International Journal of Control, Automation, and Systems 21(1) (2023) 140-150
http://dx.doi.org/10.1007/s12555-021-1018-8
ISSN:1598-6446 eISSN:2005-4092
http://www.springer.com/12555

Gradient-based Parameter Estimation for a Nonlinear Exponential Autoregressive Time-series Model by Using the Multi-innovation

Jian Pan*, **Yuqing Liu, and Jun Shu**

Abstract: The parameter estimation methods for the nonlinear exponential autoregressive model are investigated in this paper. We develop a forgetting factor gradient parameter estimation algorithm for improving the estimation accuracy. For the purpose of improving the identification accuracy further, a forgetting factor multi-innovation stochastic gradient algorithm is derived by using the multi-innovation theory. The effectiveness of the proposed algorithms is proved by a simulation example.

Keywords: Exponential autoregressive model, forgetting factor, gradient search, multi-innovation theory, parameter estimation.

1. INTRODUCTION

Mathematical models are important for studying natural sciences [1-5]. The solutions to many problems in engineering are based on the mathematical models [6-10]. Playing a key role in the system identification, parameter estimation is always a hot topic in the filed of system identification [11-15]. There are a large number of random processes in real life, most of them have typical nonlinear dynamic characteristics. The mean and variance of the time series are inherently non-stationary. For instance, exchange rates and stock transactions may change over time. Linear time-series models cannot explain the nonlinear behavior of random processes [16,17], so many scholars have proposed a large number of nonlinear time-series models since the 1970s. Nonlinear time-series analysis is an important subject in the fields of engineering, economics and natural sciences [18,19]. It is an effective tool to predict future values by using current and past data. For example, Ahmed and Kopsaftopoulos employed an output-only recursive maximum likelihood time-varying auto-regressive model to investigate the uncertainty in guided wave propagation by analyzing the time-varying model parameters [20]. The exponential autoregressive (ExpAR) model is a typical class of the nonlinear time-series model. The ExpAR model was introduced in a study of more general models of the form $X_{t+1} = \lambda\left(X_t\right) + A_{t+1}$ [21]. Meanwhile, Ozaki proposed the ExpAR model. As he was in search of a non-explosive approximation solution of $\lambda(x) = \left\{a + b\left(1 - cx^2\right)\right\}x$ and the approximate model of the threshold autoregressive models [22].

The ExpAR model was used to capture the remarkable characteristics of nonlinear vibration, like jump phenomena, amplitude dependent frequency shifts and perturbed limit cycles [23]. Besides, the ExpAR model can be successfully used in many fields, such as ecology, hydrology, and speech signal processing [24]. For instance, Sanchez *et al*. studied a problem of approximation with exponential functions and showed its relevance with economic science [25]. Because of the wide applications of the ExpAR system, the study of its properties and identification methods is meaningful. Although Ozaki gave some sufficient conditions for the stationarity of the ExpAR models [26,27], they had certain limitations. As a result, it is necessary to study the stationarity and ergodicity of the model further. Chan and Tong proved that the ergodicity and stability of the ExpAR model can be connected through the physical concept of the Lyapunov function [28]. In the past few decades, Koul and Schick constructed the asymptotic estimates for the ExpAR model by using a sample splitting technique [29]. Recently, the parameter estimation of the ExpAR model has attracted much attention. Without considering the special conditions of the model parameters, Chen *et al*. proposed a variable projection method for a generalized ExpAR model [30]. For a higher-order system, the algebraic identification algorithms are almost impossible to estimate the unknown parameters directly. At this time, adopting the recursive and iterative methods, such as the least mean square algorithm, the stochastic gradient algorithm and the maximum likelihood identifi-

Manuscript received December 2, 2021; revised February 3, 2022; accepted February 18, 2022. Recommended by Associate Editor Feng Ding under the direction of Editor Jay H. Lee. This work was supported by the National Natural Science Foundation of China (No. 61571182, 61273192).

Jian Pan, Yuqing Liu, and Jun Shu are with the School of Electrical and Electronic Engineering, Hubei University of Technology, Wuhan 430068, China (e-mails: jpan@163.com, liuyuqing@hbut.edu.cn, shujun@mail.hbut.edu.cn).
* Corresponding author.

cation method can be more available.

The stochastic gradient algorithm has a poor convergence rate because it fails to make sufficient use of data. By means of the multi-innovation identification theory, the Kalman filtering algorithm, and the maximum likelihood method, many efficient gradient-based methods are developed for online identification and off-line identification [31-34]. In that case, Chen *et al*. gave a multi-step gradient-based iterative algorithm for a controlled autoregressive model with missing outputs based on the Kalman filtering. Compared with the classical gradient-based iterative algorithm, the proposed method can not only improves the convergence rate but also improve the estimation accuracy [35]. The recursive algorithm is used to update the estimates by using new observations [36-41]. Therefore, this paper focuses on the highly efficient recursive parameter identification methods for the ExpAR model based on the gradient search. This paper studies the negative gradient search algorithm for estimating the parameters of the ExpAR model with white noise and a multi-innovation gradient identification algorithm with the forgetting factor. The proposed parameter estimation algorithms in this paper are based on this identification model in (1).

This paper is organized as follows: Sections 2 and 3 develop a gradient search algorithm and a forgetting factor multi-innovation gradient algorithm for ExpAR models. Section 4 presents an illustrative example to test the proposed algorithm. Finally, some concluding remarks are given in Section 5.

2. THE GRADIENT SEARCH ALGORITHM

Let us introduce some notation. $A =: X$ represents that X is defined by A; $X := A$ stands for X is defined by A; $\mathbf{1}_n$ denotes an n-dimensional column vector whose elements are all 1; the superscript T stands for the vector or matrix transpose; the norm of a matrix (or a column vector) X is defined as $\|X\|^2 := \mathrm{tr}[XX^{\mathrm{T}}]$.

Here derives a stochastic gradient algorithm for the exponential autoregressive model with order nth [42-44]

$$y(t)=\sum_{i=1}^{n}[\alpha_i+\beta_i \mathrm{e}^{-\gamma y^2(t-1)}]y(t-i)+v(t), \tag{1}$$

where t is defined as the time variable, $y(t)$ is the observation of the system at time t, $v(t)$ is the stochastic white noise with zero mean and variance σ^2, α_i, β_i and γ are the model parameters to be identified. Suppose that $y(t)=0$ and $v(t)=0$ for $t \leqslant 0$.

Many identification methods are derived based on the identification models of the systems [45-51] and these methods can be used to estimate the parameters of other linear systems and nonlinear systems [52-58] and can be applied to other fields [59-64]. Define the unknown parameter vector and the error between the observed output and the model output as

$$\begin{aligned}
&\theta := [\alpha^{\mathrm{T}}, \beta^{\mathrm{T}}, \gamma]^{\mathrm{T}} \in \mathbb{R}^{2n+1},\\
&\alpha := [\alpha_1, \alpha_2, \cdots, \alpha_n]^{\mathrm{T}} \in \mathbb{R}^n,\\
&\beta := [\beta_1, \beta_2, \cdots, \beta_n]^{\mathrm{T}} \in \mathbb{R}^n,\\
&v(\theta,t) := y(t)-\sum_{i=1}^{n}[\alpha_i+\beta_i \mathrm{e}^{-\gamma y^2(t-1)}]y(t-i).
\end{aligned}$$

Based on the model in (1), define the criterion function

$$J_1(\theta) := \frac{1}{2}v^2(\theta,t).$$

Taking the first-order derivative of $J_1(\theta)$ with respect to θ gives

$$\begin{aligned}
\mathrm{grad}[J_1(\theta)] &:= \frac{\partial J_1(\theta)}{\partial \theta}\\
&= \left[\frac{\partial J_1(\theta)}{\partial \alpha^{\mathrm{T}}}, \frac{\partial J_1(\theta)}{\partial \beta^{\mathrm{T}}}, \frac{\partial J_1(\theta)}{\partial \gamma}\right]^{\mathrm{T}} \in \mathbb{R}^{2n+1},\\
\frac{\partial J_1(\theta)}{\partial \alpha} &= \left[\frac{\partial J_1(\theta)}{\partial \alpha_1}, \frac{\partial J_1(\theta)}{\partial \alpha_2}, \cdots, \frac{\partial J_1(\theta)}{\partial \alpha_n}\right]^{\mathrm{T}} \in \mathbb{R}^n,\\
\frac{\partial J_1(\theta)}{\partial \beta} &= \left[\frac{\partial J_1(\theta)}{\partial \beta_1}, \frac{\partial J_1(\theta)}{\partial \beta_2}, \cdots, \frac{\partial J_1(\theta)}{\partial \beta_n}\right]^{\mathrm{T}} \in \mathbb{R}^n,\\
\frac{\partial J_1(\theta)}{\partial \gamma} &= y^2(t-1)\mathrm{e}^{-\gamma y^2(t-1)}v(\theta,t)\sum_{i=1}^{n}\beta_i y(t-i),\\
\frac{\partial J_1(\theta)}{\partial \alpha_i} &= -y(t-i)v(\theta,t),\\
\frac{\partial J_1(\theta)}{\partial \beta_i} &= -y(t-i)\mathrm{e}^{-\gamma y^2(t-1)}v(\theta,t).
\end{aligned}$$

Define the information vectors

$$\begin{aligned}
\phi(t) &:= [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}} \in \mathbb{R}^n,\\
\varphi(\theta,t) &:= \frac{\partial y(t)}{\partial \theta}\\
&= [\phi^{\mathrm{T}}(t), \phi^{\mathrm{T}}(t)\mathrm{e}^{-\gamma y^2(t-1)},\\
&\quad -y^2(t-1)\mathrm{e}^{-\gamma y^2(t-1)}\phi^{\mathrm{T}}(t)\beta]^{\mathrm{T}} \in \mathbb{R}^{2n+1}.
\end{aligned}$$

Let $\hat{\theta}(t) := [\hat{\alpha}^{\mathrm{T}}(t), \hat{\beta}^{\mathrm{T}}(t), \hat{\gamma}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n+1}$ be the estimate of the parameter vector θ at t. As a result, the gradient vector $\mathrm{grad}[J_1(\theta)]$ becomes

$$\begin{aligned}
\frac{\partial J_1(\hat{\theta}(t-1))}{\partial \alpha_i} &= -y(t-i)v(\hat{\theta}(t-1),t),\\
\frac{\partial J_1(\hat{\theta}(t-1))}{\partial \beta_i} &= -y(t-i)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}v(\hat{\theta}(t-1),t),\\
\frac{\partial J_1(\hat{\theta}(t-1))}{\partial \gamma} &= y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}v(\hat{\theta}(t-1),t)\\
&\quad \times \phi^{\mathrm{T}}(t)\hat{\beta}(t-1).
\end{aligned}$$

The information vector and innovation can be denoted as

$$\varphi(\hat{\theta}(t-1),t) = [\phi^{\mathrm{T}}(t), \phi^{\mathrm{T}}(t)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)},$$

$$-y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)}\phi^{\mathrm{T}}(t)\hat{\beta}(t-1)]^{\mathrm{T}},$$

$$v(\hat{\theta}(t-1),t)=y(t)-\sum_{i=1}^{n}[\hat{\alpha}_i(t-1)+\hat{\beta}_i(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}]y(t-i).$$

Let $0\leqslant\lambda\leqslant 1$ be the forgetting factor. We can get the stochastic gradient (SG) algorithm for estimating the parameter vector $\hat{\theta}=[\hat{\alpha}^{\mathrm{T}},\hat{\beta}^{\mathrm{T}},\hat{\gamma}]^{\mathrm{T}}$ as

$$\begin{aligned}\hat{\alpha}_i(t)&=\hat{\alpha}_i(t-1)-\mu(t)\frac{\partial J_1(\hat{\theta}(t-1))}{\partial\alpha_i}\\&=\hat{\alpha}_i(t-1)+\frac{y(t-i)}{r(t)}v(\hat{\theta}(t-1),t),\end{aligned}\tag{2}$$

$$\begin{aligned}\hat{\beta}_i(t)&=\hat{\beta}_i(t-1)-\mu(t)\frac{\partial J_1(\hat{\theta}(t-1))}{\partial\beta_i}\\&=\hat{\beta}_i(t-1)+\frac{y(t-i)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}}{r(t)}\\&\quad\times v(\hat{\theta}(t-1),t),\end{aligned}\tag{3}$$

$$\begin{aligned}\hat{\gamma}(t)&=\hat{\gamma}(t-1)-\mu(t)\frac{\partial J_1(\hat{\theta}(t-1))}{\partial\gamma}\\&=\hat{\gamma}(t-1)-\frac{y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}}{r(t)}\\&\quad\times\phi^{\mathrm{T}}(t)\hat{\beta}(t-1)v(\hat{\theta}(t-1),t),\end{aligned}\tag{4}$$

$$\begin{aligned}\hat{v}(t)&:=v(\hat{\theta}(t-1),t)\\&=y(t)-\sum_{i=1}^{n}[\hat{\alpha}_i(t-1)\\&\quad+\hat{\beta}_i(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}]y(t-i),\end{aligned}\tag{5}$$

$$r(t)=\lambda r(t-1)+\|\varphi(\hat{\theta}(t-1),t)\|^2,\ r(0)=1,\tag{6}$$

$$\phi(t)=[y(t-1),y(t-2),\cdots,y(t-n)]^{\mathrm{T}},\tag{7}$$

$$\begin{aligned}\hat{\varphi}(t)&=\varphi(\hat{\theta}(t-1),t)\\&=[\phi^{\mathrm{T}}(t),\phi^{\mathrm{T}}(t)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)},\\&\quad-y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)}\phi^{\mathrm{T}}(t)\hat{\beta}(t-1)]^{\mathrm{T}},\end{aligned}\tag{8}$$

$$\hat{\alpha}(t)=[\hat{\alpha}_1(t),\hat{\alpha}_2(t),\cdots,\hat{\alpha}_n(t)]^{\mathrm{T}},\tag{9}$$

$$\hat{\beta}(t)=[\hat{\beta}_1(t),\hat{\beta}_2(t),\cdots,\hat{\beta}_n(t)]^{\mathrm{T}},\tag{10}$$

$$\hat{\theta}(t)=[\hat{\alpha}^{\mathrm{T}}(t),\hat{\beta}^{\mathrm{T}}(t),\hat{\gamma}(t)]^{\mathrm{T}}.\tag{11}$$

The steps of the FF-SG parameter estimation algorithm in (2)-(11) are as follows.

1) Set the initial values: let $t=1$, $\hat{\theta}(0)=\mathbf{1}_{2n+1}/p_0$, $p_0=10^6$, $r(0)=1$, set the data length L and the forgetting factor λ.
2) Collect the output data $y(t)$ and form $\phi(t)$ using (7).
3) Compute the innovation $\hat{v}(t)$ by (5), the information vector $\hat{\varphi}(t)$ by (8) and $r(t)$ by (6).
4) Compute the estimates $\hat{\alpha}_i(t)$, $\hat{\beta}_i(t)$ and $\hat{\gamma}(t)$ using (2)-(4) and form $\hat{\alpha}(t)$, $\hat{\beta}(t)$ and $\hat{\theta}(t)$ using (9)-(11).
5) If $t<L$, then increase t by 1, go to Step 2; otherwise, output the parameter estimates and terminate this calculation procedure.

The characteristic of the stochastic gradient algorithm is using a single innovation to modify the parameter estimates of each recursive process. In order to make full use of the innovation to improve the accuracy of parameter estimation, we can increase the innovation length and derive a multi-innovation stochastic gradient algorithm.

3. THE MULTI-INNOVATION GRADIENT ALGORITHM

The criterion function of the stochastic gradient algorithm only uses the current observation data. To a certain degree, the more observation data are used in each recursive calculation, the higher accuracy of the parameter estimation is obtained. Hence, we consider using a batch of data to construct a criterion function. Considering the data from $j=t-p+1$ to $j=t$ and defining the criterion function. In this section, we derive a multi-innovation stochastic gradient parameter estimation algorithm based on the model in the stochastic framework.

Define the criterion function with respect to θ

$$J_2(\theta):=\frac{1}{2}\sum_{j=0}^{p-1}v^2(\theta,t-j),$$

where the integer $p\geqslant 1$ is the length of the innovation. Define the information vectors

$$\begin{aligned}\phi(t,j)&:=[y(t-j-1),\ y(t-j-2),\ \cdots,\\&\qquad y(t-j-n)]^{\mathrm{T}}\in\mathbb{R}^n,\\ \varphi(\theta,t-j)&:=[\phi^{\mathrm{T}}(t,j),\phi^{\mathrm{T}}(t,j)\mathrm{e}^{-\gamma y^2(t-j-1)},\\&\qquad-y^2(t-j-1)\mathrm{e}^{-\gamma y^2(t-j-1)}\\&\qquad\times\phi^{\mathrm{T}}(t,j)\beta]^{\mathrm{T}}\in\mathbb{R}^{2n+1}.\end{aligned}$$

Let $\hat{\theta}(t):=[\hat{\alpha}^{\mathrm{T}}(t),\ \hat{\beta}^{\mathrm{T}}(t),\ \hat{\gamma}(t)]^{\mathrm{T}}\in\mathbb{R}^{2n+1}$ be the estimate of the parameter vector θ at t. As a result, the gradient vector grad$[J_2(\theta)]$ becomes. We can obtain the following multi-innovation stochastic gradient (MISG) algorithm for estimating the parameter vector $\hat{\theta}=[\hat{\alpha}^{\mathrm{T}},\hat{\beta}^{\mathrm{T}},\hat{\gamma}]^{\mathrm{T}}$ as

$$\begin{aligned}\hat{\alpha}_i(t)&=\hat{\alpha}_i(t-1)-\mu(t)\frac{\partial J_2(\hat{\theta}(t-1))}{\partial\alpha_i}\\&=\hat{\alpha}_i(t-1)+\frac{1}{r(t)}\sum_{j=0}^{p-1}y(t-j-i)\\&\quad\times\hat{v}(t-j),\end{aligned}\tag{12}$$

$$\begin{aligned}\hat{\beta}_i(t)&=\hat{\beta}_i(t-1)-\mu(t)\frac{\partial J_2(\hat{\theta}(t-1))}{\partial\beta_i}\\&=\hat{\beta}_i(t-1)+\frac{1}{r(t)}\sum_{j=0}^{p-1}y(t-j-i)\end{aligned}$$

$$\times \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-j-1)}\hat{v}(t-j), \tag{13}$$

$$\begin{aligned}\hat{\gamma}(t) &= \hat{\gamma}(t-1) - \mu(t)\frac{\partial J_2(\hat{\theta}(t-1))}{\partial \gamma}\\ &= \hat{\gamma}(t-1) - \frac{1}{r(t)}\sum_{j=0}^{p-1} y^2(t-j-1)\\ &\quad \times \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-j-1)}\hat{v}(t-j)\\ &\quad \times \phi^{\mathrm{T}}(t,j)\hat{\beta}(t-1),\end{aligned} \tag{14}$$

$$r(t) = \lambda r(t-1) + \|\hat{\varphi}(t-j)\|^2,\ r(0)=1, \tag{15}$$

$$\phi(t,j) = [y(t-j-1), y(t-j-2), \cdots,\ y(t-j-n)]^{\mathrm{T}}, \tag{16}$$

$$\begin{aligned}\hat{\varphi}(t-j) &:= \varphi(\hat{\theta}(t-1), t-j)\\ &= [\phi^{\mathrm{T}}(t,j), \phi^{\mathrm{T}}(t,j)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-j-1)},\\ &\quad -y^2(t-j-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-j-1)}\\ &\quad \times \phi^{\mathrm{T}}(t,j)\hat{\beta}(t-1)]^{\mathrm{T}},\end{aligned} \tag{17}$$

$$\begin{aligned}\hat{v}(t-j) &:= v(\hat{\theta}(t-1), t-j)\\ &= y(t-j) - \sum_{i=1}^{n}[\hat{\alpha}_i(t-1) + \hat{\beta}_i(t-1)\\ &\quad \times \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-j-1)}]y(t-j-i),\end{aligned} \tag{18}$$

$$\hat{\alpha}(t) = [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \cdots, \hat{\alpha}_n(t)]^{\mathrm{T}}, \tag{19}$$

$$\hat{\beta}(t) = [\hat{\beta}_1(t), \hat{\beta}_2(t), \cdots, \hat{\beta}_n(t)]^{\mathrm{T}}, \tag{20}$$

$$\hat{\theta}(t) = [\hat{\alpha}^{\mathrm{T}}(t), \hat{\beta}^{\mathrm{T}}(t), \hat{\gamma}(t)]^{\mathrm{T}}. \tag{21}$$

The steps of computing the FF-MISG parameter estimates in (12)-(21) are as follows:

1) Set the initial values: let $t=1$, $\hat{\theta}(0) = \mathbf{1}_{2n+1}/p_0$, $p_0 = 10^6$, $r(0)=1$ and choose the data length L and the forgetting factor λ.
2) Collect the output data $y(t)$ and form $\phi(t,j)$ using (16).
3) Compute the innovation $\hat{v}(t-j)$ by (18), and the information vector $\hat{\varphi}(t-j)$ by (17) and $r(t)$ by (15).
4) Use (12)-(14) to compute $\hat{\alpha}_i(t)$, $\hat{\beta}_i(t)$ and $\hat{\gamma}(t)$ respectively, and form $\hat{\alpha}(t)$, $\hat{\beta}(t)$ and $\hat{\theta}(t)$ using (19)-(21).
5) If $t < L$, then increase t by 1, go to Step 2; otherwise, output the parameter estimates and terminate the computational procedure.

Compared with the stochastic gradient algorithm, the multi-innovation stochastic gradient method can obtain more accurate parameter estimates due to the dynamically changing batch data.

4. SIMULATION EXAMPLES

In this section, we present an example to illustrate the performance of the proposed algorithms. The performance of the proposed algorithms is examined for the ExpAR identification in the perspective of the estimation error based on two evaluation metrics: δ and MSE, which are defined as

$$\delta = \frac{\mathrm{norm}(\hat{\theta}(t)-\theta)}{\mathrm{norm}(\theta)},\ \mathrm{MSE} = \mathrm{mean}(\hat{\theta}(t)-\theta),$$

respectively, where θ is the parameter vector and $\hat{\theta}(t)$ represents the estimate of the parameter vector θ at time t.

Consider the following exponential autoregressive time-series model

$$\begin{aligned} y(t) &= [\alpha_1 + \beta_1 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-1)\\ &\quad + [\alpha_2 + \beta_2 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-2) + v(t),\\ \theta &= [\alpha_1,\ \alpha_2,\ \beta_1,\ \beta_2,\ \gamma]^{\mathrm{T}}\\ &= [0.83,\ -0.66,\ 0.95,\ -0.56,\ 0.25]^{\mathrm{T}}.\end{aligned}$$

The noise $\{v(t)\}$ is simulated as an uncorrelated random sequence with zero mean and variance σ^2. Then the output of the system is generated on the basis of the above model parameters. Taking the data length $L=3000$ and the noise variance $\sigma^2 = 0.20^2$, applying the FF-SG algorithm in (2)-(11) with the forgetting factor $\lambda=1$, $\lambda=0.98$, $\lambda=0.96$ and $\lambda=0.94$ to estimate the parameters of this example system. The parameter estimates and their errors are shown in Tables 1-4. The parameter estimation errors versus t of the FF-SG algorithm is shown in Fig. 1.

Considering that the parameters identification accuracy can be increased with the appropriate multi-innovation length, taking $\lambda=0.98$, we applying the FF-MISG algorithm in (12)-(21) with the innovation lengths $p=2$, $p=4$ and $p=6$ to estimate the parameters. The parameter estimates and their errors are shown in Tables 5-7 and their errors $\delta := \|\hat{\theta}(t)-\theta\|/\|\theta\|$ are shown in Fig. 2.

Table 1. The FF-SG estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda=1$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.30077	-0.18691	0.25644	-0.23162	-0.10304	71.17535
200	0.30093	-0.19130	0.25347	-0.23662	-0.08906	70.79141
500	0.30346	-0.19643	0.25726	-0.24143	-0.09170	70.37833
1000	0.30568	-0.20112	0.25984	-0.24578	-0.08987	69.95603
2000	0.30748	-0.20667	0.26152	-0.25115	-0.08630	69.50818
3000	0.30817	-0.20907	0.26007	-0.25350	-0.06858	69.08932
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 2. The FF-SG estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.98$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.37126	-0.27764	0.35424	-0.29036	-0.05280	60.37523
200	0.37614	-0.29607	0.35802	-0.30927	-0.04287	59.03501
500	0.40073	-0.33910	0.38611	-0.34901	-0.04114	55.28528
1000	0.45024	-0.39036	0.43946	-0.39488	-0.01726	48.96937
2000	0.54304	-0.45336	0.53138	-0.45860	0.01979	38.89625
3000	0.62232	-0.50367	0.60189	-0.52358	0.06127	30.64630
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 3. The FF-SG estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.96$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.31348	-0.20581	0.26205	-0.25478	-0.10374	69.61629
200	0.32541	-0.23065	0.27077	-0.27978	-0.07883	67.25411
500	0.38130	-0.30746	0.34261	-0.34127	-0.05806	59.02729
1000	0.49277	-0.40897	0.46233	-0.43147	-0.00463	45.42732
2000	0.68237	-0.49896	0.63355	-0.55182	0.08507	26.99649
3000	0.76037	-0.55833	0.72388	-0.62037	0.15739	18.07783
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 4. The FF-SG estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.94$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.32267	-0.21786	0.26918	-0.26813	-0.10677	68.53837
200	0.34500	-0.25457	0.28924	-0.30424	-0.07801	64.80400
500	0.43358	-0.35728	0.39606	-0.38625	-0.03624	52.71379
1000	0.60221	-0.47906	0.56232	-0.51023	0.04530	34.09177
2000	0.77899	-0.53959	0.73913	-0.62111	0.16226	17.42899
3000	0.78692	-0.59701	0.82061	-0.60864	0.19701	10.74382
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 5. The FF-MISG parameter estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.98$, $p = 2$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.49459	-0.41578	0.47737	-0.42189	-0.00490	44.67272
200	0.51249	-0.43140	0.49469	-0.43805	0.00414	42.57981
500	0.56467	-0.46585	0.54467	-0.47616	0.03354	36.84731
1000	0.65285	-0.51149	0.62443	-0.53857	0.07657	28.11586
2000	0.76240	-0.53957	0.72832	-0.60694	0.14803	18.33844
3000	0.77804	-0.58503	0.78782	-0.61592	0.18397	13.23753
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 6. The FF-MISG parameter estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.98, p = 4$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.61836	-0.49804	0.60202	-0.50896	0.05843	31.03227
200	0.65574	-0.51193	0.63142	-0.53358	0.06244	28.09340
500	0.71937	-0.53712	0.68982	-0.58481	0.13600	21.26255
1000	0.78102	-0.57444	0.77371	-0.61497	0.17937	14.25171
2000	0.80308	-0.58970	0.86548	-0.60068	0.20301	8.32926
3000	0.80158	-0.63329	0.90643	-0.58405	0.22730	4.33343
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 7. The FF-MISG parameter estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.98$, $p = 6$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.92369	-0.65771	0.68845	-0.59239	0.36646	19.54527
200	0.91766	-0.67283	0.72356	-0.56707	0.31296	16.20921
500	0.87885	-0.66914	0.78577	-0.53785	0.31013	11.81558
1000	0.87650	-0.67620	0.85966	-0.53544	0.28443	7.17627
2000	0.84222	-0.64073	0.91942	-0.55505	0.24192	2.53642
3000	0.82580	-0.65956	0.93966	-0.56242	0.25501	0.80516
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

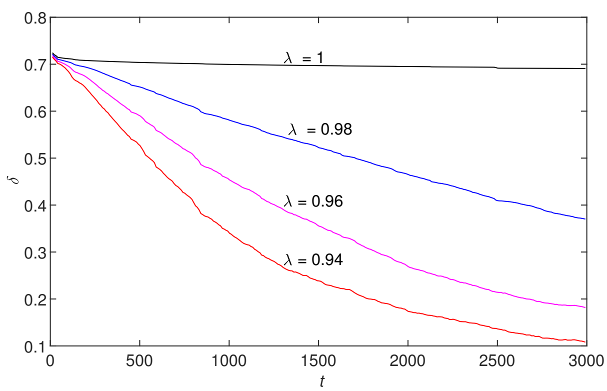

Fig. 1. The FF-SG estimation errors δ versus t ($\sigma^2 = 0.20^2$).

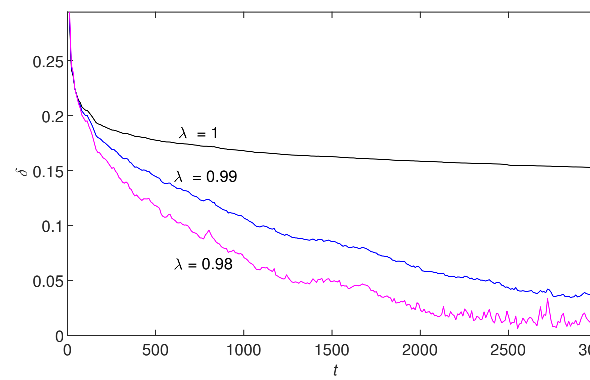

Fig. 3. The FF-MISG estimation errors δ versus t ($\sigma^2 = 0.20^2$, $p = 6$).

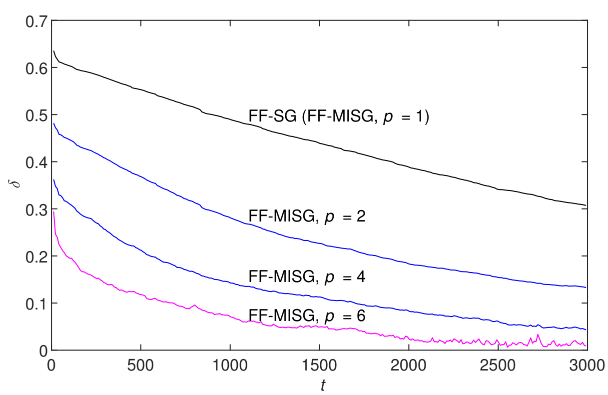

Fig. 2. The FF-MISG estimation errors δ versus t ($\sigma^2 = 0.20^2$, $\lambda = 0.98$).

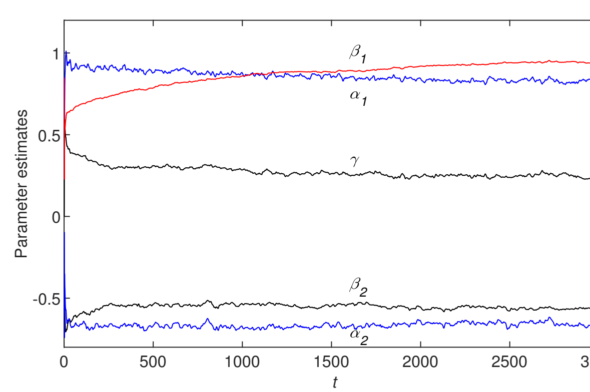

Fig. 4. The FF-MISG estimation errors δ versus t ($\sigma^2 = 0.20^2$, $\lambda = 0.98, p = 6$).

We select three different forgetting factor comparing their parameter estimates errors when $p = 6$. The results are shown in Tables 7-9, the parameter estimation errors versus t is shown in Fig. 3.

From Fig. 4, we can see that the appropriate forgetting factor and multi-innovation length, the parameter estimates of the FF-MISG algorithm tend to the corresponding true values.

The robustness of the FF-MISG identification method is examined for noise levels, through statistical analyses based on the MSE. The corresponding results are presented in Figs. 5 and 6. In Fig. 5, the x-axis represents the data length, the y-axis shows the corresponding mean MSE value for different noise variances $\sigma^2 = 0.10^2$, $\sigma^2 = 0.20^2$ and $\sigma^2 = 0.30^2$, the lower multi-innovation length p provides comparatively less accurate results. Meanwhile, the higher multi-innovation length provides more accurate results while it is more easily affected by different noise levels. The variation of the MSE versus at t is presented in Fig. 6.

In order to test the effectiveness of the proposed algorithms, we use the FF-MISG estimation method to build

Table 8. The FF-MISG parameter estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 1$, $p = 6$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.91540	-0.65018	0.67015	-0.61239	0.36322	20.52797
200	0.90988	-0.66126	0.68441	-0.60069	0.34364	19.06502
500	0.90672	-0.65921	0.69993	-0.58257	0.33345	17.77141
1000	0.90226	-0.66115	0.71200	-0.57225	0.32672	16.80981
2000	0.89955	-0.66791	0.72287	-0.56561	0.31257	15.85836
3000	0.89251	-0.67153	0.72898	-0.56062	0.30741	15.29062
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

Table 9. The FF-MISG parameter estimates and their errors ($\sigma^2 = 0.20^2$, $\lambda = 0.99$, $p = 6$).

t	α_1	α_2	β_1	β_2	γ	$\delta(\%)$
100	0.91957	-0.65363	0.67894	-0.60239	0.36527	20.05155
200	0.91502	-0.66650	0.70340	-0.58271	0.32981	17.66243
500	0.89100	-0.66875	0.74451	-0.55128	0.31504	14.47213
1000	0.88473	-0.67762	0.80128	-0.53592	0.29088	10.73151
2000	0.86508	-0.65789	0.86407	-0.53757	0.24969	6.16100
3000	0.83868	-0.66934	0.89840	-0.55014	0.25470	3.50040
True values	0.83000	-0.66000	0.95000	-0.56000	0.25000	

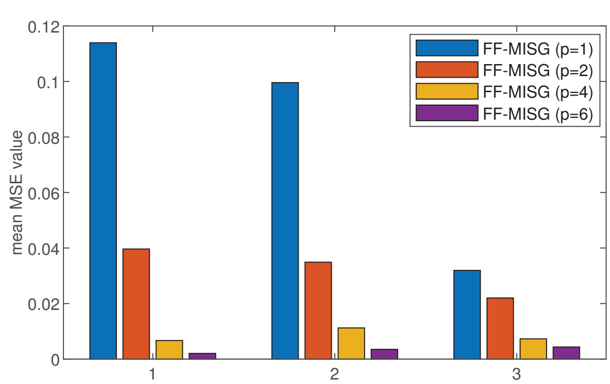

Fig. 5. The variation of the mean MSE values versus $\sigma^2(\lambda = 0.98)$.

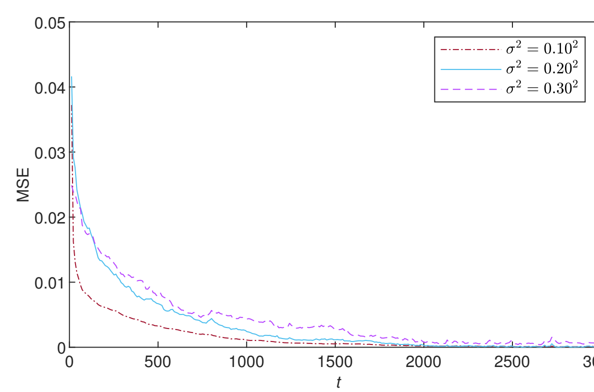

Fig. 6. The MSE against independent executions of the FF-MISG algorithm ($\lambda = 0.98$, $p = 6$).

the simulated model. The predicted output of the model can be obtained as

$$\hat{y}(t) = y(t) - [\hat{\alpha}_1(t) + \hat{\beta}_1(t)\mathrm{e}^{-\hat{\gamma}(t)y^2(t-1)}]y(t-1) - [\hat{\alpha}_2(t) + \hat{\beta}_2(t)\mathrm{e}^{-\hat{\gamma}(t)y^2(t-1)}]y(t-2).$$

Using the FF-MISG estimates with noise variance $\sigma^2 = 0.20^2$ and $t = 3000$ in Table 1 to construct the estimated model

$$\begin{aligned}\hat{y}(t) &= y(t) - [\hat{\alpha}_1(t) + \hat{\beta}_1(t)\mathrm{e}^{-\hat{\gamma}(t)y^2(t-1)}]y(t-1) \\ &\quad - [\hat{\alpha}_2(t) + \hat{\beta}_2(t)\mathrm{e}^{-\hat{\gamma}(t)y^2(t-1)}]y(t-2), \\ \hat{\theta}(t) &= [\hat{\alpha}_1(t), \hat{\alpha}_2(t), \hat{\beta}_1(t), \hat{\beta}_2(t), \hat{\gamma}(t)]^{\mathrm{T}} \\ &= [0.83705, -0.65480, 0.91406, \\ &\quad -0.55765, 0.24109]^{\mathrm{T}}.\end{aligned}$$

For model validation, we use a set of data from $t = L+1$ to $t = L+L_r$ to calculate the predicted output $\hat{y}(t)$, L_r is taken as 100. The output error between the predicted model and the actual system versus t is shown in Fig. 7. To evaluate the prediction performance, the root mean square error can be defined as

$$E(\hat{y}(t)) = \sqrt{\frac{1}{L_r}\sum_{j=L+1}^{L+L_r}[\hat{y}(t) - y(t)]^2} = 0.1802.$$

It can be seen from Fig. 7 that there is a proper multi-innovation length, the model output is close to the system actual output. It is shown that the simulated model can

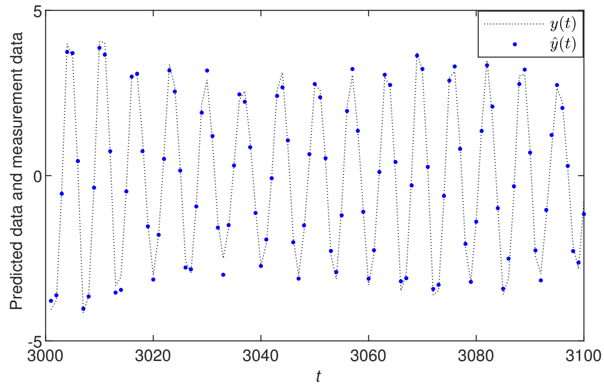

Fig. 7. The FF-MISG estimation errors δ versus t ($\sigma^2 = 0.20^2, \lambda = 0.98, p = 6$).

capture the dynamic performance of the system well.

From Tables 1-9 and Figs. 1-5, we can draw the following conclusions. 1) The estimation errors given by the FFSG algorithm and the FF-MISG algorithm become smaller as the recursive time t increases - see the estimation errors δ in Tables 1-9. 2) The FF-MISG algorithm needs less recursion number achieving the same accuracy compared with the FFSG algorithm - see Fig. 2. 3) The estimation errors given by the FFSG and the FF-MISG algorithms become smaller with the forgetting factor λ decreasing - see Fig. 1 and Fig. 3. 4) The FF-MISG algorithm provides better parameter estimates and MSE values than the FFSG algorithm and this accuracy is further enhanced by increasing the length of multi-innovation. Moreover, the FF-MISG algorithm provides better parameter estimates than the FFSG algorithm for all noise levels - see Fig. 5.

5. CONCLUSIONS

This article considers the parameter estimation of the nonlinear exponential autoregressive time-series model. By means of the negative gradient search and the multi-innovation theory, the SG algorithm and MISG algorithm are proposed to identify the unknown parameters. In order to improve the computational efficiency, we obtain the FF-MISG algorithm by introducing the forgetting factor. The simulation results demonstrate that the FF-MISG algorithm has a higher parameter estimation accuracy than the FF-SG algorithm under the same noise variance. The proposed algorithms in this paper can combine other methods to study parameter identification of different systems [65-71] and can be applied to other control and schedule areas [72-77] such as information processing and industrial process systems [78-80] and so on.

REFERENCES

[1] 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.

[2] 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.

[3] M. Li and X. Liu, "Particle filtering-based iterative identification methods for a class of nonlinear systems with interval-varying measurements," *International Journal of Control, Automation, and Systems*, vol. 20, no. 7, pp. 2239-2248, 2022.

[4] 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, August 2021.

[5] 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.

[6] 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.

[7] Y. Ji, Z. Kang, and X. Liu, "The data filtering based multiple-stage Levenberg-Marquardt algorithm for Hammerstein nonlinear systems," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 15, pp. 7007-7025, October 2021.

[8] 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.

[9] 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, February 2021.

[10] 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, October 2021.

[11] 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.

[12] L. Xu and 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.

[13] 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.

[14] 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.

[15] 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.

[16] T. Ozaki, "Non-linear time series models for non-linear random vibrations," *Journal of Applied Probability*, vol. 17, no. 1, pp. 84-93, 1980.

[17] S. Dong, T. Liu, and Q. Wang, "Identification of Hammerstein systems with time delay under load disturbance," *IET Control Theory and Applications*, vol. 12, no. 7, pp. 942-952, April 2018.

[18] W. Tang and A. Rollin, "Model identification for ARMA time series through convolutional neural networks," *Decision Support Systems*, vol. 146, Article Number: 113544, July 2021.

[19] T. Liu, W. Zhang, L. Ye, M. Ueland, S. Forbes, and S.W. Su, " A novel multi-odour identification by electronic nose using non-parametric modelling-based feature extraction and time-series classification," *Sensors and Actuators B-Chemical*, vol. 298, Article Number: 126690, November 2019.

[20] S. Ahmed and F. Kopsaftopoulos, "Stochastic identification of guided wave propagation under ambient temperature via non-stationary time series models," *Sensors*, vol. 21, no. 16, Article Number: 5672, January 2021.

[21] D. Jones and D. Cox, "Nonlinear autoregressive processes," *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, vol. 360, no. 1700, pp. 71-95, March 1978.

[22] T. Ozaki, "Non-linear threshold autoregressive models for non-linear random vibrations," *Journal of Applied Probability*, vol. 18, no. 2, pp. 443-451, June 1981.

[23] M. Al-Qassam and J. Lane, "Forecasting exponential autoregressive models of order 1," *Journal of Time Series Analysis*, vol. 10, no. 2, pp. 95-113, March 1989.

[24] M. Merzougui, H. Dridi, and A. Chadli, "Test for periodicity in restrictive EXPAR models," *Communications in Statistics-Theory and Methods*, vol. 45, no. 9, pp. 2770-2783, May 2016.

[25] J. Sanchez, J. Torvisco, and M. Arias, "TAC method for fitting exponential autoregressive models and others: Applications in economy and finance," *Mathematics*, vol. 9, no. 8, Article Number: 862, January 2021.

[26] T. Ozaki, "The statistical analysis of perturbed limit cycle processes using nonlinear time series models," *Journal of Time Series Analysis*, vol. 3, no. 1, pp. 29-41, January 1982.

[27] T. Ozaki, "2 Non-linear time series models and dynamical systems," *Handbook of Statistics*, vol. 5, no. 5, pp. 25-83, 1985.

[28] K. Chan and H. Tong, "On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations," *Advances in Applied Probability*, vol. 17, no. 3, pp. 666-678, September 1985.

[29] H. Koul and A. Schick, "Efficient estimation in nonlinear autoregressive time series models," *Bernoulli*, vol. 3, no. 3, pp. 247-277, September 1997.

[30] G. Chen, M. Gan, and G. Chen, "Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications," *Information Sciences*, pp. 46-57, Article Number: 438, April 2018.

[31] F. Ding, H. Ma, J. Pan, and E. F. Yang, "Hierarchical gradient- and least squares-based iterative algorithms for input nonlinear output-error systems using the key term separation," *Journal of the Franklin Institute*, vol. 358, no. 9, pp. 5113-5135, June 2021.

[32] M. H. Li and X. M. Liu, "The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique," *Signal Processing*, vol. 147, pp. 23-34, June 2018.

[33] Y. Ji and Z. Kang, "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, 2022.

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

[35] J. Chen and Y. Liu, "Modified Kalman filtering based multi-step-length gradient iterative algorithm for ARX models with random missing outputs," *Automatica*, vol. 118, Article Number: 109034, August 2020.

[36] J. Chen and C. Chen, "A novel reduced-order algorithm for rational models based on Arnoldi process and Krylov subspace," *Automatica*, vol. 129, Article Number: 109663, July 2021.

[37] C. Wei, X. Zhang, and E. Yang, "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, 2022.

[38] J. Pan, X. Jiang, 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.

[39] H. Ma, J. Pan, and W. Ding, "Partially-coupled least squares based iterative parameter estimation for multi-variable output-error-like autoregressive moving average systems," *IET Control Theory and Applications*, vol. 13, no. 18, pp. 3040-3051, December 2019.

[40] 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.

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

[42] T. Ozaki and H. Oda, "Non-linear time series model identification by Akaike's information criterion," *IFAC Proceedings Volumes*, vol. 10, no. 12, pp. 83-91, January 1977.

[43] R. Stine, "Non-linear time series: A dynamical system approach," *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 903-905, September 1992.

[44] H. Xu, L. Wan, and T. Hayat, "Fitting the exponential autoregressive model through recursive search," *Journal of the Franklin Institute*, vol. 356, no. 11, pp. 5801-5818, July 2019.

[45] 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, February 2021.

[46] 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.

[47] L. Xu and E. Yang, "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.

[48] X. Zhang and E. Yang, "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.

[49] X. Zhang and E. Yang, "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.

[50] L. Xu and E. Yang, "Separable recursive gradient algorithm for dynamical systems based on the impulse response signals," *International Journal of Control, Automation, and Systems*, vol. 18, no. 12, pp. 3167-3177, December 2020.

[51] L. Xu and Q. Zhu, "Decomposition strategy-based hierarchical least mean square algorithm for control systems from the impulse responses," *International Journal of Systems Science*, vol. 52, no. 9, pp. 1806-1821, 2021.

[52] Y. J. Wang, S. H. Tang, and X. B. 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.

[53] Y. J. 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, November 2021.

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

[55] J. Hou, F. Chen, and Z. Zhu, "Gray-box parsimonious subspace identification of Hammerstein-type systems," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 10, pp. 9941-9951, October 2021.

[56] N. Zhao, A. Wu, Y. Pei, and D. Niyato, "Spatial-temporal aggregation graph convolution network for efficient mobile cellular traffic prediction," *IEEE Communications Letters*, vol. 26, no. 3, pp. 587-591, 2022.

[57] Y. Chen, C. Zhang, C. Liu, Y. Wang, and X. Wan, "Atrial fibrillation detection using feedforward neural network," *Journal of Medical and Biological Engineering*, vol. 42, pp. 63-73, 2022.

[58] 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.

[59] S. Su, T. Tang, J. Xun, F. Cao, 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, pp. 189-200, February 2021.

[60] 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.

[61] 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.

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

[63] 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, April 2021.

[64] S. Su, X. Wang, Y. Cao, 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, October 2020.

[65] Y. H. Zhou and X. Zhang, "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.

[66] 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.

[67] 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.

[68] 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.

[69] X. Wang, "Modified particle filtering-based robust estimation for a networked control system corrupted by impulsive noise," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 2, pp. 830-850, January 2022.

[70] J. Pan, W. Li, and H. 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.

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

[72] Y. Cao, Y. K. 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.

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

[74] S. Su, J. F. She, and Y. Zhou, "A nonlinear safety equilibrium spacing based model predictive control for virtually coupled train set over gradient terrains," *IEEE Transactions on Transportation Electrification*, vol. 8, no. 2, pp. 2810-2824, 2022.

[75] J. M. Li, "A novel nonlinear optimization method for fitting a noisy Gaussian activation function," *International Journal of Adaptive Control and Signal Processing*, vol. 36, no. 3, pp. 690-707, 2022.

[76] Y. An, Y. J. Zhang, W. J. Cao, and Z. Q. He, "A lightweight and practical anonymous authentication protocol based on bit-self-test PUF," *Electronics*, vol. 11, no. 5, 772, 2022.

[77] J. Wang, C. Ding, M. Wu, Y. Liu, and G. Chen, "Lightweight multiple scale-patch dehazing network for real-world hazy image," *KSII Transactions on Internet and Information Systems*, vol. 15, no. 12, pp. 4420-4438, December 2022.

[78] M. Li and X. Liu, "The filtering-based maximum likelihood iterative estimation algorithms for a special class of nonlinear systems with autoregressive moving average noise using the hierarchical identification principle," *International Journal of Adaptive Control and Signal Processing*, vol. 33, no. 7, pp. 1189-1211, 2019.

[79] C. Zhang, H. Liu, 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, 2022.

[80] T. Cui, "Moving data window-based partially-coupled estimation approach for modeling a dynamical system involving unmeasurable states," *ISA Transactions*, vol. 128, Part B, pp. 437-452, 2022.

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 the School of Electrical and Electronic Engineering, Hubei University of Technology. His research interests include control science and engineering, computer control systems, and power electronics.

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 student at the Huibei University of Technology, Wuhan, China. Her research interests include nonlinear system identification and control theory.

Jun Shu was born in Wuhan, China. He received his B.Sc. degree in 2009 and an M.Sc. degree from Hubei University of Technology, Wuhan, China. He has been an Associate Professor at Hubei University of Technology. His research interests include intelligent control and process control.