Feng and W. X. Zheng, "Robust identification of stochastic linear systems with correlated noise," *Proc. Inst. Elect. Eng. Control Theory Appl.*, pt. D, vol. 138, no. 5, pp. 484–492, 1991.

[7] P. Stoica, T. Söderström, and V. Simonyte, "Study of a bias-free least-squares parameter estimator," *Proc. Inst. Elect. Eng. Control Theory Appl.*, vol. 142, no. 1, pp. 1–6, 1995.

[8] W. X. Zheng, "Least-squares identification of a class of multivariable systems with correlated disturbances," *J. Franklin Inst.*, vol. 336, no. 8, pp. 1309–1324, 1999.

[9] ——, "Fast identification of autoregressive signals from noisy observations," *IEEE Trans. Circuits Syst. II, Expr. Briefs*, vol. 52, no. 1, pp. 43–48, Jan. 2005.

[10] ——, "Transfer function estimation from noisy input and output data," *Int. J. Adaptive Contr. Signal Process.*, vol. 12, no. 4, pp. 365–380, 1998.

[11] ——, "A bias correction method for identification of linear dynamic errors-in-variables models," *IEEE Trans. Autom. Contr.*, vol. 47, no. 7, pp. 1142–1147, Jul. 2002.

[12] ——, "Parameter estimation of stochastic linear systems with noisy input," *Int. J. Syst. Sci.*, vol. 35, no. 3, pp. 185–190, 2004.

[13] ——, "A bias-correction method for indirect identification of closed-loop systems," *Automatica*, vol. 31, no. 7–24, pp. 1019–1, 1995.

[14] ——, "Identification of closed-loop systems with low-order controllers," *Automatica*, vol. 32, no. 12, pp. 1753–1757, 1996.

[15] ——, "Parametric identification of linear systems operating under feedback control," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 48, no. 4, pp. 451–458, Apr. 2001.

[16] ——, "A modified method for closed-loop identification of transfer function models with common factors," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 4, pp. 556–562, Apr. 2002.

[17] ——, "On indirect identification of feedback-control systems via the instrumental variables methods the IV methods," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 9, pp. 1232–1238, Sep. 2003.

[18] F. Ding and T. Chen, "Identification of dual-rate systems based on finite impulse response models," *Int. J. Adaptive Contr. Signal Process.*, vol. 18, no. 7, pp. 589–598, 2004.

[19] ——, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, pp. 1739–1748, 2004.

[20] ——, "Modeling and identification of multirate systems," *Acta Automatica Sinica*, vol. 31, no. 1, pp. 105–122, 2005.

[21] ——, "Hierarchical gradient-based identification of multivariable discrete-time systems," *Automatica*, vol. 41, no. 2, pp. 315–325, 2005.

[22] ——, "Performance bounds of forgetting factor least-squares algorithm for time-varying systems with finite measurement data," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 555–566, Mar. 2005.