International Journal of Control, Automation and Systems 17(6) (2019) 1547-1557
http://dx.doi.org/10.1007/s12555-018-0512-0 ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Two-stage Recursive Least Squares Parameter Estimation Algorithm for Multivariate Output-error Autoregressive Moving Average Systems

Yunze Guo, Lijuan Wan, Ling Xu, Feng Ding*, Ahmed Alsaedi, and Tasawar Hayat

Abstract: This paper focuses on the parameter estimation problem of multivariate output-error autoregressive moving average (M-OEARMA) systems. By applying the auxiliary model identification idea and the decomposition technique, we derive a two-stage recursive least squares algorithm for estimating the M-OEARMA system. Compared with the auxiliary model based recursive least squares algorithm, the proposed algorithm possesses higher identification accuracy. The simulation results confirm the effectiveness of the proposed algorithm.

Keywords: Auxiliary model, decomposition technique, least squares, parameter estimation, recursive identification.

1. INTRODUCTION

Mathematical models are basic for designing controller and system analysis [1,2]. The parameter estimation methods of models can be applied to many areas [3–9]. In recent years, with the development of control theory and the demand of engineering practice, system identification and model parameter estimation have been extensively applied in almost all natural and man made systems [10–12].

In contrast to single variable systems, multivariate systems have more complex structures, uncertain disturbances and higher dimensions [13–15]. These characteristics make multivariate system identification difficult and therefore have drawn a great deal of attention [16,17]. How to improve the identification efficiency of multivariate systems has become an essential research field in multivariate system identification [18]. As for this, Pan et al. used the filtering technique and the multi-innovation identification theory to identify the multivariable system with moving average noise, and proposed the filtering based multi-innovation extended stochastic gradient algorithm to improve the parameter estimation accuracy [19].

Many identification methods have been applied to linear systems and nonlinear systems [20–22], such as the Newton methods [23, 24] and the least squares methods [25]. Compared with the stochastic gradient algorithm [26], the recursive least squares (RLS) algorithm has a fast convergence rate and can reach a satisfactory estimation accuracy [27]. To take advantage of its high estimation accuracy, Cho et al. presented a variable data-window-size recursive least-squares algorithm for dynamic system identification and the simulations proved that the proposed algorithm has a fast tracking performance and low misalignment error under a steady state [28].

Although the RLS algorithm is known for its high estimation accuracy, there are still many means which can improve its accuracy, such as the multi-innovation theory [29], the filtering method [30] and the decomposition technique [31]. The two-stage identification algorithm is based on the decomposition technique that can transform a large scale identification problem into small subproblems which are easier to solve [32, 33].

This paper studies the parameter estimation methods for multivariate output-error systems using the decomposition technique and meanwhile presents the condition of parameter convergence of the proposed approach [34, 35]. The main idea is to decompose the identification system into two subsystems and to identify each parameter vector separately. The difficulty is that the two subsystems have unknown associated variables. To deal with this problem, we establish the auxiliary models to replace the unknown variables in the identification algorithm with the outputs

Manuscript received July 24, 2018; revised October 29, 2018 and January 11, 2019; accepted January 30, 2019. Recommended by Associate Editor Yongping Pan under the direction of Editor Jay H. Lee. This work was supported by the National Natural Science Foundation of China (No. 61873111) and the 111 Project (B12018), Qing Lan Project, and sponsored by Postdoctoral Science Foundation of Jiangsu Province (No. 1701020A), the 333 project of Jiangsu Province (No.BRA2018328).

Yunze Guo, Ling Xu, and Feng Ding are with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, P. R. China (e-mails: yzguo12@126.com, lingxu0848@163.com, fding@jiangnan.edu.cn). Lijuan Wan and Feng Ding are with the College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, P. R. China (e-mail: ljwan@qust.edu.cn). Ling Xu is with the school of Internet of Things Technology, Wuxi Vocational Institute of commerce, Wuxi 214153, P. R. China. Feng Ding, Ahmed Alsaedi, and Tasawar Hayat are with the Department of Mathematics, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mails: aalsaedi@hotmail.com, fmgpak@gmail.com).
* Corresponding author.

of the auxiliary models [36]. The main contributions of this paper are as follows:

• A two-stage recursive least squares algorithm is proposed for the multivariate output-error autoregressive moving average systems by using the decomposition technique and the auxiliary model.

• Compared with the auxiliary model based recursive least squares algorithm, the proposed algorithm can generate more accurate estimates.

The rest of this paper is organized as follows: Section 2 describes the identification model of multivariate output-error autoregressive moving average systems. Section 3 gives the auxiliary model based recursive least squares algorithm for the obtained model. Section 4 derives a two-stage recursive least squares algorithm. An illustrative example is shown to verify the effectiveness of the proposed algorithms in Section 5. Finally, we offer some concluding remarks in Section 6.

2. THE SYSTEM DESCRIPTION

Some symbols are introduced. "$A =: X$" or "$X := A$" stands for "A is defined as X"; the superscript T stands for the vector/matrix transpose; the symbol I_n denotes an identity matrix of appropriate size ($n\times n$); $\hat{\vartheta}(t)$ denotes the estimate of ϑ at time t; $\mathbf{1}_n$ stands for an n-dimensional column vector whose elements are 1; the norm of a matrix (or a column vector) X is defined by $\|X\|^2 := \mathrm{tr}[XX^{\mathrm{T}}]$.

Consider the following multivariate output-error autoregressive moving average (M-OEARMA) system:

$$y(t)=\frac{\Phi_1(t)\theta}{A(z)}+\frac{D(z)}{C(z)}v(t), \tag{1}$$

where $y(t) := [y_1(t), y_2(t), \cdots, y_m(t)]^{\mathrm{T}} \in \mathbb{R}^m$ is the output vector of the system, $\theta \in \mathbb{R}^n$ is the system parameter vector to be identified, $\Phi_1(t) \in \mathbb{R}^{m\times n}$ is the information matrix consisting of the input signal $u(t) := [u_1(t), u_2(t), \cdots, u_r(t)] \in \mathbb{R}^r$ and the output signal $y(t)$, $v(t) := [v_1(t), v_2(t), \cdots, v_m(t)]^{\mathrm{T}} \in \mathbb{R}^m$ is a white noise vector, $A(z)$, $C(z)$ and $D(z)$ are the polynomials in the unit backward shift operator z^{-1} $[z^{-1}y(t) = y(t-1)]$, and defined as

$$\begin{aligned}
A(z) &:= 1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a} \in \mathbb{R},\\
C(z) &:= 1+c_1z^{-1}+c_2z^{-2}+\cdots+c_{n_c}z^{-n_c} \in \mathbb{R},\\
D(z) &:= 1+d_1z^{-1}+d_2z^{-2}+\cdots+d_{n_d}z^{-n_d} \in \mathbb{R}.
\end{aligned}$$

Let $n_1 := n+n_a$, $n_2 := n_c+n_d$, $n_0 := n_1+n_2$. Assume that the orders n_a, n_c and n_d are known and $\Phi_1(t) = \mathbf{0}$, $v(t) = \mathbf{0}$, $y(t) = \mathbf{0}$ for $t \leqslant 0$.

Define the intermediate variables:

$$x(t) := \frac{\Phi_1(t)\theta}{A(z)} \in \mathbb{R}^m,\quad w(t) := \frac{D(z)}{C(z)}v(t) \in \mathbb{R}^m.$$

It follows that

$$\begin{aligned}
x(t) &=[1-A(z)]x(t)+\Phi_1(t)\theta\\
&=-\sum_{j=1}^{n_a}a_jx(t-j)+\Phi_1(t)\theta, \qquad (2)\\
w(t) &=[1-C(z)]w(t)+D(z)v(t)\\
&=-\sum_{j=1}^{n_c}c_jw(t-j)+\sum_{j=1}^{n_d}d_jv(t-j)+v(t). \qquad (3)
\end{aligned}$$

Define the parameter vectors:

$$\begin{aligned}
&a := [a_1, a_2, \cdots, a_{n_a}]^{\mathrm{T}} \in \mathbb{R}^{n_a},\\
&\theta_s := [\theta^{\mathrm{T}}, a^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_1},\\
&\rho := [c_1, c_2, \cdots, c_{n_c}, d_1, d_2, \cdots, d_{n_d}]^{\mathrm{T}} \in \mathbb{R}^{n_2},\\
&\vartheta := [\theta_s^{\mathrm{T}}, \rho^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_0}.
\end{aligned}$$

Define the information matrices:

$$\begin{aligned}
&\Phi_x(t) := [-x(t-1), \cdots, -x(t-n_a)] \in \mathbb{R}^{m\times n_a},\\
&\Phi_s(t) := [\Phi_1(t), \Phi_x(t)] \in \mathbb{R}^{m\times n_1},\\
&\Phi_{\mathrm{n}}(t) := [-w(t-1), \cdots, -w(t-n_c),\\
&\qquad\qquad v(t-1), \cdots, v(t-n_d)] \in \mathbb{R}^{m\times n_2},\\
&\Phi(t) := [\Phi_s(t), \Phi_{\mathrm{n}}(t)] \in \mathbb{R}^{m\times n_0}.
\end{aligned}$$

Then, equations (2) and (3) can be written as

$$x(t) = \Phi_1(t)\theta+\Phi_x(t)a, \tag{4}$$

$$w(t) = \Phi_{\mathrm{n}}(t)\rho+v(t). \tag{5}$$

Substituting (4) and (5) into (1) gives

$$\begin{aligned}
y(t) &=x(t)+w(t)\\
&=\Phi_1(t)\theta+\Phi_x(t)a+w(t) \qquad (6)\\
&=\Phi_s(t)\theta_s+\Phi_{\mathrm{n}}(t)\rho+v(t) \qquad (7)\\
&=\Phi(t)\vartheta+v(t). \qquad (8)
\end{aligned}$$

In this model, the new parameter vector ϑ contains the parameter vector θ_s of the system model and the parameter vector ρ of the noise model.

The objective of this paper is to use the auxiliary model identification idea and the decomposition technique to derive new methods for estimating the parameter vector ϑ from the observation data $y(t)$ and $\Phi_1(t)$ and to confirm the theoretical result with a simulation example.

3. THE AUXILIARY MODEL BASED RECURSIVE LEAST SQUARES ALGORITHM

According to the identification model in (8), define a least squares criterion function:

$$J(\vartheta) := \sum_{j=1}^{t}\|y(j)-\Phi(j)\vartheta\|^2. \tag{9}$$

Minimizing the criterion function $J(\vartheta)$ gives:

$$\begin{aligned}\hat{\vartheta}(t) &= \hat{\vartheta}(t-1)+P(t)\Phi^{\mathrm{T}}(t)\\ &\quad\times[y(t)-\Phi(t)\hat{\vartheta}(t-1)],\\ P^{-1}(t) &= P^{-1}(t-1)+\Phi^{\mathrm{T}}(t)\Phi(t).\end{aligned}\tag{10}$$

Applying the matrix inversion formula

$$(A+BC)^{-1}=A^{-1}-A^{-1}B(I+CA^{-1}B)^{-1}CA^{-1}$$

to (10) gives

$$\begin{aligned}P(t)=&[I_{n_0}-P(t-1)\Phi^{\mathrm{T}}(t)\\ &\times[I_m+\Phi(t)P(t-1)\Phi^{\mathrm{T}}(t)]^{-1}\Phi(t)]P(t-1).\end{aligned}$$

Then, we can obtain the following recursive least squares algorithm:

$$\hat{\vartheta}(t)=\hat{\vartheta}(t-1)+L(t)[y(t)-\Phi(t)\hat{\vartheta}(t-1)],\tag{11}$$

$$\begin{aligned}L(t)&=P(t-1)\Phi^{\mathrm{T}}(t)\\ &\quad\times[I_m+\Phi(t)P(t-1)\Phi^{\mathrm{T}}(t)]^{-1},\end{aligned}\tag{12}$$

$$P(t)=[I_{n_0}-L(t)\Phi(t)]P(t-1).\tag{13}$$

In the recursive algorithm, the initial value of the parameter estimation vector is generally taken to be zero or a very small real vector, for example, $\hat{\vartheta}(0)=\mathbf{1}_{n_0}/p_0$, $p_0=10^6$.

Here, some problems arise. The information matrix $\Phi(t)$ in (11)-(13) contains the unknown terms $x(t-i)$, $w(t-i)$ and $v(t-i)$. Therefore the estimate $\hat{\vartheta}(t)$ in (11) is impossible to compute. An effective method to solve this problem is to employ the auxiliary model identification idea. That is to establish an auxiliary model by using the measurable information in the identification algorithm and to replace the unknown variables in the system with the outputs of the auxiliary model. Establish an appropriate auxiliary model and use their outputs $x_a(t-i)$, $\hat{w}(t-i)$ and $\hat{v}(t-i)$ to define the estimates $\hat{\Phi}_x(t)$, $\hat{\Phi}_s(t)$, $\hat{\Phi}_{\mathrm{n}}(t)$, $\hat{\Phi}(t)$ of $\Phi_x(t)$, $\Phi_s(t)$, $\Phi_{\mathrm{n}}(t)$, $\Phi(t)$ as

$$\hat{\Phi}_x(t):=[-x_a(t-1),\cdots,-x_a(t-n_a)]\in\mathbb{R}^{m\times n_a},\tag{14}$$

$$\hat{\Phi}_s(t):=[\Phi_1(t),\hat{\Phi}_x(t)]\in\mathbb{R}^{m\times n_1},\tag{15}$$

$$\begin{aligned}\hat{\Phi}_{\mathrm{n}}(t):=&[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),\\ &\hat{v}(t-1),\cdots,\hat{v}(t-n_d)]\in\mathbb{R}^{m\times n_2},\end{aligned}\tag{16}$$

$$\hat{\Phi}(t):=[\Phi_1(t),\hat{\Phi}_x(t),\hat{\Phi}_{\mathrm{n}}(t)]\in\mathbb{R}^{m\times n_0}.\tag{17}$$

According to (4), use the estimates $\hat{\Phi}_x(t)$, $\hat{\theta}(t)$ and $\hat{a}(t)$ to define the outputs $x_a(t)$ of the auxiliary model as

$$x_a(t):=\Phi_1(t)\hat{\theta}(t)+\hat{\Phi}_x(t)\hat{a}(t).$$

Similarly, from (6), the estimate $\hat{w}(t)$ can be computed through

$$\begin{aligned}\hat{w}(t):=&y(t)-\Phi_1(t)\hat{\theta}(t)-\hat{\Phi}_x(t)\hat{a}(t)\\ =&y(t)-x_a(t).\end{aligned}\tag{18}$$

According to (8), the residual $\hat{v}(t)$ can be computed by

$$\hat{v}(t):=y(t)-\hat{\Phi}(t)\hat{\vartheta}(t).\tag{19}$$

Replacing $\Phi(t)$ in (11)-(13) with its estimate $\hat{\Phi}(t)$ and combining (14)-(19), we can obtain the following auxiliary model based recursive least squares (AM-RLS) algorithm:

$$\hat{\vartheta}(t)=\hat{\vartheta}(t-1)+L(t)[y(t)-\hat{\Phi}(t)\hat{\vartheta}(t-1)],\tag{20}$$

$$\begin{aligned}L(t)&=P(t-1)\hat{\Phi}^{\mathrm{T}}(t)\\ &\quad\times[I_m+\hat{\Phi}(t)P(t-1)\hat{\Phi}^{\mathrm{T}}(t)]^{-1},\end{aligned}\tag{21}$$

$$P(t)=[I_{n_0}-L(t)\hat{\Phi}(t)]P(t-1),$$

$$\hat{\Phi}(t)=[\Phi_1(t),\hat{\Phi}_x(t),\hat{\Phi}_{\mathrm{n}}(t)],\tag{22}$$

$$\hat{\Phi}_x(t)=[-x_a(t-1),\cdots,-x_a(t-n_a)],\tag{23}$$

$$\begin{aligned}\hat{\Phi}_{\mathrm{n}}(t)=&[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),\\ &\hat{v}(t-1),\cdots,\hat{v}(t-n_d)],\end{aligned}\tag{24}$$

$$x_a(t)=\Phi_1(t)\hat{\theta}(t)+\hat{\Phi}_x(t)\hat{a}(t),\tag{25}$$

$$\hat{w}(t)=y(t)-x_a(t),\tag{26}$$

$$\hat{v}(t)=y(t)-\hat{\Phi}(t)\hat{\vartheta}(t),\tag{27}$$

$$\hat{\vartheta}(t)=[\hat{\theta}^{\mathrm{T}}(t),\hat{a}^{\mathrm{T}}(t),\hat{\rho}^{\mathrm{T}}(t)]^{\mathrm{T}}.\tag{28}$$

The procedure for computing the parameter estimation vector $\hat{\vartheta}(t)$ in the AM-RLS algorithm in (20)-(28) is as follows:

1) Set the data length L ($L\gg n$). Let $t=1$, $P(0)=p_0I_{n_0}$, $\hat{\vartheta}(0)=\mathbf{1}_{n_0}/p_0$, $x_a(t-i)=\mathbf{1}_m/p_0$, $\hat{w}(t-i)=\mathbf{1}_m/p_0$, $\hat{v}(t-i)=\mathbf{1}_m/p_0$, $i=1,2,\cdots,\max[n_a,n_c,n_d]$, $p_0=10^6$.
2) Collect the observation data $y(t)$ and $\Phi_1(t)$, and construct the information matrices $\hat{\Phi}_x(t)$, $\hat{\Phi}_{\mathrm{n}}(t)$ and $\hat{\Phi}(t)$ using (22)-(24).
3) Compute the gain matrix $L(t)$ and the covariance matrix $P(t)$ according to (21)-(22).
4) Update the parameter estimation vector $\hat{\vartheta}(t)$ using (20).
5) Compute $x_a(t)$, $\hat{w}(t)$ and $\hat{v}(t)$ using (25)-(27).
6) If $t=L$, obtain the parameter estimate $\vartheta(L)$; otherwise, increase t by 1 and go to Step 2.

In order to study the convergence analysis of the proposed algorithm, assume that the noise vector $v(t)$ satisfies

(A1) $\mathrm{E}[v(t)]=\mathbf{0}$,

(A2) $\mathrm{E}[\|v(t)\|^2]\leqslant\sigma^2<\infty$.

Theorem 1: For the identification model in (8) and the AM-RLS algorithm in (20)-(28), assume that there exist positive constants α and β and large t such that the following persistent condition holds:

(A3) $\alpha I_{n_0}\leqslant\frac{1}{t}\sum_{j=1}^{t}\hat{\Phi}^{\mathrm{T}}(j)\hat{\Phi}(j)\leqslant\beta I_{n_0}$.

Then, the parameter estimation error $\|\hat{\vartheta}(t)-\vartheta\|$ converges to zero as t goes to infinity.

Proof: Referring to the method in [37], we have

$$\|\hat{\vartheta}(t)-\vartheta\|^2 = O\left(\frac{(\ln\lambda_{\max}[P^{-1}(t)])^{1+c}}{\lambda_{\min}[P^{-1}(t)]}\right), \quad \text{a.s., for } c>0. \tag{29}$$

Based on (10), we can deduce

$$\begin{aligned}P^{-1}(t) &= P^{-1}(t-1)+\hat{\Phi}^{\mathrm{T}}(t)\hat{\Phi}(t)\\ &= P^{-1}(t-2)+\hat{\Phi}^{\mathrm{T}}(t-1)\hat{\Phi}(t-1)+\hat{\Phi}^{\mathrm{T}}(t)\hat{\Phi}(t)\\ &= \sum_{j=1}^{t}\hat{\Phi}^{\mathrm{T}}(j)\hat{\Phi}(j)+P^{-1}(0)\\ &= \sum_{j=1}^{t}\hat{\Phi}^{\mathrm{T}}(j)\hat{\Phi}(j)+\frac{I_{n_0}}{p_0}.\end{aligned}$$

Using (A3), we have

$$\lambda_{\max}[P^{-1}(t)] \leqslant n_0\beta t,$$
$$\lambda_{\min}[P^{-1}(t)] \geqslant n_0\alpha t.$$

Then, equation (29) can be expressed as

$$\begin{aligned}\|\hat{\vartheta}(t)-\vartheta\|^2 &= O\left(\frac{(\ln\lambda_{\max}[P^{-1}(t)])^{1+c}}{\lambda_{\min}[P^{-1}(t)]}\right)\\ &= O\left(\frac{[\ln(n_0\beta t)]^{1+c}}{\alpha t}\right)\\ &= O\left(\frac{[\ln t]^{1+c}}{t}\right)\to 0, \text{ a.s., for } c>0.\end{aligned}$$

This proves Theorem 1. □

4. THE TWO-STAGE RECURSIVE LEAST SQUARES ALGORITHM

The basic idea of the two-stage recursive least squares identification method is to decompose the identification system in (7) into two subsystems and to identify the parameter vector of each subsystem separately. Define two intermediate output variables:

$$y_1(t) := y(t)-\Phi_{\mathrm{n}}(t)\rho \in \mathbb{R}^m, \tag{30}$$
$$y_2(t) := y(t)-\Phi_s(t)\theta_s \in \mathbb{R}^m. \tag{31}$$

This system can be decomposed into the following two fictitious subsystems:

$$y_1(t) = \Phi_s(t)\theta_s+v(t),$$
$$y_2(t) = \Phi_{\mathrm{n}}(t)\rho+v(t).$$

Define the cost functions:

$$J_1(\theta_s) := \sum_{j=1}^{t}\|y_1(j)-\Phi_s(j)\theta_s\|^2,$$
$$J_2(\rho) := \sum_{j=1}^{t}\|y_2(j)-\Phi_{\mathrm{n}}(j)\rho\|^2.$$

Let $\hat{\theta}_s(t)$ and $\hat{\rho}(t)$ be the estimates of θ_s and ρ at time t. Letting the partial derivatives of $J_1(\theta_s)$ with respect to θ_s and $J_2(\rho)$ with respect to ρ be zero gives

$$\frac{\partial J_1(\theta_s)}{\partial\theta_s} = -2\sum_{j=1}^{t}\Phi_s^{\mathrm{T}}(j)[y_1(j)-\Phi_s(j)\theta_s] = \mathbf{0},$$
$$\frac{\partial J_2(\rho)}{\partial\rho} = -2\sum_{j=1}^{t}\Phi_{\mathrm{n}}^{\mathrm{T}}(j)[y_2(j)-\Phi_{\mathrm{n}}(j)\rho] = \mathbf{0}.$$

Then, we can obtain the following least squares algorithm:

$$\hat{\theta}_s(t) = \hat{\theta}_s(t-1)+L_1(t)\times[y_1(t)-\Phi_s(t)\hat{\theta}_s(t-1)], \tag{32}$$
$$L_1(t) = P_1(t-1)\Phi_s^{\mathrm{T}}(t)[I_m+\Phi_s(t)P_1(t)\Phi_s^{\mathrm{T}}(t)]^{-1},$$
$$P_1(t) = [I_{n_1}-L_1(t)\Phi_s(t)]P_1(t-1),$$
$$\hat{\rho}(t) = \hat{\rho}(t-1)+L_2(t)\times[y_2(t)-\Phi_{\mathrm{n}}(t)\hat{\rho}(t-1)], \tag{33}$$
$$L_2(t) = P_2(t-1)\Phi_{\mathrm{n}}^{\mathrm{T}}(t)[I_m+\Phi_{\mathrm{n}}(t)P_2(t)\Phi_{\mathrm{n}}^{\mathrm{T}}(t)]^{-1},$$
$$P_2(t) = [I_{n_2}-L_2(t)\Phi_{\mathrm{n}}(t)]P_2(t-1).$$

Substituting (30) into (32) and (31) into (33) gives

$$\hat{\theta}_s(t) = \hat{\theta}_s(t-1)+L_1(t)\times[y(t)-\Phi_{\mathrm{n}}(t)\rho-\Phi_s(t)\hat{\theta}_s(t-1)], \tag{34}$$
$$\hat{\rho}(t) = \hat{\rho}(t-1)+L_2(t)\times[y(t)-\Phi_s(t)\theta_s-\Phi_{\mathrm{n}}(t)\hat{\rho}(t-1)]. \tag{35}$$

Here, we notice that the right-hand sides of (34) and (35) contain the unknown parameter vectors ρ and θ_s respectively. The solution is to replace the unknown ρ in (34) and θ_s in (35) with their corresponding estimates $\hat{\rho}(t)$ and $\hat{\theta}_s(t)$ at $t-1$. Then, we have

$$\begin{aligned}\hat{\theta}_s(t) &= \hat{\theta}_s(t-1)+L_1(t)\times[y(t)-\Phi_{\mathrm{n}}(t)\hat{\rho}(t-1)-\Phi_s(t)\hat{\theta}_s(t-1)]\\ &= \hat{\theta}_s(t-1)+L_1(t)[y(t)-\Phi(t)\hat{\vartheta}(t-1)],\end{aligned}$$
$$\begin{aligned}\hat{\rho}(t) &= \hat{\rho}(t-1)+L_2(t)\times[y(t)-\Phi_s(t)\hat{\theta}_s(t-1)-\Phi_{\mathrm{n}}(t)\hat{\rho}(t-1)]\\ &= \hat{\rho}(t-1)+L_2(t)[y(t)-\Phi(t)\hat{\vartheta}(t-1)].\end{aligned}$$

Here, we can obtain the two-stage least squares (TS-RLS) algorithm:

$$\hat{\theta}_s(t) = \hat{\theta}_s(t-1)+L_1(t)[y(t)-\hat{\Phi}(t)\hat{\vartheta}(t-1)], \tag{36}$$
$$L_1(t) = P_1(t-1)\hat{\Phi}_s^{\mathrm{T}}(t)\times[I_m+\hat{\Phi}_s(t)P_1(t-1)\hat{\Phi}_s^{\mathrm{T}}(t)]^{-1}, \tag{37}$$
$$P_1(t) = [I_{n_1}-L_1(t)\hat{\Phi}_s(t)]P_1(t-1), \tag{38}$$

$$\hat{\rho}(t)=\hat{\rho}(t-1)+L_2(t)[y(t)-\hat{\Phi}(t)\hat{\vartheta}(t-1)], \tag{39}$$

$$L_2(t)=P_2(t-1)\hat{\Phi}_{\rm n}^{\rm T}(t)$$
$$\times[I_m+\hat{\Phi}_{\rm n}(t)P_2(t-1)\hat{\Phi}_{\rm n}^{\rm T}(t)]^{-1}, \tag{40}$$

$$P_2(t)=[I_{n_2}-L_2(t)\hat{\Phi}_{\rm n}(t)]P_2(t-1), \tag{41}$$

$$\hat{\Phi}_s(t)=[\Phi_1(t),\hat{\Phi}_x(t)], \tag{42}$$

$$\hat{\Phi}_x(t)=[-x_a(t-1),\cdots,-x_a(t-n_a)], \tag{43}$$

$$\hat{\Phi}_{\rm n}(t)=[-\hat{w}(t-1),\cdots,-\hat{w}(t-n_c),$$
$$\hat{v}(t-1),\cdots,\hat{v}(t-n_d)], \tag{44}$$

$$\hat{\Phi}(t)=[\Phi_1(t),\hat{\Phi}_x(t),\hat{\Phi}_{\rm n}(t)], \tag{45}$$

$$x_a(t)=\Phi_1(t)\hat{\theta}(t)+\hat{\Phi}_x(t)\hat{a}(t), \tag{46}$$

$$\hat{w}(t)=y(t)-x_a(t), \tag{47}$$

$$\hat{v}(t)=y(t)-\hat{\Phi}(t)\hat{\vartheta}(t). \tag{48}$$

The steps of implementing the TS-RLS algorithm in (36)-(48) to estimate θ_s and ρ are listed in the following:

1) Set the data length L ($L\gg n$). Let $t=1$, $P_1(0)=p_0 I_{n_1}$, $P_2(0)=p_0 I_{n_2}$, $\hat{\theta}_s(0)=\mathbf{1}_{n_1}/p_0$, $\hat{\rho}(0)=\mathbf{1}_{n_2}/p_0$, $\hat{w}(t-i)=\mathbf{1}_m/p_0$, $\hat{v}(t-i)=\mathbf{1}_m/p_0$, $i=1,2,\cdots,\max[n_a,n_c,n_d]$, $p_0=10^6$.
2) Collect the input/output data $y(t)$ and $\Phi_1(t)$, and construct the information matrices $\hat{\Phi}_x(t)$, $\hat{\Phi}_{\rm n}(t)$, $\hat{\Phi}_s(t)$ and $\hat{\Phi}(t)$ using (42)-(45).
3) Compute the gain matrices $L_1(t)$ and $L_2(t)$ by (37) and (40), and update the covariance matrices $P_1(t)$ and $P_2(t)$ through (38) and (41).
4) Update the parameter estimates $\hat{\theta}_s(t)$ and $\hat{\rho}(t)$ using (36) and (39), respectively.
5) Compute the outputs $x_a(t)$, $\hat{w}(t)$ and $\hat{v}(t)$ of the auxiliary models by (46)-(48).
6) If $t=L$, obtain the parameter estimation vectors $\hat{\theta}_s(t)$ and $\hat{\rho}(t)$; otherwise, increase t by 1 and go to Step 2.

Theorem 2: For the identification model in (7) and the TS-RLS algorithm in (36)-(48), assume that there exist positive constants α and β and large t such that the following persistent condition holds:

$$\text{(A4)}\quad \alpha I_{n_1}\leqslant\frac{1}{t}\sum_{j=1}^{t}\hat{\Phi}_s^{\rm T}(j)\hat{\Phi}_s(j)\leqslant\beta I_{n_1},$$

$$\text{(A5)}\quad \alpha I_{n_2}\leqslant\frac{1}{t}\sum_{j=1}^{t}\hat{\Phi}_{\rm n}^{\rm T}(j)\hat{\Phi}_{\rm n}(j)\leqslant\beta I_{n_2}.$$

Then the parameter estimation errors $\|\hat{\theta}_s(t)-\theta_s\|$ and $\|\hat{\rho}(t)-\rho\|$ converge to zero as t goes to infinity.

The proof can be done in a similar way in [38, 39] and is omitted here.

In system identification, one usually uses the flop to assess the amount of computational efficiency of an algorithm. The computational efficiency of the proposed algorithms at every recursive calculation is shown in Tables 1-2. Their total numbers are as follows:

$$N_1:=3n_0m^2+m(4n_0^2+5n_0+2q),$$

$$N_2:=3n_0m^2+m(4q^2+4r^2+5n_0+2q).$$

It is clear that $N_e:=N_2-N_1=8qrm>0$, which means the TS-RLS has higher computational efficiency.

Remark 1: It is worth pointing out that least squares algorithms are suitable for linear regressive models. The multivariate output-error autoregressive moving average (M-OEARMA) system in this paper is a linear-parameter system instead of a nonlinear-parameter system. Therefore, the least squares can be applied to present new two-stage methods.

Remark 2: There are many other two-stage estimation methods, such as the two-stage least squares iterative algorithm and the two-stage stochastic gradient algorithm. Compared with the two-stage least squares iterative algorithm, the proposed method requires less computational load. Compared with the two-stage stochastic gradient algorithm, the proposed method possesses higher identification accuracy.

Many methods have been proposed to deal with the linear multivariable systems, such as the filtering based stochastic gradient algorithm [40] and the filtering based recursive least squares algorithm [39]. Unlike the filtering method, which is used to change the structure of the disturbance noise model, the decomposition technique is used to decompose the identification model into the system model and the noise model and identify each parameter vector separately.

5. EXAMPLE

Consider the following M-OEARMA system:

$$y(t)=\frac{\Phi_1(t)\theta}{A(z)}+\frac{D(z)}{C(z)}v(t),$$

$$\Phi_1(t)=\begin{bmatrix} y_2(t-1) & u_1(t-1) & u_2(t-1)\cos(t/\pi)\\ y_1(t-2) & u_2(t-2)\sin(t/\pi) & u_1(t-1)\end{bmatrix},$$

$$A(z)=1+0.69z^{-1}+0.20z^{-2},$$

$$C(z)=1+0.38z^{-1}+0.83z^{-2},$$

$$D(z)=1-0.58z^{-1}+0.38z^{-2},$$

$$\theta=[\theta_1,\theta_2,\theta_3]^{\rm T}=[-0.12,0.38,-0.48]^{\rm T},$$

$$\vartheta=[-0.12,0.38,-0.48,0.69,0.20,0.38,$$
$$0.83,-0.58,0.38]^{\rm T}.$$

In simulation, the inputs $u_1(t)$ and $u_2(t)$ are taken as two independent persistent excitation signal sequences with zero mean and unit variance. $v_1(t)$ and $v_2(t)$ are taken as two white noise sequences with zero mean and variances σ_1^2 for $v_1(t)$ and σ_2^2 for $v_2(t)$. The sequence input obtained by using the Matlab function 'idinput' is pseudorandom. Take $\sigma_1^2=\sigma_2^2=\sigma^2=0.10^2$ and $\sigma_1^2=\sigma_2^2=\sigma^2=0.30^2$, respectively. Based on the above model, we generate the system output signals $y(t)=[y_1(t),y_2(t)]^{\rm T}$. By using $u_1(t)$,

Table 1. The computational efficiency of the AM-RLS algorithm ($n_0 := q+r$).

Expressions	Multiplications	Additions
$\hat{\vartheta}(t) = \hat{\vartheta}(t-1) + L(t)e(t) \in \mathbb{R}^{n_0}$	mn_0	mn_0
$e(t) := y(t) - \hat{\Phi}(t)\hat{\vartheta}(t-1) \in \mathbb{R}^m$	mn_0	mn_0
$Q(t) := P(t-1)\hat{\Phi}^{\mathrm{T}}(t) \in \mathbb{R}^{n_0 \times m}$	mn_0^2	$m(n_0^2 - n_0)$
$L(t) = Q(t)/[I_m + \hat{\Phi}(t)Q(t)] \in \mathbb{R}^{n_0 \times m}$	$2m^2 n_0$	$m^2 n_0$
$P(t) = P(t-1) - L(t)Q^{\mathrm{T}}(t) \in \mathbb{R}^{n_0 \times n_0}$	mn_0^2	mn_0^2
$x_a(t) = \Phi_s(t)\hat{\theta}_s(t) \in \mathbb{R}^m$	mq	$m(q-1)$
$\hat{w}(t) = y(t) - x_a(t) \in \mathbb{R}^m$	0	m
$\hat{v}(t) = y(t) - \hat{\Phi}(t)\hat{\vartheta}(t) \in \mathbb{R}^m$	mn_0	mn_0
Sum	$2m^2 n_0 + m(2n_0^2 + 3n_0 + q)$	$m^2 n_0 + m(2n_0^2 + 2n_0 + q)$
Total flops	$N_1 := 3n_0 m^2 + m(4n_0^2 + 5n_0 + 2q)$	

Table 2. The computational efficiency of the TS-RLS algorithm ($n_0 := q+r$).

Expressions	Multiplications	Additions
$\hat{\theta}_s(t) = \hat{\theta}_s(t-1) + L_1(t)e(t) \in \mathbb{R}^q$	mq	mq
$e(t) := y(t) - \hat{\Phi}(t)\hat{\vartheta}(t-1) \in \mathbb{R}^m$	mn_0	mn_0
$Q_1(t) := P_1(t-1)\hat{\Phi}_s^{\mathrm{T}}(t) \in \mathbb{R}^{q \times m}$	mq^2	$m(q^2 - q)$
$L_1(t) = Q_1(t)/[I_m + \hat{\Phi}_s(t)Q_1(t)] \in \mathbb{R}^{q \times m}$	$2qm^2$	qm^2
$P_1(t) = P_1(t-1) - L_1(t)Q_1^{\mathrm{T}}(t) \in \mathbb{R}^{q \times q}$	mq^2	mq^2
$\hat{\rho}(t) = \hat{\rho}(t-1) + L_2(t)e(t) \in \mathbb{R}^r$	mr	mr
$Q_2(t) := P_2(t-1)\hat{\Phi}_{\mathrm{n}}^{\mathrm{T}}(t) \in \mathbb{R}^{r \times m}$	mr^2	$m(r^2 - r)$
$L_2(t) = Q_2(t)/[I_m + \hat{\Phi}_{\mathrm{n}}(t)Q_2(t)] \in \mathbb{R}^{r \times m}$	$2rm^2$	rm^2
$P_2(t) = P_2(t-1) - L_2(t)Q_2^{\mathrm{T}}(t) \in \mathbb{R}^{r \times r}$	mr^2	mr^2
$x_a(t) = \Phi_s(t)\hat{\theta}_s(t) \in \mathbb{R}^m$	mq	$m(q-1)$
$\hat{w}(t) = y(t) - x_a(t) \in \mathbb{R}^m$	0	m
$\hat{v}(t) = y(t) - \hat{\Phi}(t)\hat{\vartheta}(t) \in \mathbb{R}^m$	mn_0	mn_0
Sum	$2n_0 m^2 + m(2q^2 + 2r^2 + 3n_0 + q)$	$n_0 m^2 + m(2q^2 + 2r^2 + 2n_0 + q)$
Total flops	$N_2 := 3n_0 m^2 + m(4q^2 + 4r^2 + 5n_0 + 2q)$	

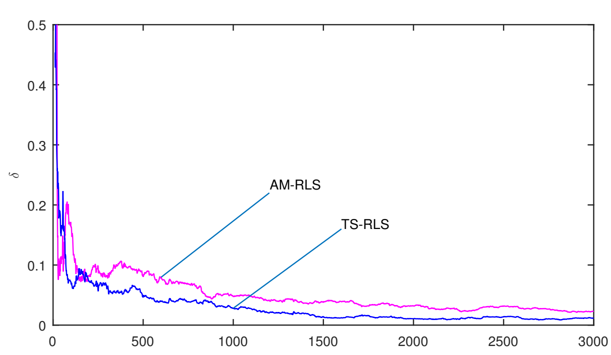


Fig. 1. The AM-RLS and TS-RLS estimation errors δ versus t with $\sigma^2 = 0.10^2$.

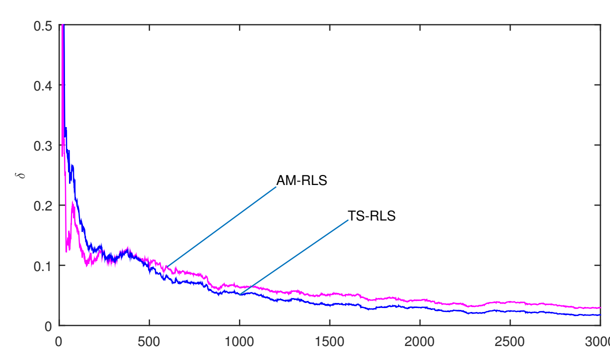


Fig. 2. The AM-RLS and TS-RLS estimation errors δ versus t with $\sigma^2 = 0.30^2$.

$u_2(t)$, $y_1(t)$ and $y_2(t)$ and applying the proposed algorithms to estimate the parameters of this system, the simulation results are shown in Tables 3-6 and Figs. 1-5. The results of Monte-Carlo simulations are shown in Tables 7-8.

Remark 3: One commonly uses the estimation error δ to evaluate the parameter estimation accuracy. In other words, the smaller the estimation errors, the more accurate the parameter estimates. Due to the interference from colored noise, the estimation error has fluctuation. But generally the parameter estimation error become smaller with the data length t increasing.

From Tables 3-8 and Figs. 1-2, we can draw the following conclusions.

- The parameter estimation errors of the AM-RLS algorithm and the TS-RLS algorithm become smaller with the data length t increasing - see Tables 3-8.
- The TS-RLS algorithm leads to smaller parameter es-

Table 3. The AM-RLS estimates and their errors with $\sigma^2 = 0.10^2$.

t	θ_1	θ_2	θ_3	a_1	a_2	c_1	c_2	d_1	cd_2	δ (%)
100	-0.13598	0.39612	-0.48205	0.68172	0.20595	0.43921	0.89637	-0.38154	0.27177	16.39194
200	-0.12448	0.38623	-0.48528	0.67908	0.20674	0.41995	0.86164	-0.47586	0.34782	8.14669
500	-0.12328	0.38321	-0.48146	0.68649	0.20338	0.39829	0.83080	-0.51051	0.28486	8.01656
1000	-0.12444	0.38346	-0.48058	0.68977	0.20128	0.39955	0.83029	-0.55568	0.31672	4.75200
2000	-0.12386	0.37849	-0.47819	0.69279	0.20160	0.40050	0.83761	-0.55486	0.34612	3.21136
3000	-0.12157	0.38001	-0.47845	0.69101	0.20062	0.39439	0.83030	-0.55310	0.36476	2.29523
True values	-0.12000	0.38000	-0.48000	0.69000	0.20000	0.38000	0.83000	-0.58000	0.38000	

Table 4. The TS-RLS estimates and their errors with $\sigma^2 = 0.10^2$.

t	θ_1	θ_2	θ_3	a_1	a_2	c_1	c_2	d_1	d_2	δ (%)
100	-0.13221	0.38513	-0.47322	0.68595	0.17344	0.32776	0.76216	-0.57237	0.33689	6.89661
200	-0.12438	0.38040	-0.48055	0.68052	0.18591	0.30521	0.76565	-0.61609	0.40120	7.24370
500	-0.12442	0.38168	-0.47922	0.68862	0.19558	0.32400	0.80223	-0.58849	0.34867	4.67388
1000	-0.12572	0.38304	-0.47929	0.69087	0.19665	0.35070	0.82496	-0.60386	0.36384	2.80778
2000	-0.12473	0.37819	-0.47756	0.69323	0.19882	0.37466	0.84338	-0.57913	0.38073	1.06270
3000	-0.12231	0.37975	-0.47806	0.69127	0.19867	0.37734	0.83777	-0.56934	0.39006	1.15659
True values	-0.12000	0.38000	-0.48000	0.69000	0.20000	0.38000	0.83000	-0.58000	0.38000	

Table 5. The AM-RLS estimates and their errors with $\sigma^2 = 0.30^2$.

t	θ_1	θ_2	θ_3	a_1	a_2	c_1	c_2	d_1	d_2	δ (%)
100	-0.13468	0.41441	-0.48772	0.67974	0.22184	0.46856	0.92479	-0.39256	0.30949	16.30097
200	-0.12274	0.39420	-0.49719	0.66980	0.22598	0.45302	0.87561	-0.45127	0.35305	10.88550
500	-0.12691	0.38743	-0.48520	0.68273	0.21246	0.41800	0.83272	-0.48805	0.27103	9.98308
1000	-0.12749	0.38904	-0.48208	0.68912	0.20222	0.41265	0.82752	-0.54037	0.30295	6.27103
2000	-0.12541	0.37489	-0.47463	0.69733	0.20209	0.40675	0.83309	-0.54745	0.33522	4.20923
3000	-0.12048	0.37963	-0.47540	0.69201	0.19943	0.39827	0.82606	-0.54801	0.35642	2.96778
True values	-0.12000	0.38000	-0.48000	0.69000	0.20000	0.38000	0.83000	-0.58000	0.38000	

Table 6. The TS-RLS estimates and their errors with $\sigma^2 = 0.30^2$.

t	θ_1	θ_2	θ_3	a_1	a_2	c_1	c_2	d_1	d_2	δ (%)
100	-0.15389	0.41004	-0.49079	0.73757	0.37113	0.47108	0.89002	-0.45436	0.22061	19.89815
200	-0.14361	0.39730	-0.50159	0.70544	0.33245	0.44711	0.85270	-0.49697	0.32589	12.54814
500	-0.13848	0.38961	-0.48935	0.69750	0.27238	0.41228	0.81748	-0.53503	0.28490	9.19497
1000	-0.13321	0.39035	-0.48412	0.69705	0.23919	0.40330	0.81663	-0.57335	0.32451	5.10824
2000	-0.12860	0.37542	-0.47594	0.70165	0.22363	0.39727	0.82614	-0.56842	0.35276	3.00464
3000	-0.12278	0.37982	-0.47631	0.69524	0.21550	0.38913	0.82143	-0.56569	0.37307	1.76754
True values	-0.12000	0.38000	-0.48000	0.69000	0.20000	0.38000	0.83000	-0.58000	0.38000	

Table 7. The AM-RLS parameter estimates and errors based on 20 Monte-Carlo runs ($\sigma^2 = 0.10^2$).

t	100	200	500	1000	2000	3000	True values
θ_1	-0.13255±0.01157	-0.12177±0.00963	-0.12255±0.00657	-0.12095±0.00448	-0.1206±0.00254	-0.12042±0.00199	-0.12000
θ_2	0.39308±0.01076	0.38766±0.00657	0.38439±0.00245	0.38279±0.00153	0.38093±0.00142	0.38069±0.00134	0.38000
θ_3	-0.4802±0.00754	-0.48045±0.00515	-0.47993±0.00423	-0.48052±0.00158	-0.48152±0.00176	-0.48149±0.00138	-0.48000
a_1	0.69361±0.0147	0.68368±0.00851	0.68709±0.00614	0.68754±0.0032	0.68799±0.00328	0.68901±0.00358	-0.69000
a_2	0.22176±0.02362	0.21169±0.01113	0.20661±0.00983	0.20337±0.00533	0.20194±0.00358	0.20126±0.00283	0.20000
c_1	0.48276±0.09851	0.40188±0.08033	0.35477±0.03602	0.34368±0.01618	0.34259±0.01689	0.34949±0.01628	0.38000
c_2	0.67111±0.07796	0.70181±0.06171	0.75194±0.04056	0.79009±0.02173	0.81431±0.01464	0.82397±0.01067	0.83000
d_1	-0.49252±0.1249	-0.53595±0.09961	-0.60018±0.03376	-0.60378±0.01592	-0.60852±0.01827	-0.60323±0.01107	-0.58000
d_2	0.11254±0.18201	0.26719±0.12736	0.33594±0.06823	0.37311±0.03638	0.39095±0.02355	0.39444±0.02083	0.38000
δ (%)	26.50264±9.8533	16.32498±5.49335	8.35766±2.96942	4.95135±2.25232	3.84058±1.73951	3.31727±1.48867	

Table 8. The TS-RLS parameter estimates and errors based on 20 Monte-Carlo runs ($\sigma^2 = 0.10^2$).

t	100	200	500	1000	2000	3000	True values
θ_1	-0.12627±0.00894	-0.11828±0.00757	-0.12176±0.00585	-0.1204±0.00365	-0.12032±0.00194	-0.12037±0.00166	-0.12000
θ_2	0.38106±0.01225	0.38113±0.00712	00.38152±0.00383	0.38138±0.0017	0.38027±0.0017	0.38026±0.00141	0.38000
θ_3	-0.48517±0.00788	-0.48288±0.00512	-0.48125±0.00339	-0.48093±0.00084	-0.48163±0.00138	-0.48153±0.00104	-0.48000
a_1	0.69093±0.01865	0.68236±0.01002	0.68637±0.00667	0.68725±0.00274	0.68765±0.00266	0.68864±0.00289	-0.69000
a_2	0.19651±0.02773	0.19591±0.01208	0.19813±0.00699	0.1979±0.00313	0.19818±0.00179	0.19809±0.00146	0.20000
c_1	0.41457±0.06928	0.40709±0.05503	0.39458±0.02887	0.3846±0.01633	0.37433±0.01079	0.37411±0.00963	0.38000
c_2	0.84468±0.09824	0.82792±0.06215	0.81838±0.01838	0.82218±0.01873	0.8249±0.01272	0.82768±0.00933	0.83000
d_1	-0.53495±0.12166	-0.52304±0.07687	-0.55377±0.0324	-0.56185±0.01896	-0.57595±0.01319	-0.57858±0.00797	-0.58000
d_2	0.34887±0.13309	0.37173±0.09374	0.3613±0.04234	0.36502±0.02236	0.36882±0.01406	0.37275±0.01155	0.38000
δ (%)	14.37789±4.09728	10.15813±2.81647	4.59738±1.5813	2.85029±0.95983	1.76306±0.79823	1.33988±0.54453	

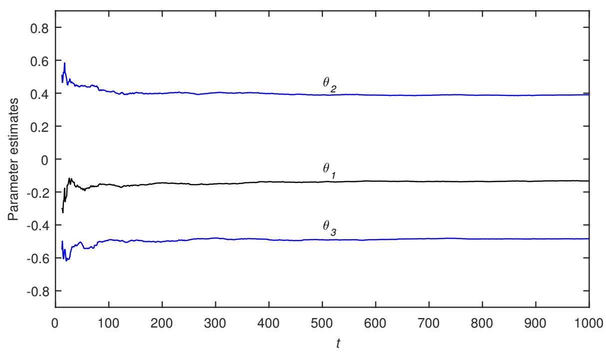


Fig. 3. The TS-RLS estimates θ_1, θ_2, θ_3 versus t with $\sigma^2 = 0.10^2$.

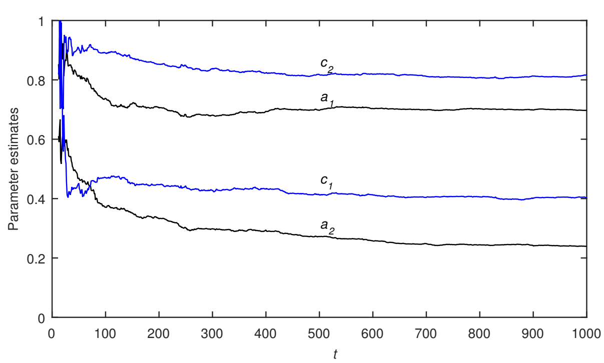


Fig. 4. The TS-RLS estimates a_1, a_2, c_1, c_2 versus t with $\sigma^2 = 0.10^2$.

timation errors than the AM-RLS algorithm - see Figs. 1-2.

6. CONCLUSIONS

The main contribution of this paper is to derive the AM-RLS algorithm and the TS-RLS algorithm for the M-OEARMA systems. The simulation indicates that the proposed TS-RLS algorithm can generate higher accurate parameter estimates compared with the AM-RLS algorithm in MatLab. The proposed methods in this paper can be extended to study the identification problems of other multivariate systems with different structures and disturbance noise. The identification method presented in this paper can be extended to study the parameter estimation algorithms of different systems [41–51] and can be applied to other fields [52–58].

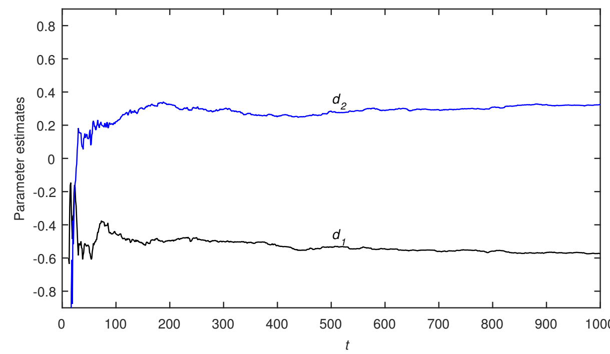


Fig. 5. The TS-RLS estimates d_1, d_2 versus t with $\sigma^2 = 0.10^2$.

REFERENCES

[1] Y. Gu, J. Liu, X. Li, Y. Chou, and Y. Ji, "State space model identification of multirate processes with time-delay using the expectation maximization," *Journal of the Franklin Institute*, vol. 356, no. 3, pp. 1623-1639, February 2019.

[2] Y. Gu, Y. Chou, J. Liu, and Y. Ji, "Moving horizon estimation for multirate systems with time-varying time-delays," *Journal of the Franklin Institute*, vol. 356, no. 4, pp. 2325-2345, March 2019.

[3] Y. Cao, P. Li, and Y. Zhang, "Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing," *Future Generation Computer Systems*, vol. 88, pp. 279-283, November 2018.

[4] Y. Z. Zhang, Y. Cao, Y. H. Wen, L. Liang, and F. Zou, "Optimization of information interaction protocols in cooperative vehicle-infrastructure systems," *Chinese Journal of Electronics*, vol. 27, no. 2, pp. 439-444, March 2018.

[5] Y. Cao, L. C. Ma, S. Xiao, X. Zhang, and W. Xu, "Standard analysis for transfer delay in CTCS-3," *Chinese Journal*

of Electronics, vol. 26, no. 5, pp. 1057-1063, September 2017.

[6] Y. Cao, Y. Wen, X. Meng, and W. Xu, "Performance evaluation with improved receiver design for asynchronous coordinated multipoint transmissions," *Chinese Journal of Electronics*. vol. 25, no. 2, pp. 372-378, March 2016.

[7] F. Liu, "Rough maximal functions supported by subvarieties on Triebel-Lizorkin spaces," *Revista De La Real Academia De Ciencias Exactas Fisicas Y Naturales Serie A-Matematicas*, vol. 112, no. 2, pp. 593-614, April 2018.

[8] F. Liu, Z. Fu, and S. Jhang, "Boundedness and continuity of Marcinkiewicz integrals associated to homogeneous mappings on Triebel-Lizorkin spaces," *Frontiers of Mathematics in China*, vol. 14, no. 1, pp. 95-122, January 2019.

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

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

[11] M. Gan, H. X. Li, and H. Peng, "A variable projection approach for efficient estimation of RBF-ARX model," *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 462-471, March 2015.

[12] A. Brouri, F. Giri, F. Ikhouane, F. Z. Chaoui, and O. Amdouri, "Identification of Hammerstein-Wiener systems with backlask input nonlinearity bordered by straight lines," *Proc. of 19th IFAC World Congress*, vol. 47, no. 3, pp. 475-480, August 2014.

[13] C. P. Yu, L. H. Xie, and Y. C. Soh, "Blind channel and source estimation in networked systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 17, pp. 4611-4626, September 2014.

[14] R. N. Mahia, M. Singh, and D. M. Fulwani, "Identification of optimal set of driver nodes in complex networked systems using region of attraction," *International Journal of Control Automation and Systems*, vol. 16, no. 1, pp. 97-107, February 2018.

[15] L. Xu, "The parameter estimation algorithms based on the dynamical response measurement data," *Advances in Mechanical Engineering*, vol. 9, no. 11, pp. 1-12, November 2017. doi: 10.1177/1687814017730003

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

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

[18] Z. P. Zhou and X. F. Liu, "State and fault estimation of sandwich systems with hysteresis," *International Journal of Robust and Nonlinear Control*, vol. 28, no. 13, pp. 3974-3986, September 2018.

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

[20] A. Brouri, L. Kadi, and S. Slassi, "Frequency identification of Hammerstein-Wiener systems with backlash input nonlinearity," *International Journal of Control Automation and Systems*, vol. 15, no. 5, pp. 2222-2232, October 2017.

[21] A. Brouri, O. Amdouri, F. Z. Chaoui, and F. Giri, "Frequency identification of Hammerstein-Wiener systems with piecewise affine input nonlinearity," *Proc. of 19th IFAC World Congress*, vol. 47, no. 3, pp. 10030-10035, August 2014.

[22] J. Na, J. Yang, X. Wu, and Y. Guo, "Robust adaptive parameter estimation of sinusoidal signals," *Automatica*, vol. 53, pp. 376-384, March 2015.

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

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

[25] J. L. Ding, "Recursive and iterative least squares parameter estimation algorithms for multiple-input-output-error systems with autoregressive noise," *Circuits Systems, and Signal Processing*, vol. 37, no. 5, pp. 1884-1906, May 2018.

[26] X. Zhang, L. Xu, F. Ding, and T. Hayat, "Combined state and parameter estimation for a bilinear state space system with moving average noise," *Journal of the Franklin Institute*, vol. 355, no. 6, pp. 3079-3103, April 2018.

[27] J. Chen, B. Jiang, and J. Li, "Missing output identification model based recursive least squares algorithm for a distributed parameter system," *International Journal of Control Automation and Systems*, vol. 16, no. 1, pp. 150-157, February 2018.

[28] H. Cho and S. C. Yu, "Variable data-window-size recursive least-squares algorithm for dynamic system identification," *Electronics Letters*, vol. 51, no. 4, pp. 341-343, February 2015.

[29] Y. J. Wang and F. Ding, "A filtering based multi-innovation gradient estimation algorithm and performance analysis for nonlinear dynamical systems," *IMA Journal of Applied Mathematics*, vol. 82, no. 6, pp. 1171-1191, November 2017.

[30] X. Zhang, F. Ding, L. Xu, and E. F. Yang, "State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle," *IET Control Theory and Applications*, vol. 12, no. 12, pp. 1704-1713, August 2018.

[31] Z. W. Ge, F. Ding, L. Xu, A. Alsaedi, and T. Hayat, "Gradient-based iterative identification method for multivariate equation-error autoregressive moving average systems using the decomposition technique," *Journal of the Franklin Institute*, vol. 356, no. 3, pp. 1658-1676, February 2019.

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

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

[34] Y. P. Pan, X. Li, and H. Y. Yu, "Least-squares learning control with guaranteed parameter convergence," *Proc. of 22nd International Conference on Automation and Computing (ICAC)*, Colchester, UK, 2016.

[35] Y. P. Pan and H. Y. Yu, "Composite learning from adaptive dynamic surface control," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2603-2609, September 2016.

[36] Q. Y. Liu and F. Ding, "Auxiliary model-based recursive generalized least squares algorithm for multivariate output-error autoregressive systems using the data filtering," *Circuits Systems and Signal Processing*, vol. 38, no. 2, pp. 590-610, February 2019.

[37] T. L. Lai and C. Z. Wei, "least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, vol. 10, no. 1, pp. 154-166, January 1982.

[38] F. Ding, "Coupled-least-squares identification for multivariable systems," *IET Control Theory and Applications*, vol. 7, no. 1, pp. 68-79, January 2013.

[39] Y. J. Wang and F. Ding, "Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model," *Automatica*, vol. 71, pp. 308-313, September 2016.

[40] Y. J. Wang and F. Ding, "The auxiliary model based hierarchical gradient algorithms and convergence analysis using the filtering technique,"*Signal Processing*, vol. 128, pp. 212-221, November 2016.

[41] Y. J. Wang, F. Ding, and L. Xu, "Some new results of designing an IIR filter with colored noise for signal processing," *Digital Signal Processing*, vol. 72, pp. 44-58, January 2018.

[42] Y. Ji and F. Ding, "Multiperiodicity and exponential attractivity of neural networks with mixed delays," *Circuits Systems and Signal Processing*, vol. 36, no. 6, pp. 2558-2573, June 2017.

[43] F. Ding, Y. J. Wang, J. Y. Dai, Q. S Li, and Q. J. Chen, "A recursive least squares parameter estimation algorithm for output nonlinear autoregressive systems using the input-output data filtering," *Journal of the Franklin Institute*, vol. 354, no. 15, pp. 6938-6955, October 2017.

[44] H. B. Chen, Y. S. Xiao, and F. Ding, "Hierarchical gradient parameter estimation algorithm for Hammerstein nonlinear systems using the key term separation principle," *Applied Mathematics and Computation*, vol. 247, pp. 1202-1210, November 2014.

[45] Y. W. Mao and F. Ding, "A novel parameter separation based identification algorithm for Hammerstein systems," *Applied Mathematics Letters*, vol. 60, pp. 21-27, October 2016.

[46] Y. Ji and X. M. Liu, "Unified synchronization criteria for hybrid switching-impulsive dynamical networks," *Circuits Systems Signal Processing*, vol. 34, no. 5, pp. 1499-1517, May 2015.

[47] L. Xu, "A proportional differential control method for a time-delay system using the Taylor expansion approximation," *Applied Mathematics and Computation*, vol. 236, pp. 391-399, June 2015.

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

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

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

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

[52] F. Liu, "Boundedness and continuity of maximal operators associated to polynomial compound curves on Triebel-Lizorkin spaces," *Mathematical Inequalities & Applications*, vol. 22, no. 1, pp. 25-44, January 2019.

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

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

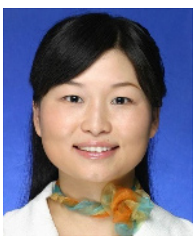
[55] N. Zhao, R. Liu, Y. Chen, M. Wu, Y. Jiang, W. Xiong, and C. Liu, "Contract design for relay incentive mechanism under dual asymmetric information in cooperative networks," *Wireless Networks*, vol. 24, no. 8, pp. 3029-3044, November 2018.

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

[57] Y. Wang, Y. Si, B. Huang, and S. X. Ding, "Survey on the theoretical research and engineering applications of multivariate statistics process monitoring algorithms: 2008-2017," *The Canadian Journal of Chemical Engineering*, vol. 96, no. 10, pp. 2073-2085, October 2018.

[58] X. Y. Li, H. X. Li, and B. Y. Wu, "Piecewise reproducing kernel method for linear impulsive delay differential equations with piecewise constant arguments," *Applied Mathematics and Computation*, vol. 349, pp. 304-313, May 2019.

Yunze Guo was born in Luoyang (Henan Province, China) in 1996. He received his B.Sc. degree from the North China University of Water Resources and Electric Power (Zhengzhou, China) in 2017. He is currently a master student in the School of Internet of Things Engineering, Jiangnan University, (Wuxi, China). His interests include system modeling, system identification and parameter estimation.



Lijuan Wan received her B.Sc. and M.Sc. degrees from Wuhan University, in 2006 and 2008, respectively. She has been a lecturer in the School of Mathematics and Physics at the Qingdao University of Science and Technology (Qingdao, China) since 2008. She is currently a Ph.D. student in the College of Automation and Electronic Engineering, Qingdao University of Science and Technology. Her research interests include system modeling and system identification.

Ling Xu was born in Tianjin, China. She received her Master and Ph.D. degrees from the Jiangnan University (Wuxi, China), in 2005 and 2015, respectively. She has been an Associate Professor since 2015. She is a Colleges and Universities,Blue Project, Young Teacher (Jiangsu, China). Her research interests include process control, parameter estimation and signal modeling.

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

Ahmed Alsaedi obtained his Ph.D. degree from Swansea University (UK) in 2002. He has a broad experience of research in applied mathematics. His fields of interest include dynamical systems, nonlinear analysis involving ordinary differential equations, fractional differential equations, boundary value problems, mathematical modeling, biomathematics, Newtonian and Non-Newtonian fluid mechanics. He has published several articles in peer-reviewed journals. He has supervised several M.S. students and executed many research projects successfully. He is reviewer of several international journals. He served as the chairman of the mathematics department at KAU and presently he is serving as director of the research program at KAU. Under his great leadership, this program is running quite successfully and it has attracted a large number of highly rated researchers and distinguished professors from all over the world. He is also the head of NAAM international research group at KAU.

Tasawar Hayat was born in Khanewal, Punjab, Distinguished National Professor and Chairperson of Mathematics Department at Quaid-I-Azam University is renowned worldwide for his seminal, diversified and fundamental contributions in models relevant to physiological systems, control engineering. He has a honor of being fellow of Pakistan Academy of Sciences, Third World Academy of Sciences (TWAS) and Islamic World Academy of Sciences in the mathematical Sciences.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.