International Journal of Control, Automation, and Systems 21(5) (2023) 1455-1464
http://dx.doi.org/10.1007/s12555-022-0253-y
ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Multivariable CAR-like System Identification with Multi-innovation Gradient and Least Squares Algorithms

Jian Pan*, Huijian Zhang, Hongzhan Guo, Sunde Liu, and Yuqing Liu

Abstract: This paper focuses on the identification of a multivariable controlled autoregressive-like (CAR-like) system. A joint identification algorithm of stochastic gradient and least squares is deduced for estimating the system parameters by decomposing the multivariable CAR-like system into two subsystems, which avoids the calculation of the matrix inversion. To further improve the parameter estimation accuracy, a joint identification algorithm of hierarchical multi-innovation stochastic gradient and least squares is proposed by using the multi-innovation identification theory. The simulation results confirm that these proposed algorithms are effective.

Keywords: Hierarchical identification, least squares, multi-innovation, multivariable system, parameter estimation, stochastic gradient.

1. INTRODUCTION

System identification is a modeling method to minimize the error criterion function by using the input and output data of the system [1-5] and by means of some optimization tools [6-10]. Mathematical models are important for studying natural sciences [11-15]. According to the number of input and output variables, the system can be divided into univariate systems and multivariable systems [16-20]. Compared with the univariate systems, the multivariable systems have complex structures, multiple variables, and large dimensionality. Additionally, there is a coupling relationship between parameters, and these factors result in a heavy computational burden on identification algorithms. With the development of the modern industrial control field, multivariable systems are widely used in various applications, such as UAV flight control and weather forecast. As far as lithium battery is concerned, its actual system is complex and has high coupling characteristics. Therefore, the calculation efficiency of parameter estimation is low. Without rapid parameter identification technology, it is difficult to accurately evaluate the actual safe operation range of lithium batteries, and it is difficult to prevent the occurrence of lithium battery safety accidents. Therefore, improving the calculation efficiency and the parameter estimation accuracy is a hot issue in the above multivariable system identification.

At present, a great deal of research has been done on the identification of the multivariable systems. Katayama *et al*. showed that the best linear approximation of the MIMO Wiener-Hammerstein systems in the mean square sense can be obtained by the orthogonal projection subspace identification method [21]. Multivariable-like systems are a special class of multivariable systems. The identification model of the multivariable-like systems contain parameter vector and parameter matrix, so the traditional least squares algorithm can not directly estimate the parameters of the system. The Kronecker product can convert the parameter vector and the parameter matrix into a large parameter vector, to convert a multivariable-like identification model into a multivariable linear regression identification model [22-25]. However, the dimension of the parameter vector and information matrix of this method is large, and the information matrix contains multiple zero terms, which will cause the problem of the parameter redundancy [26,27].

In the field of multivariable system identification, the hierarchical identification pinciple has also attracted extensive attention recently. The basic idea of hierarchical identification principle is to decompose the original system into multiple subsystems with simple structures and independent of each other, which can reduce the amount of calculation [22-25]. The amount of calculation of the hierarchical stochastic gradient (H-SG) algorithm is small, but the convergence speed and the estimation accuracy are not as good as the H-RLS algorithm. However, the H-RLS algorithm needs to calculate the inverse of the matrix, which produces a large amount of calculation and reduces the calculation efficiency.

Based on the above research, a hierarchical stochas-

Manuscript received March 22, 2022; revised May 26, 2022; accepted June 6, 2022. Recommended by Associate Editor Seongmin Heo under the direction of Senior Editor Jong Min Lee. This work was supported by the National Natural Science Foundation of China (No. 61571182).

Jian Pan, Huijian Zhang, Hongzhan Guo, Sunde Liu, and Yuqing Liu are with the School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China (e-mails: {jpan, zhj19990614, 15098015305}@163.com, {sdliu, liuyuqing}@hbut.edu.cn).
* Corresponding author.

tic gradient and least squares identification algorithm (H-SG-RLS) is proposed to alternately estimate the system parameter matrix and the parameter vector of the multivariable controlled autoregressive-like (CAR-like) system. The main contributions of this paper are listed as follows:

1) This paper decomposes the multivariable CAR-like system into two subsystems using the hierarchical identification principle.
2) The proposed H-SG-RLS algorithm avoids matrix inversion. Compared with the H-RLS algorithm, the parameter estimation accuracy is close, but the amount of calculation is lower.
3) To further improve the accuracy of the identification algorithm, a hierarchical multi-innovation stochastic gradient least squares algorithm (H-MI-SG-RLS) is used for the parameter estimation.

The rest of the paper is organized as follows: In Section 2, the multivariable CAR-like system is decomposed into two sub-identification models. Section 3 and Section 4 propose a H-SG-RLS algorithm and a H-MI-SG-RLS algorithm. An example is given in Section 5 to illustrate the effectiveness of the proposed algorithms. Finally, a conclusion is given in Section 6.

2. SYSTEM DESCRIPTION AND IDENTIFICATION MODEL

Let us introduce some notations first. The symbol $\boldsymbol{I}_n$ stands for an identity matrix of order n and $\boldsymbol{I}$ is an identity matrix of appropriate sizes; $\mathbf{1}_n$ (or $\mathbf{1}_{m\times n}$) represents an n-dimensional column vector (or an $m\times n$ dimensional matrix) whose entries are $\mathbf{1}$; the superscript T denotes the matrix transpose; the norm of a matrix $\boldsymbol{X}$ is defined by $\|\boldsymbol{X}\|^2 := \mathrm{tr}[\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}]$.

Consider a multivariable CAR-like system,

$$\alpha(z)\boldsymbol{y}(k) = \boldsymbol{Q}(z)\boldsymbol{u}(k) + \boldsymbol{v}(k), \tag{1}$$

where $\{\boldsymbol{u}(k)\} \in \mathbb{R}^r$ and $\{\boldsymbol{y}(k)\} \in \mathbb{R}^m$ are the input and output vectors of the system, respectively, $\{\boldsymbol{v}(k)\} \in \mathbb{R}^m$ is the stochastic noise vector with zero mean, $\alpha(z)$ is the system characteristic polynomial in the unit backward shift operator z^{-1} $[z^{-1}y(k) = y(k-1)]$, $\boldsymbol{Q}(z)$ is a matrix polynomial in z^{-1}, which are defined as

$$\alpha(z) := 1 + \alpha_1 z^{-1} + \cdots + \alpha_{n_\alpha} z^{-n_\alpha},\ \alpha_i \in \mathbb{R}, \tag{2}$$

$$\boldsymbol{Q}(z) := \boldsymbol{Q}_1 z^{-1} + \cdots + \boldsymbol{Q}_{n_q} z^{-n_q},\ \boldsymbol{Q}_i \in \mathbb{R}^{m\times r}. \tag{3}$$

Bring (2) and (3) into (1) to obtain

$$\begin{aligned}&(1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \cdots + \alpha_{n_\alpha} z^{-n_\alpha})\boldsymbol{y}(k)\\ &= (\boldsymbol{Q}_1 z^{-1} + \boldsymbol{Q}_2 z^{-2} + \cdots + \boldsymbol{Q}_{n_q} z^{-n_q})\boldsymbol{u}(k) + \boldsymbol{v}(k).\end{aligned} \tag{4}$$

Then, (4) can be transformed into

$$\boldsymbol{y}(k) = -\sum_{i=1}^{n_\alpha} \alpha_i \boldsymbol{y}(k-i) + \sum_{i=1}^{n_q} \boldsymbol{Q}_i \boldsymbol{u}(k-i) + \boldsymbol{v}(k). \tag{5}$$

Define the parameter vector $\boldsymbol{\alpha}$, the parameter matrix $\boldsymbol{\theta}$, and the input information vector $\boldsymbol{\varphi}(k)$ and the information matrix $\boldsymbol{\phi}(k)$ as

$$\begin{aligned}
\boldsymbol{\alpha} &:= [\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_{n_\alpha}]^{\mathrm{T}} \in \mathbb{R}^{n_\alpha},\\
\boldsymbol{\theta}^{\mathrm{T}} &:= [\boldsymbol{Q}_1,\ \boldsymbol{Q}_2,\ \cdots,\ \boldsymbol{Q}_{n_q}] \in \mathbb{R}^{m\times n_q r},\\
\boldsymbol{\varphi}(k) &:= [\boldsymbol{u}^{\mathrm{T}}(k-1),\ \boldsymbol{u}^{\mathrm{T}}(k-2),\ \cdots,\ \boldsymbol{u}(k-n_q)^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n_q r},\\
\boldsymbol{\phi}(k) &:= [\boldsymbol{y}(k-1),\ \boldsymbol{y}(k-2),\ \cdots,\ \boldsymbol{y}(k-n_\alpha)] \in \mathbb{R}^{m\times n_\alpha}.
\end{aligned}$$

Then, (1) may be rewritten as

$$\boldsymbol{y}(k) = -\boldsymbol{\phi}(k)\boldsymbol{\alpha} + \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(k) + \boldsymbol{v}(k). \tag{6}$$

Equation (6) is the identification model of the multivariable CAR-like system of (1), and contains a parameter vector $\boldsymbol{\alpha}$ which consists of the coefficients of polynomials, and a parameter matrix $\boldsymbol{\theta}$ which consists of the coefficients of the matrix polynomial. Introduce two intermediate variables, which are defined as

$$\boldsymbol{y}_1(k) := \boldsymbol{y}(k) - \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(k),\ \boldsymbol{y}_2(k) := \boldsymbol{y}(k) + \boldsymbol{\phi}(k)\boldsymbol{\alpha}. \tag{7}$$

According to the hierarchical identification principle, (6) can be decomposed into two fictitious subsystems

$$\boldsymbol{y}_1(k) = -\boldsymbol{\phi}(k)\boldsymbol{\alpha} + \boldsymbol{v}(k), \tag{8}$$

$$\boldsymbol{y}_2(k) = \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(k) + \boldsymbol{v}(k). \tag{9}$$

The proposed parameter estimation algorithms in this paper are based on the two sub-identification models in (8) and (9). Many identification methods are derived based on the identification models of the systems [28-34] and these methods can be used to estimate the parameters of other linear systems and nonlinear systems [35-38] and can be applied to other fields [39-44].

3. THE H-SG-RLS ALGORITHM

The recursive algorithms update the estimates by using new observations at each recursion [45-48] and the iterative algorithms update the estimates by using a fixed batch of observations [49-53]. Here uses the hierarchical stochastic gradient and least squares identification algorithm for the parameter estimation.

Refer to the works in [22-25] and based on the identification models in (8) and (9), define two quadratic criterion functions,

$$J_1(\boldsymbol{\alpha}) := \sum_{j=1}^{k} [\boldsymbol{y}_1(j) + \boldsymbol{\phi}(j)\boldsymbol{\alpha}]^2,$$

$$J_2(\boldsymbol{\theta}) := \sum_{j=1}^{k}[\boldsymbol{y}_2(j)-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(j)]^2.$$

Let $\hat{\boldsymbol{\alpha}}(k)$ and $\hat{\boldsymbol{\theta}}(k)$ be the estimations of parameter vector $\boldsymbol{\alpha}$ and parameter matrix $\boldsymbol{\theta}$ at time k, respectively. Referring to the hierarchical identification principle, minimizing the quadratic standard functions $J_1(\boldsymbol{\alpha})$ and $J_2(\boldsymbol{\theta})$, we have the following least squares algorithm [22-25]:

$$\hat{\boldsymbol{\alpha}}(k) = \hat{\boldsymbol{\alpha}}(k-1)+\boldsymbol{L}_1(k) \times [\boldsymbol{y}(k)+\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1)-\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\varphi}(k)], \tag{10}$$

$$\boldsymbol{L}_1(k) = -\boldsymbol{P}_1(k-1)\boldsymbol{\phi}^{\mathrm{T}}(k) \times [\boldsymbol{I}+\boldsymbol{\phi}(k)\boldsymbol{P}_1(k-1)\boldsymbol{\phi}^{\mathrm{T}}(k)]^{-1}, \tag{11}$$

$$\boldsymbol{P}_1(k) = [\boldsymbol{I}+\boldsymbol{L}_1(k)\boldsymbol{\phi}(k)]\boldsymbol{P}_1(k-1), \tag{12}$$

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}_2(k) \times [\boldsymbol{y}(k)+\boldsymbol{\phi}(k)\boldsymbol{\alpha}-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)]^{\mathrm{T}}, \tag{13}$$

$$\boldsymbol{L}_2(k) = \frac{\boldsymbol{P}_2(k-1)\boldsymbol{\varphi}(k)}{[1+\boldsymbol{\varphi}^{\mathrm{T}}(k)\boldsymbol{P}_2(k-1)\boldsymbol{\varphi}(k)]}, \tag{14}$$

$$\boldsymbol{P}_2(k) = [\boldsymbol{I}-\boldsymbol{L}_2(k)\boldsymbol{\varphi}^{\mathrm{T}}(k)]\boldsymbol{P}_2(k-1). \tag{15}$$

However, the right-hand sides of (10) and (13) contain the unknown parameter matrices $\boldsymbol{\theta}$ and unknown parameter vector $\boldsymbol{\alpha}$ so the algorithm in (10)-(15) cannot be implemented. Replacing $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ with their estimates $\hat{\boldsymbol{\theta}}(k-1)$ and $\hat{\boldsymbol{\alpha}}(k)$, we can obtain the hierarchical least squares (H-RLS) algorithm for estimating $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$

$$\hat{\boldsymbol{\alpha}}(k) = \hat{\boldsymbol{\alpha}}(k-1)+\boldsymbol{L}_1(k)[\boldsymbol{y}(k) +\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)], \tag{16}$$

$$\boldsymbol{L}_1(k) = -\boldsymbol{P}_1(k-1)\boldsymbol{\phi}^{\mathrm{T}}(k)[\boldsymbol{I}+\boldsymbol{\phi}(k)\boldsymbol{P}_1(k-1)\boldsymbol{\phi}^{\mathrm{T}}(k)]^{-1}, \tag{17}$$

$$\boldsymbol{P}_1(k) = [\boldsymbol{I}+\boldsymbol{L}_1(k)\boldsymbol{\phi}(k)]\boldsymbol{P}_1(k-1), \tag{18}$$

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}_2(k)[\boldsymbol{y}(k) +\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)]^{\mathrm{T}}, \tag{19}$$

$$\boldsymbol{L}_2(k) = \frac{\boldsymbol{P}_2(k-1)\boldsymbol{\varphi}(k)}{[1+\boldsymbol{\varphi}^{\mathrm{T}}(k)\boldsymbol{P}_2(k-1)\boldsymbol{\varphi}(k)]}, \tag{20}$$

$$\boldsymbol{P}_2(k) = [\boldsymbol{I}-\boldsymbol{L}_2(k)\boldsymbol{\varphi}^{\mathrm{T}}(k)]\boldsymbol{P}_2(k-1), \tag{21}$$

$$\boldsymbol{\phi}(k) = [\boldsymbol{y}(k-1),\boldsymbol{y}(k-2),\cdots,\boldsymbol{y}(k-n_\alpha)], \tag{22}$$

$$\boldsymbol{\varphi}(k) = [\boldsymbol{u}^{\mathrm{T}}(k-1),\ \boldsymbol{u}^{\mathrm{T}}(k-2),\ \cdots,\ \boldsymbol{u}^{\mathrm{T}}(k-n_q)]^{\mathrm{T}}. \tag{23}$$

It can be seen from the above equations that the inverse matrix needs to be calculated when using the least squares algorithm to identify parameter vector $\boldsymbol{\alpha}$. When the number of inputs is large, massive calculation calculations will be generated. So we use the stochastic gradient algorithm instead of the least squares algorithm to identify the parameter vector $\boldsymbol{\alpha}$. In this way, the calculation of gain matrix $\boldsymbol{L}_1(k)$ can be avoided, to avoid the inverse operation of calculation matrix. Thus, we can get the hierarchical stochastic gradient and least squares identification (H-SG-RLS) algorithm for estimating $\boldsymbol{\alpha}$ and $\boldsymbol{\theta}$ [22-25]

$$\hat{\boldsymbol{\alpha}}(k) = \hat{\boldsymbol{\alpha}}(k-1)-\frac{\boldsymbol{\phi}^{\mathrm{T}}(k)}{r_1(k)}[\boldsymbol{y}(k) +\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)], \tag{24}$$

$$r(k) = r(k-1)+\|\boldsymbol{\phi}(k)\|^2, \tag{25}$$

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}(k)[\boldsymbol{y}(k) +\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)]^{\mathrm{T}}, \tag{26}$$

$$\boldsymbol{L}(k) = \frac{\boldsymbol{P}(k-1)\boldsymbol{\varphi}(k)}{[1+\boldsymbol{\varphi}^{\mathrm{T}}(k)\boldsymbol{P}(k-1)\boldsymbol{\varphi}(k)]}, \tag{27}$$

$$\boldsymbol{P}(k) = [\boldsymbol{I}-\boldsymbol{L}(k)\boldsymbol{\varphi}^{\mathrm{T}}(k)]\boldsymbol{P}(k-1), \tag{28}$$

$$\boldsymbol{\phi}(k) = [\boldsymbol{y}(k-1),\boldsymbol{y}(k-2),\cdots,\boldsymbol{y}(k-n_\alpha)], \tag{29}$$

$$\boldsymbol{\varphi}(k) = [\boldsymbol{u}^{\mathrm{T}}(k-1),\boldsymbol{u}^{\mathrm{T}}(k-2),\cdots,\boldsymbol{u}^{\mathrm{T}}(k-n_q)]^{\mathrm{T}}. \tag{30}$$

The proposed algorithms in this paper can combine other methods to study parameter identification of different systems [54-59] and can be applied to other control and schedule areas [60-65] such as information processing and industrial process systems and so on. The procedures of computing $\hat{\boldsymbol{\alpha}}(k)$ and $\hat{\boldsymbol{\theta}}(k)$ in the H-SG-RLS algorithm are listed as follows:

1) Let $k=1$, set the initial values $\hat{\boldsymbol{\alpha}}(0)=I_{n_\alpha}/p_0$, $\hat{\boldsymbol{\theta}}(0)=I_{(n_q r)\times m}/p_0$, $r(0)=1$, $p_0=10^6$, $\boldsymbol{u}(i)=0$, $\boldsymbol{y}(i)=0$ as $i\leq 0$.
2) Collect the input-output data $\boldsymbol{u}(k)$ and $\boldsymbol{y}(k)$, form $\boldsymbol{\phi}(k)$ by (29) and $\boldsymbol{\varphi}(k)$ by (30).
3) Compute $r(k)$, $\boldsymbol{L}(k)$, and $\boldsymbol{P}(k)$ by (25), (27) and (28), respectively.
4) Update the parameter estimates $\hat{\boldsymbol{\alpha}}(k)$ and $\hat{\boldsymbol{\theta}}(k)$ by (24) and (26), respectively.
5) Increase k by 1 and go to Step 2.

Tables 1 and 2 show the computational efficiency of the H-RLS algorithm and the H-SG-RLS algorithm, respectively. When the number of inputs m is large, the calculation amount of the H-SG-RLS algorithm is much lower than that of the H-RLS algorithm.

4. THE H-MI-SG-RLS ALGORITHM

The multi-innovation identification theory can extract more useful information from the observation data to improve the parameter estimation accuracy. Based on the H-SG-RLS algorithm in (24) to (30), the innovation vector in (24) and (26) are

$$\boldsymbol{e}_1(k) := [\boldsymbol{y}(k)+\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m, \tag{31}$$

$$\boldsymbol{e}_2(k) := [\boldsymbol{y}(k)+\boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k)-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m. \tag{32}$$

According to the multi-innovation identification theory, we introduce the innovation length p to expand the vector

Table 1. The computational efficiency of the H-RLS algorithm.

Expressions	Multiplications	Additions
$\hat{\boldsymbol{\alpha}}(k) = \hat{\boldsymbol{\alpha}}(k-1) + \boldsymbol{L}_1(k)[\hat{\boldsymbol{e}}_1(k)] \in \mathbb{R}^{n_\alpha}$	mn_α	mn_α
$\hat{\boldsymbol{e}}_1(k) := [\boldsymbol{y}(k) + \boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1) - \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m$	$m(n_\alpha + n_q r)$	$m(n_\alpha + n_q r)$
$\boldsymbol{L}_1(k) = -\boldsymbol{H}_1(k)\boldsymbol{R}_1'(k) \in \mathbb{R}^{n_\alpha \times m}$	$m^2 n_\alpha$	$m^2 n_\alpha - mn_\alpha$
$\boldsymbol{H}_1(k) = \boldsymbol{P}_1(k-1)\boldsymbol{\phi}^{\mathrm{T}}(k) \in \mathbb{R}^{n_\alpha \times m}$	mn_α^2	$mn_\alpha^2 - mn_\alpha$
$\boldsymbol{R}_1(k) = \boldsymbol{I} + \boldsymbol{\phi}(k)\boldsymbol{H}_1(k) \in \mathbb{R}^{m \times m}$	$m^2 n_\alpha$	$m^2 n_\alpha - mn_\alpha$
$\boldsymbol{R}_1'(k) = \boldsymbol{R}_1^{-1}(k) \in \mathbb{R}^{m \times m}$	m^3	$m^3 - m^2$
$\boldsymbol{P}_1(k) = \boldsymbol{P}_1(k-1) + \boldsymbol{L}_1(k)\boldsymbol{H}_1^{\mathrm{T}}(k) \in \mathbb{R}^{n_\alpha \times n_\alpha}$	mn_α^2	mn_α^2
$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \boldsymbol{L}_2(k)[\hat{\boldsymbol{e}}_2(k)]^{\mathrm{T}} \in \mathbb{R}^{n_q r \times m}$	$mn_q r$	$mn_q r$
$\hat{\boldsymbol{e}}_2(k) := [\boldsymbol{y}(k) + \boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k) - \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m$	$m(n_\alpha + n_q r)$	$m(n_\alpha + n_q r)$
$\boldsymbol{L}_2(k) = \frac{\boldsymbol{H}_2(k)}{\boldsymbol{R}_2(k)} \in \mathbb{R}^{n_q r}$	$n_q r$	0
$\boldsymbol{H}_2(k) = \boldsymbol{P}_2(k-1)\boldsymbol{\varphi}(k) \in \mathbb{R}^{n_q r}$	$(n_q r)^2$	$n_q r(n_q r - 1)$
$\boldsymbol{R}_2(k) = 1 + \boldsymbol{\varphi}^{\mathrm{T}}(k)\boldsymbol{H}_2(k) \in \mathbb{R}$	$n_q r$	$n_q r$
$\boldsymbol{P}_2(k) = \boldsymbol{P}_2(k-1) + \boldsymbol{L}_2(k)\boldsymbol{H}_2^{\mathrm{T}}(k) \in \mathbb{R}^{(n_q r)\times(n_q r)}$	$(n_q r)^2$	$(n_q r)^2$
Sum	$m^3 + 2m^2 n_\alpha + 2mn_\alpha^2 + 2(n_q r)^2 + 3mn_\alpha + 3mn_q r + 2n_q r$	$m^3 - m^2 + 2m^2 n_\alpha + 2mn_\alpha^2 + 2(n_q r)^2 + 2mn_q r$
Total flops	$2m^3 - m^2 + 4m^2 n_\alpha + 4mn_\alpha^2 + 4(n_q r)^2 + 3mn_\alpha + 5mn_q r + 2n_q r$	

Table 2. The computational efficiency of the H-SG-RLS algorithm.

Expressions	Multiplications	Additions
$\hat{\boldsymbol{\alpha}}(k) = \hat{\boldsymbol{\alpha}}(k-1) - \frac{\boldsymbol{\phi}^{\mathrm{T}}(k)}{r_1(k)}[\hat{\boldsymbol{e}}_1(k)] \in \mathbb{R}^{n_\alpha}$	$mn_\alpha + n_\alpha$	mn_α
$\hat{\boldsymbol{e}}_1(k) := [\boldsymbol{y}(k) + \boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k-1) - \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m$	$m(n_\alpha + n_q r)$	$m(n_\alpha + n_q r)$
$r(k) = r(k-1) + \lVert\boldsymbol{\phi}(k)\rVert^2 \in \mathbb{R}$	mn_α	mn_α
$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \boldsymbol{L}(t)[\hat{\boldsymbol{e}}_2(k)]^{\mathrm{T}} \in \mathbb{R}^{n_q r \times m}$	$mn_q r$	$mn_q r$
$\hat{\boldsymbol{e}}_2(k) := [\boldsymbol{y}(k) + \boldsymbol{\phi}(k)\hat{\boldsymbol{\alpha}}(k) - \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\varphi}(k)] \in \mathbb{R}^m$	$m(n_\alpha + n_q r)$	$m(n_\alpha + n_q r)$
$\boldsymbol{L}(k) = \frac{\boldsymbol{H}(k)}{\boldsymbol{R}(k)} \in \mathbb{R}^{n_q r}$	$n_q r$	0
$\boldsymbol{H}(k) = \boldsymbol{P}(k-1)\boldsymbol{\varphi}(k) \in \mathbb{R}^{n_q r}$	$(n_q r)^2$	$n_q r(n_q r - 1)$
$\boldsymbol{R}(k) = 1 + \boldsymbol{\varphi}^{\mathrm{T}}(k)\boldsymbol{H}(k) \in \mathbb{R}$	$n_q r$	$n_q r$
$\boldsymbol{P}(k) = \boldsymbol{P}(k-1) + \boldsymbol{L}(k)\boldsymbol{H}^{\mathrm{T}}(k) \in \mathbb{R}^{n_q r \times n_q r}$	$(n_q r)^2$	$(n_q r)^2$
Sum	$2(n_q r)^2 + 4mn_\alpha + 3mn_q r + 2n_q r + n_\alpha$	$2(n_q r)^2 + 4mn_\alpha + 3mn_q r$
Total flops	$4(n_q r)^2 + 8mn_\alpha + 6mn_q r + 2n_q r + n_\alpha$	

innovations $\boldsymbol{e}_1(k)$ and $\boldsymbol{e}_2(k)$ to a larger innovation vector $\boldsymbol{E}_1(p,k)$ and a large innovation matrix $\boldsymbol{E}_2(p,k)$ as

$$\begin{aligned}\boldsymbol{E}_1(p,k) &:= \boldsymbol{Y}_1(p,k) + \boldsymbol{\Phi}_1(p,k)\hat{\boldsymbol{\alpha}}(k-1) \\ &\quad - (\boldsymbol{I}_p \otimes \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1))\boldsymbol{\Psi}_1(p,k), \qquad (33)\\ \boldsymbol{E}_2(p,k) &:= \boldsymbol{Y}_2(p,k) + \boldsymbol{\Phi}_2^{\mathrm{T}}(p,k)(\boldsymbol{I}_p \otimes \hat{\boldsymbol{\alpha}}(k)) \\ &\quad - \hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\Psi}_2(p,k), \qquad (34)\end{aligned}$$

where the stacked output vector $\boldsymbol{Y}_1(p,k)$, the stacked output matrix $\boldsymbol{Y}_2(p,k)$, the stacked information vector $\boldsymbol{\Phi}_1(p,k)$ and the stacked information matrices $\boldsymbol{\Phi}_2(p,k)$, $\boldsymbol{\Psi}_1(p,k)$ and $\boldsymbol{\Psi}_2(p,k)$ are defined as

$$\begin{aligned}\boldsymbol{Y}_1(p,k) &:= [\boldsymbol{y}^{\mathrm{T}}(k), \boldsymbol{y}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{y}^{\mathrm{T}}(k-p+1)]^{\mathrm{T}},\\ \boldsymbol{Y}_2(p,k) &:= [\boldsymbol{y}(k), \boldsymbol{y}(k-1), \cdots, \boldsymbol{y}(k-p+1)],\\ \boldsymbol{\Phi}_1(p,k) &:= [\boldsymbol{\phi}^{\mathrm{T}}(k), \boldsymbol{\phi}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{\phi}^{\mathrm{T}}(k-p+1)]^{\mathrm{T}},\\ \boldsymbol{\Psi}_1(p,k) &:= [\boldsymbol{\varphi}^{\mathrm{T}}(k), \boldsymbol{\varphi}^{\mathrm{T}}(k-1), \cdots, \boldsymbol{\varphi}^{\mathrm{T}}(k-p+1)]^{\mathrm{T}},\\ \boldsymbol{\Phi}_2(p,k) &:= [\boldsymbol{\phi}(k), \boldsymbol{\phi}(k-1), \cdots, \boldsymbol{\phi}(k-p+1)],\\ \boldsymbol{\Psi}_2(p,k) &:= [\boldsymbol{\varphi}(k), \boldsymbol{\varphi}(k-1), \cdots, \boldsymbol{\varphi}(k-p+1)].\end{aligned}$$

Table 3. The computational efficiency of the H-MI-SG-RLS algorithm.

Expressions	Multiplications	Additions
$\hat{\boldsymbol{\alpha}}(k)=\hat{\boldsymbol{\alpha}}(k-1)-\frac{\boldsymbol{\Phi}_1^{\mathrm{T}}(p,k)}{r(k)}[\boldsymbol{E}_1(p,k)]\in\mathbb{R}^{n_\alpha}$	$mn_\alpha p+n_\alpha$	$mn_\alpha p$
$\boldsymbol{E}_1(p,k):=[\boldsymbol{Y}_1(p,k)+\boldsymbol{\Phi}_1(p,k)\hat{\boldsymbol{\alpha}}(k-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1))\boldsymbol{\Psi}_1(p,k)]\in\mathbb{R}^{mp}$	$mp(n_\alpha+n_q rp)$	$mp(n_\alpha+n_q rp)$
$r(k)=r(k-1)+\|\boldsymbol{\Phi}_1(p,k)\|^2\in\mathbb{R}$	$mn_\alpha p$	$mn_\alpha p$
$\hat{\boldsymbol{\theta}}(k)=\hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}(k)[\boldsymbol{E}_2(p,k)]^{\mathrm{T}}\in\mathbb{R}^{(n_q r)\times m}$	$mn_q rp$	$mn_q rp$
$\boldsymbol{E}_2(p,k):=[\boldsymbol{Y}_2(p,k)+\boldsymbol{\Phi}_2(p,k)(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(k))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\Psi}_2(p,k)]\in\mathbb{R}^{m\times p}$	$mp(n_\alpha p+n_q r)$	$mp(n_\alpha p+n_q r)$
$\boldsymbol{L}(k)=\boldsymbol{H}(k)\boldsymbol{R}'(k)\in\mathbb{R}^{n_q r\times p}$	$n_q rp^2$	$n_q rp(p-1)$
$\boldsymbol{H}(k)=\boldsymbol{P}(k-1)\boldsymbol{\Psi}_2(p,k)\in\mathbb{R}^{n_q r\times p}$	$(n_q r)^2 p$	$n_q rp(n_q r-1)$
$\boldsymbol{R}(k)=\boldsymbol{I}_p+\boldsymbol{\Psi}_2^{\mathrm{T}}(k)\boldsymbol{H}(k)\in\mathbb{R}^{p\times p}$	$n_q rp^2$	$n_q rp^2$
$\boldsymbol{R}'(k)=\boldsymbol{R}^{-1}(k)\in\mathbb{R}^{p\times p}$	p^3	p^3-p^2
$\boldsymbol{P}(k)=\boldsymbol{P}(k-1)+\boldsymbol{L}(k)\boldsymbol{H}^{\mathrm{T}}(k)\in\mathbb{R}^{(n_q r)\times(n_q r)}$	$(n_q r)^2 p$	$(n_q r)^2 p$
Multiplications sum	$p^3+p^2(2n_q r+mn_\alpha+mn_q r)+p[2(n_q r)^2+2mn_q r+3mn_\alpha]+n_\alpha$	
Additions sum	$p^3+p^2(2n_q r+mn_\alpha+mn_q r-1)+p[2(n_q r)^2+2mn_q r+3mn_\alpha-2n_q r]$	
Total flops	$2p^3+p^2(4n_q r+2mn_\alpha+2mn_q r-1)+p[4(n_q r)^2+4mn_q r+6mn_\alpha-2n_q r]+n_\alpha$	

Thus, we can obtain the following multi-innovation based stochastic gradient and least squares (H-MI-SG-RLS) algorithm [22,23]

$$\hat{\boldsymbol{\alpha}}(k)=\hat{\boldsymbol{\alpha}}(k-1)-\frac{\boldsymbol{\Phi}_1^{\mathrm{T}}(p,k)}{r_1(k)}[\boldsymbol{Y}_1(p,k)+\boldsymbol{\Phi}_1(p,k)\times\hat{\boldsymbol{\alpha}}(k-1)-(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1))\boldsymbol{\Psi}_1(p,k)],\tag{35}$$

$$r(k)=r(k-1)+\|\boldsymbol{\Phi}_1(p,k)\|^2, r(0)=1,\tag{36}$$

$$\hat{\boldsymbol{\theta}}(k)=\hat{\boldsymbol{\theta}}(k-1)+\boldsymbol{L}(k)[\boldsymbol{Y}_2(k)+\boldsymbol{\Phi}_2^{\mathrm{T}}(p,k)\times(\boldsymbol{I}_p\otimes\hat{\boldsymbol{\alpha}}(k))-\hat{\boldsymbol{\theta}}^{\mathrm{T}}(k-1)\boldsymbol{\Psi}_2(p,k)]^{\mathrm{T}},\tag{37}$$

$$\boldsymbol{L}(k)=\boldsymbol{P}(k-1)\boldsymbol{\Psi}_2(p,k)[\boldsymbol{I}_p+\boldsymbol{\Psi}_2^{\mathrm{T}}(p,k)\boldsymbol{P}(k-1)\boldsymbol{\Psi}_2(p,k)]^{-1},\tag{38}$$

$$\boldsymbol{P}(k)=[\boldsymbol{I}-\boldsymbol{L}(k)\boldsymbol{\Psi}_2^{\mathrm{T}}(p,k)]\boldsymbol{P}(k-1),\tag{39}$$

$$\boldsymbol{\phi}(k)=[\boldsymbol{y}(k-1),\ \boldsymbol{y}(k-2),\ \cdots,\ \boldsymbol{y}(k-n_\alpha)],\tag{40}$$

$$\boldsymbol{\varphi}(k)=[\boldsymbol{u}^{\mathrm{T}}(k-1),\ \boldsymbol{u}^{\mathrm{T}}(k-2),\ \cdots,\ \boldsymbol{u}^{\mathrm{T}}(k-n_q)]^{\mathrm{T}},\tag{41}$$

$$\boldsymbol{Y}_1(p,k)=[\boldsymbol{y}^{\mathrm{T}}(k),\ \boldsymbol{y}^{\mathrm{T}}(k-1),\ \cdots,\ \boldsymbol{y}^{\mathrm{T}}(k-p+1)],\tag{42}$$

$$\boldsymbol{\Phi}_1(p,k)=[\boldsymbol{\phi}^{\mathrm{T}}(k),\ \boldsymbol{\phi}^{\mathrm{T}}(k-1),\ \cdots,\ \boldsymbol{\phi}^{\mathrm{T}}(k-p+1)],\tag{43}$$

$$\boldsymbol{\Psi}_1(p,k)=[\boldsymbol{\varphi}^{\mathrm{T}}(k),\ \boldsymbol{\varphi}^{\mathrm{T}}(k-1),\ \cdots,\ \boldsymbol{\varphi}^{\mathrm{T}}(k-p+1)],\tag{44}$$

$$\boldsymbol{Y}_2(p,k)=[\boldsymbol{y}(k),\ \boldsymbol{y}(k-1),\ \cdots,\ \boldsymbol{y}(k-p+1)],\tag{45}$$

$$\boldsymbol{\Phi}_2(p,k)=[\boldsymbol{\phi}(k),\ \boldsymbol{\phi}(k-1),\ \cdots,\ \boldsymbol{\phi}(k-p+1)],\tag{46}$$

$$\boldsymbol{\Psi}_2(p,k)=[\boldsymbol{\varphi}(k),\ \boldsymbol{\varphi}(k-1),\ \cdots,\ \boldsymbol{\varphi}(k-p+1)].\tag{47}$$

By extending the innovation vectors $\boldsymbol{e}_1(k)$ and $\boldsymbol{e}_2(k)$ in the H-SG-RLS algorithm to a large innovation vector $\boldsymbol{E}_1(p,k)$ and an innovation matrix $\boldsymbol{E}_2(p,k)$ in the H-MI-SG-RLS algorithm, respectively. The procedures of computing $\hat{\boldsymbol{\alpha}}(t)$ and $\hat{\boldsymbol{\theta}}(t)$ in the H-MI-SG-RLS algorithm are listed as follows:

1) Let $k=1$, set the initial values $\boldsymbol{P}(0)=p_0\boldsymbol{I}_{n_q r}$, $\hat{\boldsymbol{\alpha}}(0)=I_{n_\alpha}/p_0$, $\hat{\boldsymbol{\theta}}(0)=I_{n_q r\times m}/p_0$, $r(0)=1$, $p_0=10^6$, $\boldsymbol{u}(i)=0$, $\boldsymbol{y}(i)=0$ as $i\le 0$.
2) Collect the input-output data $\boldsymbol{u}(k)$ and $\boldsymbol{y}(k)$, construct the information vector $\boldsymbol{\phi}(k)$ and the information matrix $\boldsymbol{\varphi}(k)$ using (40) and (41).
3) Construct stacked output vector $\boldsymbol{Y}_1(p,k)$ and stacked output matrix $\boldsymbol{Y}_2(p,k)$ by (42) and (45), construct stacked information vector $\boldsymbol{\Psi}_1(p,k)$ and stacked information matrix $\boldsymbol{\Phi}_1(p,k)$, $\boldsymbol{\Phi}_2(p,k)$ and $\boldsymbol{\Psi}_2(p,k)$ by (43) to (47).
4) Compute $r(k)$, $\boldsymbol{L}(k)$, and $\boldsymbol{P}(k)$ by (36), (38) and (39), respectively.
5) Update the parameter estimates $\hat{\boldsymbol{\alpha}}(k)$ and $\hat{\boldsymbol{\theta}}(k)$ by (35) and (37), respectively.
6) Increase k by 1 and go to Step 2.

Table 3 shows the computational efficiency of the H-MI-SG-RLS algorithm. The amount of computation of the H-MI-SG-RLS algorithm will increase with the increase of the innovation length p. When $p=1$, the H-MI-SG-RLS algorithm will degenerate into the H-SG-RLS algorithm.

5. EXAMPLE

Consider a two-input two-output CAR-like model,

$$\alpha(z)\boldsymbol{y}(k)=\boldsymbol{Q}(z)\boldsymbol{u}(k)+\boldsymbol{v}(k),$$

Table 4. The H-SG, H-RLS and H-SG-RLS estimates and errors ($\sigma^2 = 0.1^2$).

Algorithms	k	$\hat{\alpha}_1(k)$	$\hat{\alpha}_2(k)$	$\hat{Q}_1(1,1)$	$\hat{Q}_1(1,2)$	$\hat{Q}_2(2,1)$	$\hat{Q}_2(2,2)$	δ (%)
H-SG	100	0.12073	1.00630	1.28170	0.27718	-1.05330	0.24319	46.515
	200	0.15542	0.99994	1.46020	0.34842	-1.13250	0.25237	37.894
	500	0.19097	0.99881	1.63130	0.40068	-1.16540	0.27073	27.783
	1000	0.19574	0.99909	1.70020	0.42844	-1.17340	0.28223	22.985
	2000	0.19675	0.99918	1.74460	0.44918	-1.18200	0.28645	19.308
	3000	0.19746	0.99927	1.76450	0.45801	-1.18630	0.28715	17.401
H-RLS	100	0.18217	0.99591	1.59560	0.40657	-1.19710	0.29721	29.877
	200	0.19438	0.99782	1.71560	0.45212	-1.20920	0.29861	21.443
	500	0.20000	1.00010	1.79900	0.47926	-1.20610	0.29906	13.084
	1000	0.19995	0.99999	1.82590	0.48698	-1.20330	0.30343	9.2098
	2000	0.19998	1.00000	1.83810	0.49161	-1.20460	0.30043	6.8055
	3000	0.20005	1.00000	1.84220	0.49391	-1.20210	0.29781	5.6569
H-SG-RLS	100	0.12428	1.01470	1.55690	0.38635	-1.28930	0.27393	35.347
	200	0.15598	1.00540	1.72180	0.45472	-1.28140	0.28286	25.059
	500	0.19035	0.99970	1.85280	0.47755	-1.24880	0.29210	13.499
	1000	0.19540	0.99949	1.85970	0.48596	-1.22420	0.29931	9.8219
	2000	0.19649	0.99954	1.85440	0.49174	-1.21600	0.29804	7.8781
	3000	0.19724	0.99956	1.85390	0.49397	-1.21050	0.29605	6.6068
True values		0.20000	1.00000	1.85000	0.50000	-1.20000	0.30000	

Table 5. The H-MI-SG-RLS estimates and errors ($\sigma^2 = 0.1^2$).

p	k	$\hat{\alpha}_1(k)$	$\hat{\alpha}_2(k)$	$\hat{Q}_1(1,1)$	$\hat{Q}_1(1,2)$	$\hat{Q}_2(2,1)$	$\hat{Q}_2(2,2)$	δ (%)
1	100	0.12428	1.01470	1.55690	0.38635	-1.28930	0.27393	35.347
	500	0.19035	0.99970	1.85280	0.47755	-1.24880	0.2921	13.499
	1000	0.19540	0.99949	1.85970	0.48596	-1.22420	0.29931	9.8219
	2000	0.19649	0.99954	1.85440	0.49174	-1.21600	0.29804	7.8781
	3000	0.19724	0.99956	1.85390	0.49397	-1.21050	0.29605	6.6068
2	100	0.17752	0.97351	1.79800	0.46944	-1.15820	0.28896	17.911
	500	0.19623	0.99571	1.85750	0.49122	-1.22510	0.29462	9.9535
	1000	0.19801	0.99783	1.85710	0.49299	-1.21300	0.30060	7.5234
	2000	0.19844	0.99833	1.85160	0.49448	-1.20920	0.29865	5.9897
	3000	0.19877	0.99864	1.85120	0.49576	-1.20520	0.29656	5.0628
8	100	0.19043	0.97410	1.84310	0.49302	-1.14280	0.29516	16.230
	500	0.19799	0.99617	1.85460	0.49584	-1.21550	0.29711	7.9670
	1000	0.19886	0.99813	1.85490	0.49533	-1.20840	0.30160	6.2476
	2000	0.19910	0.99858	1.85060	0.49553	-1.20690	0.29912	5.1626
	3000	0.19929	0.99885	1.85040	0.49661	-1.20360	0.29698	4.4864
True values		0.20000	1.00000	1.85000	0.50000	-1.20000	0.30000	

$$\alpha(z) = 1 + 0.2z^{-1} + 1z^{-2},$$

$$\boldsymbol{Q}(z) = \begin{bmatrix} 1.85 & 0.5 \\ -1.2 & 0.3 \end{bmatrix} z^{-1}.$$

The inputs $\{u_1(k)\}$ and $\{u_2(k)\}$ are taken as two random sequences with zero mean and unit variances, $\{v_1(k)\}$ and $\{v_2(k)\}$ are taken as white noise sequences with zero mean and variances $\sigma_1^2 = \sigma_2^2 = 0.1^2$ and $\sigma_1^2 = \sigma_2^2 = 0.5^2$, respectively. All the simulations are conducted in MATLAB 2018 running on Hp Pavilion Gaming (15-dk1000), with 2.40 GHz, Core-i5 processor and 8 GB RAM. Applying the H-SG-RLS algorithm and the H-MI-SG-RLS algorithm to estimate the parameters of this example system, the parameter estimates and their errors δ are shown in Tables 4 and 5, and the parameter estimation errors versus k are shown in Figs. 1 to 3, where

$$\delta := \left\{ \frac{\|\hat{\boldsymbol{\alpha}}(k) - \boldsymbol{\alpha}\|^2 + \|\hat{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}\|^2}{\|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\theta}\|^2} \right\}^{1/2} \times 100\%.$$

Fig. 4 shows the time consumption of the H-SG algorithm, the H-MI-SG-RLS algorithm, and other algorithms. From

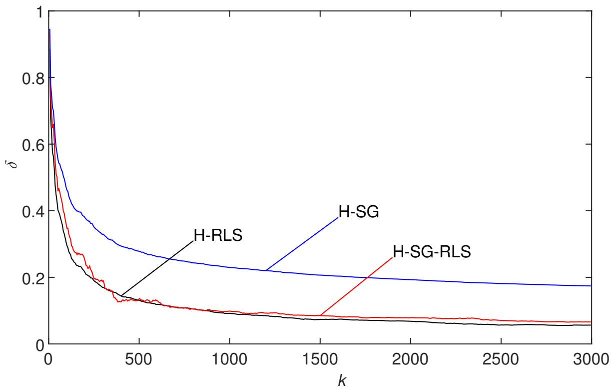


Fig. 1. The estimation errors δ versus k ($\sigma^2 = 0.1^2$).

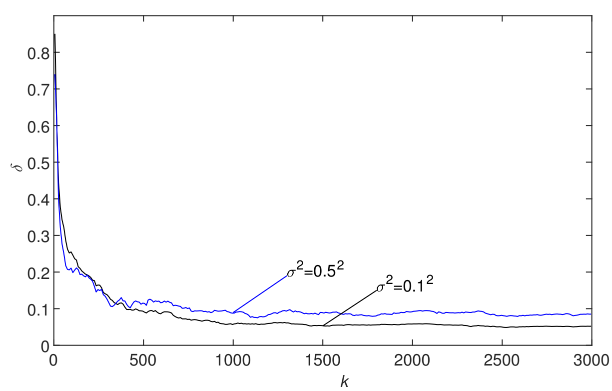


Fig. 2. The H-SG-RLS errors δ versus k with σ^2.

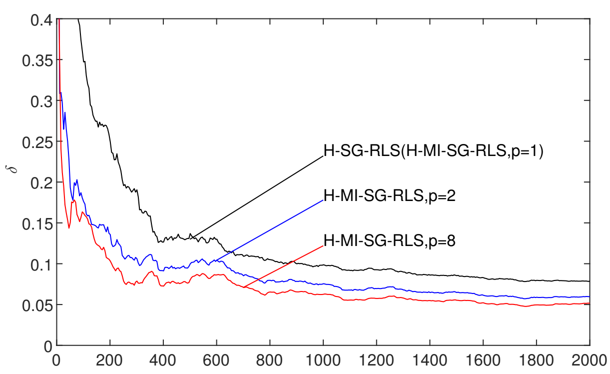


Fig. 3. The H-MI-SG-RLS errors δ versus k with p.

Tables 1-5 and Figs. 1-3, we can draw the following conclusions.

1) The H-SG-RLS algorithm requires less computational load than the H-RLS algorithm because the H-SG-RLS algorithm does not need to calculate the gain matrix $L_1(k) \in \mathbb{R}^{n_\alpha \times m}$ to avoid the calculation of the matrix inversion for larger m - see Tables 1 and 2.
2) The parameter estimation errors of the H-SG, H-RLS, H-SG-RLS and H-MI-SG-RLS algorithms become (generally) smaller with the data length k increasing - see Tables 4 and 5.
3) The parameter estimation accuracy is related to the noise variance. The larger the noise variance is, the lower the estimation accuracy is - see Fig. 2.
4) The innovation length p can effectively improve the parameter estimation accuracy of the H-SG-RLS algorithm. With the increase of the innovation length p, the parameter estimation becomes more stable and the calculation time becomes longer - see Fig. 3.

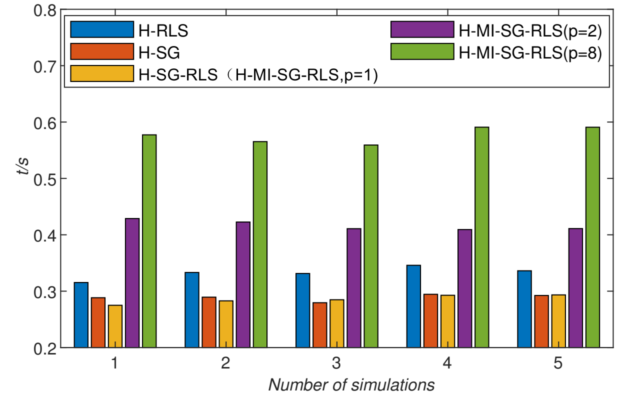


Fig. 4. The actual time consumption of the H-SG, H-RLS, H-SG-RLS, and H-MI-SG-RLS algorithms.

6. CONCLUSIONS

This paper derives the H-SG-RLS algorithm and the H-MI-SG-RLS algorithm for identifying the multivariable CAR-like system based on the hierarchical identification theory and the multi-innovation identification theory. Simulation results show that the proposed algorithms are effective. The proposed H-SG-RLS algorithm combines the advantages of the H-SG algorithm and the H-RLS algorithm, with a small amount of calculation, short calculation time, and high precision. The H-MI-SG-RLS algorithm provides more accurate parameter estimation than the H-SG-RLS algorithm. The algorithms in this paper can be extended to study parameter identification issues of linear and nonlinear stochastic processes [66-72] and can be applied to engineering systems such as process control and machine learning.

REFERENCES

[1] J. Wang, Y. Ji, and C. Zhang, "Iterative parameter and order identification for fractional-order nonlinear finite impulse response systems using the key term separation," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 8, pp. 1562-1577, August 2021.

[2] Y. Ji, Z. Kang, and X. Liu, "The data filtering based multiple-stage Levenberg-Marquardt algorithm for Ham-

merstein nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 15, pp. 7007-7025, October 2021

[3] M. Li and X. Liu, "Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique," *International Journal of Control, Automation, and Systems*, vol. 18, no. 6, pp. 1581-1592, June 2020.

[4] M. Li and X. Liu, "Maximum likelihood hierarchical least squares-based iterative identification for dual-rate stochastic systems," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 2, pp. 240-261, 2021.

[5] M. Li and X. Liu, "Iterative identification methods for a class of bilinear systems by using the particle filtering technique," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 10, pp. 2056-2074, 2021.

[6] Y. Fan and X. Liu, "Two-stage auxiliary model gradient-based iterative algorithm for the input nonlinear controlled autoregressive system with variable-gain nonlinearity," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 14, pp. 5492-5509, September 2020.

[7] X. Liu and Y. Fan, "Maximum likelihood extended gradient-based estimation algorithms for the input nonlinear controlled autoregressive moving average system with variable-gain nonlinearity," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 9, pp. 4017-4036, June 2021.

[8] G. Chen, M. Gan, C. Chen, and H. 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.

[9] L. Xu, G. Song, "A recursive parameter estimation algorithm for modeling signals with multi-frequencies," *Circuits Systems and Signal Processing*, vol. 39, no. 8, pp. 4198-4224, August 2020.

[10] L. Xu, "Separable multi-innovation Newton iterative modeling algorithm for multi-frequency signals based on the sliding measurement window," *Circuits Systems and Signal Processing*, vol. 41, no. 2, pp. 805-830, February 2022.

[11] Y. Ji, C. Zhang, Z. Kang, and T. Yu, "Parameter estimation for block-oriented nonlinear systems using the key term separation," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 9, pp. 3727-3752, June 2020.

[12] P. Ma and L. Wang, "Filtering-based recursive least squares estimation approaches for multivariate equation-error systems by using the multiinnovation theory," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 9, pp. 1898-1915, September 2021.

[13] F. Ding, *System Identification - Auxiliary Model Identification Idea and Methods*, Science Press, Beijing, 2017.

[14] Y. Ji, Z. Kang, and C. Zhang, "Two-stage gradient-based recursive estimation for nonlinear models by using the data filtering," *International Journal of Control, Automation, and Systems*, vol. 19, no. 8, pp. 2706-2715, 2021.

[15] Y. Ji, X. Jiang, and L. Wan, "Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems," *Journal of the Franklin Institute*, vol. 357, no. 8, pp. 5019-5032, May 2020.

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

[17] J. You, C. Yu, J. Sun, and J. Chen, "Generalized maximum entropy based identification of graphical ARMA models," *Automatica*, vol. 141, p. 110319, 2022.

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

[19] C. Yu, Y. Li, H. Fang, and J. Chen, "System identification approach for inverse optimal control of finite-horizon linear quadratic regulators," *Automatica*, vol. 129, 109636, 2021.

[20] 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.

[21] T. Katayama and H. Ase, "Linear approximation and identification of MIMO Wiener-Hammerstein systems," *Automatica*, vol. 71, pp. 118-124, September 2016.

[22] F. Ding, *System Identification - Hierarchical Identification Principle and Methods*, Science Press, Beijing, 2024.

[23] F. Ding and Y.J. Wang, "Hierarchical multi-innovation identification methods for multivariable equation-error-like type systems," *Journal of Nanjing University of Information Science and Technology* (Natural Science Edition), vol. 6, no. 5, pp. 385-404, 2014.

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

[25] F. Ding and T. Chen, "Hierarchical least squares identification methods for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 397-402, March 2005.

[26] Y. Ji, Z. Kang, and L. Xu, "Model recovery for multi-input signal-output nonlinear systems based on the compressed sensing recovery theory," *Journal of the Franklin Institute*, vol. 359, no. 5, pp. 2317-2339, March 2022.

[27] C. Zhang and Y. Ji, "Gradient parameter estimation of a class of nonlinear systems based on the maximum likelihood principle," *International Journal of Control, Automation, and Systems*, vol. 20, no. 5, pp. 1393-1404, May 2022.

[28] Y. Ji and Z. Kang, "Three-stage forgetting factor stochastic gradient parameter estimation methods for a class of nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 3, pp. 871-987, 2021.

[29] L. Xu, "Separable Newton recursive estimation method through system responses based on dynamically discrete measurements with increasing data length," *International Journal of Control, Automation, and Systems*, vol. 20, no. 2, pp. 432-443, February 2022.

[30] L. Xu, "Auxiliary model multiinnovation stochastic gradient parameter estimation methods for nonlinear sandwich systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 1, pp. 148-165, January 2021.

[31] J. Pan, S. Liu, and X. Wan, "Hierarchical recursive least squares estimation algorithm for secondorder Volterra nonlinear systems," *International Journal of Control, Automation, and Systems*, vol. 20, no. 12, pp. 3940-3950, 2022.

[32] X. Zhang, "Highly computationally efficient state filter based on the delta operator," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 6, pp. 875-889, June 2019.

[33] J. Pan, Y. Liu, and J. Shu, "Gradient-based parameter estimation for an exponential nonlinear autoregressive time-series model by using the multi-innovation," *International Journal of Control, Automation, and Systems*, vol. 21, no. 1, pp. 140-150, 2023.

[34] X. Zhang, "State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1157-1173, July 2019.

[35] Y. Wang, S. Tang, and X. Gu, "Parameter estimation for nonlinear Volterra systems by using the multi-innovation identification theory and tensor decomposition," *Journal of the Franklin Institute*, vol. 359, no. 2, pp. 1782-1802, January 2022.

[36] Y. Wang and L. Yang, "An efficient recursive identification algorithm for multilinear systems based on tensor decomposition," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 16, pp. 7920-7936, 2021.

[37] X. Zhang, "Optimal adaptive filtering algorithm by using the fractional-order derivative," *IEEE Signal Processing Letters*, vol. 29, pp. 399-403, 2022.

[38] Y. Zhou, "Partially-coupled nonlinear parameter optimization algorithm for a class of multivariate hybrid models," *Applied Mathematics and Computation*, vol. 414, p. 126663, Februray 2022.

[39] S. Su, T. Tang, and Y. Wang, "Design of running grades for energy-efficient train regulation: A case study for beijing yizhuang line," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 2, 189-200, February 2021.

[40] Y. Cao, Z. Wang, and G. Xie, "Bio-inspired speed curve optimization and sliding mode tracking control for subway trains," IEEE Transactions on Vehicular Technology, vol. 68, no. 7, pp. 6331-6342, July 2019.

[41] Y. Cao, Y. Sun, and T. Wen, "Fault diagnosis of train plug door based on a hybrid criterion for IMFs selection and fractional wavelet package energy entropy," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7544-7551, August 2019.

[42] Y. Cao, L. Ma, and W. Xu, "Standard analysis for transfer delay in CTCS-3," *Chinese Journal of Electronics*, vol. 26, no. 5, pp. 1057-1063, September 2017.

[43] Y. Cao, J. Wen, and L. Ma, "Tracking and collision avoidance of virtual coupling train control system," *Alexandria Engineering Journal*, vol. 60, no. 2, pp. 2115-2125, 2021.

[44] S. Su, X. Wang, and J. Yin, "An energy-efficient train operation approach by integrating the metro timetabling and eco-driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 10, pp. 4252-4268, 2020.

[45] F. Ding and T. Chen, "Combined parameter and output estimation of dual-rate systems using an auxiliary model," *Automatica*, vol. 40, no. 10, pp. 1739-1748.

[46] F. Ding, "Filtered auxiliary model recursive generalized extended parameter estimation methods for Box-Jenkins systems," *International Journal of Robust and Nonlinear Control*, vol. 33, 2023. DOI: 10.1002/rnc.6657

[47] J. Chen, "Modified Kalman filtering based multi-step-length gradient iterative algorithm for ARX models with random missing outputs," *Automatica*, vol. 118, 109034, 2020.

[48] Y. Wang, "Recursive parameter estimation algorithm for multivariate output-error systems," *Journal of the Franklin Institute*, vol. 355, no. 12, pp. 5163-5181, 2018.

[49] 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.

[50] L. J. Wan, "Decomposition- and gradient-based iterative identification algorithms for multivariable systems using the multi-innovation theory," *Circuits Systems and Signal Processing*, vol. 38, no. 7, pp. 2971-2991, July 2019.

[51] H. J. Wang, "Singularly degenerate heteroclinic cycles with nearby apple-shape attractors," *International Journal of Bifurcation and Chaos*, vol. 33, no. 1, 2350011, 2023,

[52] 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.

[53] L. Xu, Separable synchronous multi-innovation gradient-based iterative signal modeling from on-line measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, 6501313, 2022.

[54] J. Wang, Y. Ji, and L. Xu, "Two-stage gradient-based iterative algorithms for the fractional-order nonlinear systems by using the hierarchical identification principle," *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 7, pp. 1778-1796, 2022.

[55] Y. Zhou, "Hierarchical estimation approach for RBF-AR models with regression weights based on the increasing data length," *IEEE Transactions on Circuits and Systems-II: Express Briefs*, vol. 68, no, 12, pp. 3597-3601, December 2021.

[56] Z. Kang, Y. Ji, and X. Liu, "Hierarchical recursive least squares algorithms for Hammerstein nonlinear autoregressive output-error systems," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 11, pp. 2276-2295, November 2021.

[57] Y. Zhou, "Modeling nonlinear processes using the radial basis function-based state-dependent autoregressive models," *IEEE Signal Processing Letters*, vol. 27, pp. 1600-1604, 2020.

[58] J. Ding and W. Zhang, "Finite-time adaptive control for nonlinear systems with uncertain parameters based on the command filters," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 9, pp. 1754-1767, September 2021.

[59] X. Zhang, "Hierarchical parameter and state estimation for bilinear systems,' *International Journal of Systems Science*, vol. 51, no. 2, 275-290, 2020.

[60] Y. Cao, Y. Sun, and P. Li, "A sound-based fault diagnosis method for railway point machines based on two-stage feature selection strategy and ensemble classifier," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12074-12083, 2022.

[61] Y. Cao, J. Wen, and T. Wen, "Parameter-varying artificial potential field control of virtual coupling system with nonlinear dynamics," *Fractals*, vol. 30, no. 2, 2240099, 2022.

[62] X. Zhang, "Recursive identification of bilinear timedelay systems through the redundant rule," *Journal of the Franklin Institute*, vol. 357, no. 1, pp. 726-747, 2020.

[63] X. Zhang, "Adaptive parameter estimation for a general dynamical system with unknown states," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 4, pp. 1351-1372, March 2020.

[64] Y. Cao, "Research on virtual coupled train control method based on GPC & VAPF," *Chinese Journal of Electronics*, vol. 31, no. 5, pp. 897-905, 2022.

[65] X. Zhang, "Recursive parameter estimation methods and convergence analysis for a special class of nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 4, pp. 1373-1393, March 2020.

[66] F. Ding, *System Identification - Multi-Innovation Identification Theory and Methods*, Science Press, Beijing, 2016.

[67] F. Ding, "Least squares parameter estimation and multiinnovation least squares methods for linear fitting problems from noisy data," *Journal of Computational and Applied Mathematics*, vol. 426, 115107, July 2023.

[68] J. Xiong, J. Pan, and G. Chen, "Sliding mode dual-channel disturbance rejection attitude control for a quadrotor," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 10, pp. 10489-10499, October 2022.

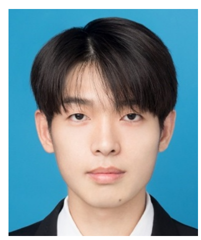
[69] J. Ding, "Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data," *IEEE Transactions on Automatic Control*, vol. 56, no. 11, pp. 2677-2683, November 2011.

[70] J. Pan, Q. Chen, J. Xiong, and G. Chen, "A novel quadruple boost nine level switched capacitor inverter," *Journal of Electrical Engineering & Technology*, vol. 18, pp. 467-480, 2023.

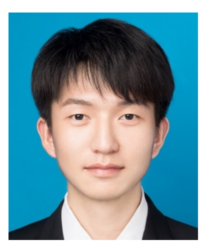
[71] L. Xu, "Hierarchical recursive signal modeling for multifrequency signals based on discrete measured data," *International Journal of Adaptive Control and Signal Processing*, vol. 35, no. 5, pp. 676-693, May 2021.

[72] C. Wei, "Overall recursive least squares and overall stochastic gradient algorithms and their convergence for feedback nonlinear controlled autoregressive systems," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 9, pp. 5534-5554, June 2020.

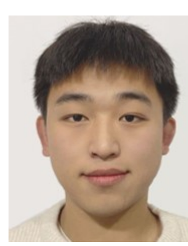
Jian Pan was born in Wuhan, China. He received his B.Sc. degree from Hubei University of Technology (Wuhan, China) in 1984. He has been a Professor in School of Electrical and Electronic Engineering, Hubei University of Technology. His research interests include control science and engineering, computer control systems, and power electronics.



Huijian Zhang was born in Xianning (Hubei Province, China) in 1999. He received his B.D. degree from Hubei University of Technology (Wuhan, China) in 2020. He is currently a Master student in the School of Electrical and Electronic Engineering, Hubei University of Technology. His interests include system identification and parameter estimation.



Hongzhan Guo was born in Yingshan (Hubei Province, China) in 1998. He received his B.Sc degree from Huanggang Normal University (Huanggang, China) in 2019. He is currently a master's degree student in the School of Electrical and Electronic Engineering, Hubei University of Technology. His interests include system modeling and resonant converters.



Sunde Liu was born in Anqing, Anhui Province, China. He received his B.Sc. degree from the Jiangxi University of Science and Technology (Ganzhou, China) in 2018. He is now a master's degree student at the Huibei University of Technology, Wuhan, China. His research interests include system identification and control theory.

Yuqing Liu was born in Xiangtan, Hunan Province, China. She received her B.Sc. degree from the Hunan Institute of Science and Technology (Yueyang, China) in 2019. She is now a master's degree student at the Huibei University of Technology, Wuhan, China. Her research interests include nonlinear system identification and control theory.