IET Research Journals

Regular Paper

IET Journals
The Institution of Engineering and Technology

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Recursive search-based identification algorithms for the exponential autoregressive time series model with colored noise

Huan Xu[1], *Feng Ding*[1,2*], *Erfu Yang*[3]

[1] *Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, People's Republic of China*
[2] *College of Automation and Electronic Engineering, Qingdao University of Science and Technology, Qingdao 266061, People's Republic of China*
[3] *Department of Design, Manufacture and Engineering Management, Space Mechatronic Systems Technology Laboratory, Strathclyde Space Institute, University of Strathclyde, Glasgow G1 1XJ, Scotland, United Kingdom*
** E-mail: fding@jiangnan.edu.cn*

Abstract: This paper focuses on the recursive parameter estimation problems for the nonlinear exponential autoregressive model with moving average noise (the ExpARMA model for short). By means of the gradient search, an extended stochastic gradient (ESG) algorithm is derived. Considering the difficulty of determining the step-size in the ESG algorithm, a numerical approach is proposed to obtain the optimal step-size. In order to improve the parameter estimation accuracy, we employ the multi-innovation identification theory to develop a multi-innovation extended stochastic gradient (MI-ESG) algorithm for the ExpARMA model. Introducing a forgetting factor into the MI-ESG algorithm, the parameter estimation accuracy can be further improved. With an appropriate innovation length and forgetting factor, the variant of the MI-ESG algorithm is effective to identify all the unknown parameters of the ExpARMA model. A simulation example is provided to test the proposed algorithms.

1 Introduction

Nonlinearities are common in financial, hydrologic, industrial and many other practical systems [1]. The exponential autoregressive (ExpAR) family is an important kind of nonlinear time series models. The ExpAR models are applied to the statistical analysis of the ship rolling data, the Canadian lynx data and other nonlinear time series. The famous Canadian lynx data are the records of Canadian lynx trapped in the years 1821-1934. Recently, Chen *et al.* discussed the stationary conditions of several generalised ExpAR models, and adopted these models to model and predict the Canadian lynx data [2]. On the other hand, the ExpAR family has shown the appropriateness in capturing certain well-known features of nonlinear vibration theory, such as amplitude-dependent frequency, jump phenomena and limit cycles [3, 4]. Given a time series $\{y(t), y(t-1), y(t-2), \cdots\}$, the ExpAR model can be described as a stochastic difference equation

$$\begin{aligned}y(t) &= [\alpha_1+\beta_1\mathrm{e}^{-\gamma y^2(t-1)}]y(t-1)+[\alpha_2+\beta_2\mathrm{e}^{-\gamma y^2(t-1)}]\\ &\quad\times y(t-2)+\cdots+[\alpha_n+\beta_n\mathrm{e}^{-\gamma y^2(t-1)}]y(t-n)+v(t),\end{aligned}$$

where $v(t)$ is a stochastic white noise, n is the system degree, and α_i, β_i and γ are the model parameters. The form in the above equation represents the classic ExpAR model, some modified versions have been presented. For instance, in order to give a more sophisticated specification for the dynamics of the characteristic roots of autoregressive (AR) models, Ozaki derived a variant of the ExpAR model in [5] using the Hermite type polynomials:

$$\begin{aligned}y(t) &= \sum_{i=1}^{n}\{\alpha_i+[\beta_{i0}+\sum_{j=1}^{m_i}\beta_{ij}y^j(t-1)]\mathrm{e}^{-\gamma y^2(t-1)}\}\\ &\quad\times y(t-i)+v(t).\end{aligned}$$

Introducing a time-delay d and a scalar parameter ζ, Teräsvirta developed a different variant of the ExpAR model in [6]:

$$\begin{aligned}y(t) &= \{\alpha_0+\beta_0\mathrm{e}^{-\gamma[y(t-d)-\zeta]^2}\}\\ &\quad+\sum_{i=1}^{n}\{\alpha_i+\beta_i\mathrm{e}^{-\gamma[y(t-d)-\zeta]^2}\}y(t-i)+v(t).\end{aligned}$$

Other generalised ExpAR models were summarised in [2]. However, all the ExpAR models mentioned above involve the white noise. This is not the case in many situations in practice. Colored noises with different structures, such as the autoregressive (AR) noise, the moving average (MA) noise and the autoregressive moving average (ARMA) noise [7], are also required to be considered in modeling. Many parameter estimation methods have been developed for linear and nonlinear systems with colored noises [8, 9] and can be applied to many areas [10–14].

System identification and parameter estimation are effective tools to establish the mathematical models of many dynamical systems [15, 16], and have been widely used in the area of sliding mode control [17], fault diagnosis [18] and so on. In the field of system identification, a great deal of publications are devoted to the identification methods, such as moving horizon estimation [19] and hierarchical identification [20, 21]. On the identification of bilinear systems, Li *et al.* presented the least squares based iterative algorithms by using the data filtering technique [22], the auxiliary model based least squares iterative algorithms by using interval-varying measurements [23] and the filtering-based maximum likelihood iterative estimation algorithms by using the hierarchical identification principle [24]. In respect of the ExpAR model identification, setting the parameter of the nonlinear part at a specific value, Haggan and Ozaki obtained the least squares (LS) estimates for the ExpAR model [25]; with a certain definition of the parameter of the nonlinear part, Shi *et al.* gave the parameter estimates of the linear part by the LS estimator [26]. Imposing no special conditions on the model parameters, Chen *et al.* developed a variable projection method for generalised ExpAR models, where the parameters of the linear part are obtained by the LS algorithm and the parameters of the nonlinear one are estimated by a line search procedure [2]. In the previous work in [27, 28], decomposing the classic ExpAR model into two sub-identification (Sub-ID) models by the hierarchical identification principle, decomposition-based recursive parameter estimation algorithms were developed for the ExpAR model. The parameters of the linear and nonlinear Sub-ID models are estimated by an interactive way. Different from all the above-mentioned work, setting no special conditions on the parameters of the ExpARMA model, we aim to investigate recursive identification algorithms by using a novel identification technique in this paper.

In recent decades, many techniques have been be devoted to improving the identification accuracy and convergence rate [29, 30] such as the multi-innovation identification theory . The innovation is the useful information which can improve the parameter and state estimation accuracy [31]. The multi-innovation identification theory has been developed as a significant branch of system identification. The key idea is to expand the scalar innovation into a multi-dimensional innovation vector and to make full use of the measurement data [32]. Recently, using the multi-innovation identification theory, Zhang *et al.* derived a multi-innovation extended stochastic gradient algorithm for estimating the unknown parameters of bilinear systems [33]. Moreover, gradient-based methods are widely used in system identification [34]. For instance, Xu and Ding presented a gradient-based iterative algorithm for identifying the unknown amplitudes, angular frequencies and phases of multi-frequency signals [35] and derived a filtering-based gradient iterative algorithm for pseudo-linear autoregressive moving average systems [36]. In this paper, through the gradient search and the multi-innovation identification theory, we study the recursive identification algorithms for the ExpAR model with MA noise. The main contributions of this paper are as follows.

• After parameterization, the ExpARMA model is written as an identification model. Defining a criterion function, the identification problem is transformed into a nonlinear optimization problem. Using the gradient search to minimize the optimization problem, we derive an extended stochastic gradient (ESG) algorithm and its variant for the ExpARMA model.
• Note that the information vector of the identification model contains the unknown parameter, the computation of the step-size in the ESG algorithm is a highly nonlinear optimization problem. In this paper, we employ the one-dimensional search method to obtain the optimal step-size.
• Applying the multi-innovation identification theory, we expand the scalar innovation into a multi-dimensional innovation vector, such that we can make full use of the sampled data and innovations, and develop a multi-innovation extended stochastic gradient (MI-ESG) algorithm and its variant for the ExpARMA model.

In summary, the rest of this paper is organised as follows. Section 2 describes the identification problem for the ExpARMA model. An extended stochastic gradient algorithm and its variant are derived in Section 3. A multi-innovation extended stochastic gradient algorithm and its variant are presented in Section 4. Section 5 provides a numerical example for testing the effectiveness of the proposed algorithms. Finally, some concluding remarks are given in Section 6.

2 Problem description

Given a time series $\{y(t), y(t-1), y(t-2), \cdots\}$, the ExpARMA model of order n can be described as

$$\begin{aligned} y(t) &= [\alpha_1 + \beta_1 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-1) + \cdots \\ &\quad + [\alpha_n + \beta_n \mathrm{e}^{-\gamma y^2(t-1)}]y(t-n) + D(z)v(t), \qquad (1) \end{aligned}$$

where $v(t)$ is a stochastic white noise with zero mean and variance σ^2, $D(z)$ is the polynomial in the unit backward shift operator z^{-1} $[z^{-1}y(t) = y(t-1)]$:

$$D(z) := 1 + d_1 z^{-1} + d_2 z^{-2} + \cdots + d_m z^{-m}, \quad d_j \in \mathbb{R},$$

and α_i, β_i $(i = 1, 2, \cdots, n)$, γ and d_j $(j = 1, 2, \cdots, m)$ are the parameters to be estimated.

The nonlinearity of the ExpARMA model comes from the exponential dependence on $\gamma y^2(t-1)$. When the parameter γ is large enough, Equation (1) reduces to an autoregressive moving average (ARMA) model with respect to the parameters α_i and d_j; when $\gamma = 0$, the ExpARMA model reduces to an ARMA model with respect to the parameters $(\alpha_i + \beta_i)$ and d_j. Neither of these two ARMA models can describe any nonlinear dynamics. Thus, the parameter γ is essentially a scaling factor, and should be limited in a range so that $\mathrm{e}^{-\gamma y^2(t-1)}$ is different from both zero and one.

Assume that the orders m and n are known, $y(t)$ is measurable. Without loss of generality, the initial values are set to be $y(t) = 0$ and $v(t) = 0$ for $t \leq 0$.

It is obvious that $y(t)$ is linear with respect to the parameters α_i, β_i and d_j, and is nonlinear with respect to the parameter γ. Define the parameter vectors and the information vectors as

$$\begin{aligned} \boldsymbol{\vartheta} &:= [\boldsymbol{\theta}^{\mathrm{T}}, \boldsymbol{d}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n+m}, \\ \boldsymbol{\theta} &:= [\boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n}, \\ \boldsymbol{\alpha} &:= [\alpha_1, \alpha_2, \cdots, \alpha_n]^{\mathrm{T}} \in \mathbb{R}^{n}, \\ \boldsymbol{\beta} &:= [\beta_1, \beta_2, \cdots, \beta_n]^{\mathrm{T}} \in \mathbb{R}^{n}, \\ \boldsymbol{d} &:= [d_1, d_2, \cdots, d_m]^{\mathrm{T}} \in \mathbb{R}^{m}, \\ \boldsymbol{\phi}(\gamma, t) &:= [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\gamma, t), \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n+m}, \\ \boldsymbol{\phi}_{\mathrm{s}}(\gamma, t) &:= [\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n}, \\ \boldsymbol{\varphi}(t) &:= [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}} \in \mathbb{R}^{n}, \\ \boldsymbol{\phi}_{\mathrm{n}}(t) &:= [v(t-1), v(t-2), \cdots, v(t-m)]^{\mathrm{T}} \in \mathbb{R}^{m}. \end{aligned}$$

Then, Equation (1) can be written as

$$\begin{aligned} y(t) &= \sum_{i=1}^{n} \alpha_i y(t-i) + \mathrm{e}^{-\gamma y^2(t-1)} \sum_{i=1}^{n} \beta_i y(t-i) \\ &\quad + \sum_{j=1}^{m} d_j v(t-j) + v(t) \\ &= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\alpha} + \mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\beta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} + v(t) \\ &= \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\gamma, t)\boldsymbol{\theta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} + v(t) \\ &= \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta} + v(t). \qquad (2) \end{aligned}$$

Equation (2) is the identification model of the ExpARMA process. In the field of system identification, the least squares method has been used for certain nonlinear systems. For example, Hafezi and Arefi derived a recursive generalised extended least squares and recursive maximum likelihood algorithm for bilinear systems [37]. Kazemi and Arefi presented an iterative recursive least squares and a robust recursive least squares algorithm for Wiener systems [38]. Moreover, the least squares method has shown the appropriateness in controller design and some other areas [39, 40]. Note that $y(t)$ is highly nonlinear with respect to the parameter γ, the identification problem becomes a complex nonlinear optimization problem and the least squares method cannot be used directly for the ExpARMA model. The objective of this paper is to study and present new recursive identification algorithms for the ExpARMA model by using the gradient search and the multi-innovation identification theory.

3 Extended stochastic gradient algorithm

In this section, applying the gradient search, we derive an extended stochastic gradient (ESG) algorithm and its variant for the ExpARMA model.

For convenience, we let $\boldsymbol{\Theta} := [\boldsymbol{\vartheta}^{\mathrm{T}}, \gamma]^{\mathrm{T}} \in \mathbb{R}^{2n+m+1}$. Then, the parameter vector $\boldsymbol{\Theta}$ involves all the model parameters to be estimated. Define the cost function

$$J_1(\boldsymbol{\Theta}) := \frac{1}{2}[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta}]^2.$$

Computing the gradient of $J(\boldsymbol{\Theta})$ gives

$$\begin{aligned}\mathrm{grad}[J_1(\boldsymbol{\Theta})] &= \begin{bmatrix} \frac{\partial J_1(\boldsymbol{\Theta})}{\partial \boldsymbol{\vartheta}} \\ \frac{\partial J_1(\boldsymbol{\Theta})}{\partial \gamma} \end{bmatrix} \\ &= -\begin{bmatrix} \boldsymbol{\phi}(\gamma, t)[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta}] \\ \boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{\phi}'(\gamma, t)[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta}] \end{bmatrix} \in \mathbb{R}^{2n+m+1},\end{aligned}$$

where $\boldsymbol{\phi}'(\gamma, t)$ is the derivative of $\boldsymbol{\phi}(\gamma, t)$ with respect to γ, i.e.,

$$\begin{aligned}\boldsymbol{\phi}'(\gamma, t) &:= \frac{\partial \boldsymbol{\phi}(\gamma, t)}{\partial \gamma} = \begin{bmatrix} \frac{\partial \boldsymbol{\phi}_{\mathrm{s}}(\gamma, t)}{\partial \gamma} \\ \frac{\partial \boldsymbol{\phi}_{\mathrm{n}}(t)}{\partial \gamma} \end{bmatrix} \\ &= [\mathbf{0}_n, -y^2(t-1)\mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathbf{0}_m]^{\mathrm{T}} \in \mathbb{R}^{2n+m}.\end{aligned}$$

Define the generalised information vector

$$\boldsymbol{\psi}(\boldsymbol{\Theta}, t) := \begin{bmatrix} \boldsymbol{\phi}(\gamma, t) \\ \boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{\phi}'(\gamma, t) \end{bmatrix} \in \mathbb{R}^{2n+m+1}.$$

Then, the gradient of $J_1(\boldsymbol{\Theta})$ can be written as

$$\mathrm{grad}[J_1(\boldsymbol{\Theta})] = -\boldsymbol{\psi}(\boldsymbol{\Theta}, t)[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta}].$$

Let $\hat{\boldsymbol{\Theta}}(t) := [\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t), \hat{\gamma}(t)]^{\mathrm{T}} \in \mathbb{R}^{2n+m+1}$ denote the estimate of $\boldsymbol{\Theta}$ at time t. Using the gradient search to minimize the cost function $J_1(\boldsymbol{\Theta})$, we obtain the following recursive algorithm,

$$\begin{aligned}\hat{\boldsymbol{\Theta}}(t) &= \hat{\boldsymbol{\Theta}}(t-1) - \mu(t)\mathrm{grad}[J_1(\hat{\boldsymbol{\Theta}}(t-1))] \\ &= \hat{\boldsymbol{\Theta}}(t-1) + \mu(t)\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t) \\ &\quad \times [y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1)], \qquad (3)\end{aligned}$$

where $\mu(t)$ represents the step-size of the recursive algorithm. One method of determining $\mu(t)$ is to apply the one-dimensional search, i.e., to solve the optimization problem

$$\mu(t) = \arg\min_{\mu(t) \geq 0} J_1[\hat{\boldsymbol{\Theta}}(t)].$$

Let $g[\mu(t)] := J_1[\hat{\boldsymbol{\Theta}}(t)]$ and define the innovation $e(t) := y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1) \in \mathbb{R}$. Substituting (3) into $J_1(\boldsymbol{\Theta})$ gives

$$\begin{aligned}g[\mu(t)] &= \frac{1}{2}[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t), t)\hat{\boldsymbol{\vartheta}}(t)]^2 \\ &= \frac{1}{2}\{y(t) - \boldsymbol{\phi}^{\mathrm{T}}[\hat{\gamma}(t-1) \\ &\quad + \mu(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)e(t), t] \\ &\quad \times [\hat{\boldsymbol{\vartheta}}(t-1) + \mu(t)\boldsymbol{\phi}(\hat{\gamma}(t-1), t)e(t)]\}^2.\end{aligned}$$

Substituting the first-order Taylor expansion of $\boldsymbol{\phi}(\gamma, t)$ at $\gamma = \hat{\gamma}(t-1)$ into the above equation, we have

$$\begin{aligned}g[\mu(t)] &= \frac{1}{2}\{y(t) - [\boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t) + [\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)]^{\mathrm{T}} \\ &\quad \times [\hat{\gamma}(t) - \hat{\gamma}(t-1)] + o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]]\hat{\boldsymbol{\vartheta}}(t)\}^2 \\ &= \frac{1}{2}\{y(t) - [\boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t) + [\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)]^{\mathrm{T}} \\ &\quad \times [\mu(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)e(t)] \\ &\quad + o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]]\hat{\boldsymbol{\vartheta}}(t)\}^2 \\ &= \frac{1}{2}\{y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t)[\hat{\boldsymbol{\vartheta}}(t-1) \\ &\quad + \mu(t)\boldsymbol{\phi}(\hat{\gamma}(t-1), t)e(t)] \\ &\quad - [\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)]^{\mathrm{T}}[\mu(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)e(t)] \\ &\quad \times [\hat{\boldsymbol{\vartheta}}(t-1) + \mu(t)\boldsymbol{\phi}(\hat{\gamma}(t-1), t)e(t)] \\ &\quad - o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]\}^2 \\ &= \frac{1}{2}\{[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1)] \\ &\quad - e(t)\|\boldsymbol{\phi}(\hat{\gamma}(t-1), t)\|^2\mu(t) \\ &\quad - e(t)\|\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2\mu(t) \\ &\quad - e^2(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}(\hat{\gamma}(t-1), t)\|\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2\mu^2(t) \\ &\quad - o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]\}^2 \\ &= \frac{1}{2}e^2(t)\{1 - [\|\boldsymbol{\phi}(\hat{\gamma}(t-1), t)\|^2 \\ &\quad + \|\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2]\mu(t) \\ &\quad - e(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}(\hat{\gamma}(t-1), t) \\ &\quad \times \|\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2\mu^2(t) - o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]\}^2 \\ &= \frac{1}{2}e^2(t)[1 - \|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2\mu(t) - \xi(t)\mu^2(t)]^2 \\ &\quad + o[\hat{\gamma}(t) - \hat{\gamma}(t-1)]^2,\end{aligned}$$

where

$$\xi(t) := e(t)\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}(\hat{\gamma}(t-1), t)\|\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2 \in \mathbb{R}.$$

Minimizing $g[\mu(t)]$ is equal to solving the equation

$$1 - \|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2\mu(t) - \xi(t)\mu^2(t) = 0.$$

Thus, we have

$$\begin{aligned}\mu(t) &= \frac{\sqrt{\|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^4 + 4\xi(t)} - \|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2}{2\xi(t)} \\ &= \frac{2}{\sqrt{\|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^4 + 4\xi(t)} + \|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2}. \qquad (4)\end{aligned}$$

Equation (4) for computing the step-size $\mu(t)$ is complicated. Referring to the method of finding the optimal step-size for Hammerstein nonlinear systems in [41], Equation (4) can be simplified as

$$\mu(t) = \frac{1}{\|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2}.$$

In order to avoid the denominator being zero, we take the step-size as

$$\mu(t) = \frac{1}{1 + \|\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)\|^2}. \qquad (5)$$

Since the generalised information vector $\boldsymbol{\psi}(\hat{\boldsymbol{\Theta}}(t-1), t)$ in (3) and (5) and the information vector $\boldsymbol{\phi}(\hat{\gamma}(t-1), t)$ in (3) involve the unmeasurable noise $v(t-j)$, Equations (3) and (5) cannot be directly used to compute the estimate $\hat{\boldsymbol{\Theta}}(t)$. To address this problem, we replace $v(t-j)$ with the estimate $\hat{v}(t-j)$. From (2), we have

$$v(t)=y(t)-\boldsymbol{\phi}^{\mathrm{T}}(\gamma,t)\boldsymbol{\vartheta}.$$

Replacing $\boldsymbol{\vartheta}$ and γ in the above equation with $\hat{\boldsymbol{\vartheta}}(t)$ and $\hat{\gamma}(t-1)$ gives

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t), \tag{6}$$

where

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)=[\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\hat{\gamma}(t-1),t),\hat{\boldsymbol{\phi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{7}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)=[\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-m)]^{\mathrm{T}}. \tag{8}$$

Replacing $\boldsymbol{\vartheta}$ and γ with $\hat{\boldsymbol{\vartheta}}(t-1)$ and $\hat{\gamma}(t-1)$, the estimate of the generalised information vector $\boldsymbol{\psi}(\boldsymbol{\Theta},t)$ can be computed by

$$\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)=\begin{bmatrix}\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)\\ \hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\phi}}'(\hat{\gamma}(t-1),t)\end{bmatrix}. \tag{9}$$

Substituting (6)–(9) into (3)–(4), we have

$$\begin{aligned}\hat{\boldsymbol{\Theta}}(t)&=\hat{\boldsymbol{\Theta}}(t-1)+\mu(t)\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\\ &\quad\times[y(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1)],\end{aligned} \tag{10}$$

$$\mu(t)=\frac{1}{1+\|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2}. \tag{11}$$

Equations (10)–(11) and (6)–(9) form the projection algorithm for the ExpARMA model. The projection algorithm is sensitive to the noise since the gain vector $\mu(t)\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)=\frac{\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)}{1+\|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2}$ does not approach zero. In order to adjust the gain vector of the projection algorithm, we take the step-size to be $\mu(t)=\frac{1}{r(t)}$ and summarise the following recursive algorithm,

$$\hat{\boldsymbol{\Theta}}(t)=\hat{\boldsymbol{\Theta}}(t-1)+\frac{1}{r(t)}\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)e(t), \tag{12}$$

$$e(t)=y(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{13}$$

$$r(t)=r(t-1)+\|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2, \tag{14}$$

$$\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)=\begin{bmatrix}\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)\\ \hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1),t)\end{bmatrix}, \tag{15}$$

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)=[\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\hat{\gamma}(t-1),t),\hat{\boldsymbol{\phi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{16}$$

$$\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t)=[\boldsymbol{\varphi}^{\mathrm{T}}(t),\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{17}$$

$$\boldsymbol{\varphi}(t)=[y(t-1),y(t-2),\cdots,y(t-n)]^{\mathrm{T}}, \tag{18}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)=[\hat{v}(t-1),\hat{v}(t-2),\cdots,\hat{v}(t-m)]^{\mathrm{T}}, \tag{19}$$

$$\boldsymbol{\phi}'(\hat{\gamma}(t-1),t)=[\mathbf{0}_n,-y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t),\mathbf{0}_m]^{\mathrm{T}}, \tag{20}$$

$$\hat{v}(t)=y(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t), \tag{21}$$

$$\hat{\boldsymbol{\Theta}}(t)=[\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t),\hat{\boldsymbol{\beta}}^{\mathrm{T}}(t),\hat{\boldsymbol{d}}^{\mathrm{T}}(t),\hat{\gamma}(t)]^{\mathrm{T}}. \tag{22}$$

Equations (12)–(22) form the ESG algorithm for the ExpARMA model.

Remark 1: In (12) and (14), $r(t)$ represents the reciprocal of the step-size $\mu(t)$. Compared with the denominator of the gain vector $\frac{\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)}{1+\|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2}$ in the projection algorithm, the denominator of $\frac{1}{r(t)}\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)$ in the ESG algorithm involves both the current and the preceding $(t-1)$ information, which makes the new gain vector approach zero.

Remark 2: In order to improve the parameter estimation accuracy, we introduce a forgetting factor (FF) λ into (14), i.e.

$$r(t)=\lambda r(t-1)+\|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2,\quad 0<\lambda<1. \tag{23}$$

Replacing (14) in the ESG algorithm with (23), we obtain the forgetting factor extended stochastic gradient (FF-ESG) algorithm for the ExpARMA model. When $\lambda=1$, the FF-ESG degenerates into the ESG algorithm.

To summarise, we list the steps for computing the FF-ESG parameter estimation vector $\hat{\boldsymbol{\Theta}}(t)$ as follows.

1. To initialise, let $t=1$, $\hat{\boldsymbol{\Theta}}(0)=\mathbf{1}_{2n+m+1}/p_0$, $\hat{v}(t-j)=1/p_0, j=1,2,\cdots,m$, $r(0)=1$, $p_0=10^6$, give an error tolerance $\varepsilon>0$.
2. Collect the measurement data $y(t)$, form the information vectors $\boldsymbol{\varphi}(t)$ and $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ by (18) and (19).
3. Compute the information vector $\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t)$ by (17), form the information vector $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)$ by (16).
4. Compute the derivative vector $\boldsymbol{\phi}'(\hat{\gamma}(t-1),t)$ by (20), form the generalised information vector $\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)$ by (15).
5. Compute the innovation $e(t)$ and the reciprocal of the step-size $r(t)$ by (13) and (23).
6. Update the parameter estimation vector $\hat{\boldsymbol{\Theta}}(t)$ by (12), read out $\hat{\boldsymbol{\alpha}}(t)$, $\hat{\boldsymbol{\beta}}(t)$, $\hat{\boldsymbol{d}}(t)$ and $\hat{\gamma}(t)$ from $\hat{\boldsymbol{\Theta}}(t)$ in (22).
7. Compute the noise estimate $\hat{v}(t)$ by (21).
8. Compare $\hat{\boldsymbol{\Theta}}(t)$ with $\hat{\boldsymbol{\Theta}}(t-1)$: if $\|\hat{\boldsymbol{\Theta}}(t)-\hat{\boldsymbol{\Theta}}(t-1)\|>\varepsilon$, increase t by 1 and go to Step 2; otherwise, terminate this procedure.

Multi-innovation identification is an important branch of system identification and has been widely used in parameter estimation for many systems, such as bilinear state space systems, multivariate systems and so on. Expanding the scalar innovation into a multi-dimensional innovation vector, not only the current observation and innovation, but also the preceding observations and innovations are included in identification algorithms. The innovation $e(t)$ of the ESG algorithm is a scalar. Based on the multi-innovation identification theory, the scalar innovation can be expanded into an innovation vector, and the resulting multi-innovation identification algorithm has improved convergence rate and parameter estimation accuracy [42–44].

4 Multi-innovation extended stochastic gradient algorithm

The innovation is the useful information which can improve the parameter estimation accuracy. The traditional identification methods for scalar systems contains the single-innovation which is a scalar. The multi-innovation identification is the innovation expansion based identification [42–45]. That is, expanding the scalar innovation of the traditional identification method into an innovation vector, the resulting algorithm uses more innovations and has an improved parameter estimation accuracy. In this section, the innovation $e(t)$ in (13) is expanded into an innovation vector, and thus a multi-innovation extended stochastic gradient (MI-ESG) algorithm and its variant are derived for the ExpARMA model. The details are as follows.

Expanding the innovation $e(t)$ into the innovation vector

$$\boldsymbol{E}(p,t):=\begin{bmatrix}y(t)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1)\\ y(t-1)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t-1)\hat{\boldsymbol{\vartheta}}(t-1)\\ \vdots\\ y(t-p+1)-\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t-p+1)\hat{\boldsymbol{\vartheta}}(t-1)\end{bmatrix}\in\mathbb{R}^p, \tag{24}$$

where p is the innovation length. Define the stacked vector $\boldsymbol{Y}(p,t)$ and the stacked information matrix $\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t)$ as

$$\boldsymbol{Y}(p,t) := \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-p+1) \end{bmatrix} \in \mathbb{R}^{p}, \tag{25}$$

$$\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t) := \begin{bmatrix} \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t) \\ \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t-1) \\ \vdots \\ \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t-p+1) \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{(2n+m)\times p}. \tag{26}$$

Then, the innovation vector $\boldsymbol{E}(p,t)$ in (24) can be written as

$$\boldsymbol{E}(p,t) = \boldsymbol{Y}(p,t) - \hat{\boldsymbol{\Phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1). \tag{27}$$

In order to guarantee the compatibility of the matrix multiplication dimension, the generalised information vector $\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)$ in (12) must be expanded into a generalised stacked information matrix

$$\hat{\boldsymbol{\Psi}}(p,t) := \begin{bmatrix} \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\hat{\boldsymbol{\Theta}}(t-1),t) \\ \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\hat{\boldsymbol{\Theta}}(t-1),t-1) \\ \vdots \\ \hat{\boldsymbol{\psi}}^{\mathrm{T}}(\hat{\boldsymbol{\Theta}}(t-1),t-p+1)] \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{(2n+m+1)\times p}. \tag{28}$$

Applying the multi-innovation identification theory [42], we expand the scalar innovation $e(t)$ in (12) into the innovation vector $\boldsymbol{E}(p,t)$ and the information vector $\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)$ into the information matrix $\hat{\boldsymbol{\Psi}}(p,t)$, and obtain

$$\hat{\boldsymbol{\Theta}}(t) = \hat{\boldsymbol{\Theta}}(t-1) + \frac{1}{r(t)}\hat{\boldsymbol{\Psi}}(p,t)\boldsymbol{E}(p,t). \tag{29}$$

Combining (25)–(29) and (14)–(22), we obtain the MI-ESG algorithm for the ExpARMA model:

$$\hat{\boldsymbol{\Theta}}(t) = \hat{\boldsymbol{\Theta}}(t-1) + \frac{1}{r(t)}\hat{\boldsymbol{\Psi}}(p,t)\boldsymbol{E}(p,t), \tag{30}$$

$$\boldsymbol{E}(p,t) = \boldsymbol{Y}(p,t) - \hat{\boldsymbol{\Phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{31}$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2, \tag{32}$$

$$\boldsymbol{Y}(p,t) = [y(t), y(t-1), \cdots, y(t-p+1)]^{\mathrm{T}}, \tag{33}$$

$$\begin{aligned}\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t) = [&\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t), \hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t-1), \cdots, \\ &\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t-p+1)],\end{aligned} \tag{34}$$

$$\begin{aligned}\hat{\boldsymbol{\Psi}}(p,t) = [&\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t), \hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t-1), \cdots, \\ &\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t-p+1)],\end{aligned} \tag{35}$$

$$\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t) = \begin{bmatrix} \hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t) \\ \hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1),t) \end{bmatrix}, \tag{36}$$

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t) = [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\hat{\gamma}(t-1),t), \hat{\boldsymbol{\phi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{37}$$

$$\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t) = [\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{38}$$

$$\boldsymbol{\varphi}(t) = [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}}, \tag{39}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t) = [\hat{v}(t-1), \hat{v}(t-2), \cdots, \hat{v}(t-m)]^{\mathrm{T}}, \tag{40}$$

$$\boldsymbol{\phi}'(\hat{\gamma}(t-1),t) = [\mathbf{0}_n, -y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathbf{0}_m]^{\mathrm{T}}, \tag{41}$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t), \tag{42}$$

$$\hat{\boldsymbol{\Theta}}(t) = [\hat{\boldsymbol{\alpha}}^{\mathrm{T}}(t), \hat{\boldsymbol{\beta}}^{\mathrm{T}}(t), \hat{\boldsymbol{d}}^{\mathrm{T}}(t), \hat{\gamma}(t)]^{\mathrm{T}}. \tag{43}$$

When $p=1$, the MI-ESG algorithm reduces to the ESG algorithm.

Remark 3: Similarly, introducing a forgetting factor (FF) λ into (32) gives

$$r(t) = \lambda r(t-1) + \|\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)\|^2, \quad 0<\lambda<1. \tag{44}$$

Replacing (32) in the MI-ESG algorithm with (44), we obtain the forgetting factor multi-innovation extended stochastic gradient (FF-MI-ESG) algorithm for the ExpARMA model. When $\lambda=1$, the FF-MI-ESG degenerates into the MI-ESG algorithm. The methods proposed in this paper can be extended to study the parameter estimation problems of different systems with colored noises [46–49] such as signal modeling and communication networked systems [50–54].

The process of computing $\hat{\boldsymbol{\Theta}}(t)$ by the FF-MI-ESG algorithm is summarised as follows.

1. Choose the innovation length p and initialise: let $t=1$, $\hat{\boldsymbol{\Theta}}(0) = \mathbf{1}_{2n+m+1}/p_0$, $\hat{v}(t-j) = 1/p_0$, $j=1,2,\cdots,m$, $r(0)=1$, $p_0 = 10^6$, give an error tolerance $\varepsilon > 0$.
2. Collect the measurement data $y(t)$, form the stacked vector $\boldsymbol{Y}(p,t)$ by (33), and the information vectors $\boldsymbol{\varphi}(t)$ and $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ by (39) and (40).
3. Compute the information vector $\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t)$ by (38), form the information vector $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)$ by (37).
4. Compute the derivative vector $\hat{\boldsymbol{\phi}}'(\hat{\gamma}(t-1),t)$ by (41), form the generalised information vector $\hat{\boldsymbol{\psi}}(\hat{\boldsymbol{\Theta}}(t-1),t)$ by (36).
5. Form the stacked information matrixes $\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t)$ and $\hat{\boldsymbol{\Psi}}(p,t)$ by (34) and (35).
6. Compute the innovation $\boldsymbol{E}(p,t)$ and the reciprocal of the step-size $r(t)$ by (31) and (44).
7. Update the parameter estimation vector $\hat{\boldsymbol{\Theta}}(t)$ by (30), read out $\hat{\boldsymbol{\alpha}}(t)$, $\hat{\boldsymbol{\beta}}(t)$, $\hat{\boldsymbol{d}}(t)$ and $\hat{\gamma}(t)$ from $\hat{\boldsymbol{\Theta}}(t)$ in (43).
8. Compute the noise estimate $\hat{v}(t)$ by (42).
9. Compare $\hat{\boldsymbol{\Theta}}(t)$ with $\hat{\boldsymbol{\Theta}}(t-1)$: if $\|\hat{\boldsymbol{\Theta}}(t) - \hat{\boldsymbol{\Theta}}(t-1)\| > \varepsilon$, increase t by 1 and go to Step 2; otherwise, terminate this procedure.

Compared with the traditional single-innovation algorithms, the new recursive algorithm proposed in this paper uses the multi-innovation identification theory to expand the single-innovation into a multi-innovation vector, such that both the current and preceding innovations are employed to estimate the unknown parameters. Thus, the proposed multi-innovation identification algorithm makes full advantage of the identification innovations and has an improved parameter estimation accuracy.

5 Example

Consider the following ExpARMA model

$$\begin{aligned} y(t) &= [\alpha_1 + \beta_1\mathrm{e}^{-\gamma y^2(t-1)}]y(t-1) + \cdots \\ &\quad + [\alpha_n + \beta_n\mathrm{e}^{-\gamma y^2(t-1)}]y(t-n) + D(z)v(t) \\ &= [1.23 + 2.00\mathrm{e}^{-2.76y^2(t-1)}]y(t-1) \\ &\quad + [-0.26 + 1.86\mathrm{e}^{-2.76y^2(t-1)}]y(t-2) \\ &\quad + 0.11v(t-1) - 0.15v(t-2) + v(t). \end{aligned}$$

The parameters to be estimated are

$$\begin{aligned} \boldsymbol{\Theta} &= [\alpha_1, \alpha_2, \beta_1, \beta_2, d_1, d_2, \gamma]^{\mathrm{T}} \\ &= [1.23, -0.26, 2.00, 1.86, 0.11, -0.15, 2.76]^{\mathrm{T}}. \end{aligned}$$

In simulation, the noise $\{v(t)\}$ is taken as a white noise sequence with zero mean and variance σ^2, and the data length is taken as $L_e = 3000$.

To exhibit the advantage of the proposed multi-innovation identification algorithm, taking the variance $\sigma^2 = 0.20^2$ and the forgetting factor $\lambda = 0.98$, we use the FF-ESG algorithm and the FF-MI-ESG algorithm with $p = 10$ and $p = 15$ to identify this ExpARMA model, respectively. The parameter estimates and their errors are shown in Tables 1–3, the parameter estimation errors $\delta := \|\hat{\boldsymbol{\Theta}}(t) - \boldsymbol{\Theta}\|/\|\boldsymbol{\Theta}\| \times 100\%$ against t are shown in Figure 1.

To demonstrate the influence of the forgetting factor on the parameter estimation accuracy, taking $\sigma^2 = 0.20^2$ and the innovation length $p = 15$, we use the MI-ESG algorithm and the FF-MI-ESG algorithm with $\lambda = 0.99$ to identify this ExpARMA model, respectively. The parameter estimates and their errors are shown in Tables 4–5, the parameter estimation errors δ against t are shown in Figure 2.

To show how the performance of the proposed FF-MI-ESG algorithm depends on the noise level, we fix the innovation length $p = 15$ and the forgetting factor $\lambda = 0.98$, and adopt the FF-MI-ESG algorithm with the noise variance $\sigma^2 = 0.23^2$ and $\sigma^2 = 0.26^2$ to identify this ExpARMA model, respectively. The parameter estimates and their errors are shown in Tables 6–7, the parameter estimation errors δ against t are shown in Figure 3.

To show how the FF-MI-ESG estimates fluctuate against t, we set $p = 15$, $\lambda = 0.98$ and $\sigma^2 = 0.20^2$. The identification results are shown in Figure 4.

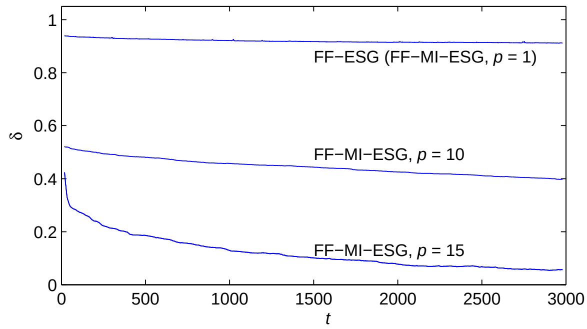

Fig. 1: FF-ESG and FF-MI-ESG errors δ against t ($\lambda = 0.98$, $\sigma^2 = 0.20^2$)

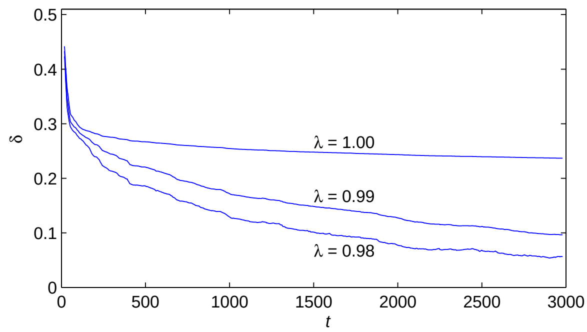

Fig. 2: MI-ESG and FF-MI-ESG errors δ against t ($p = 15$, $\sigma^2 = 0.20^2$)

The vertical axes of Figures 1–3 depict the parameter estimation errors, which are corresponding to the last columns in Tables 1–7, and the abscissa vertical axes depict the recursion numbers, which are corresponding to the first columns in Tables 1–7. The vertical axis of Figure 4 depicts the parameter estimates of the FF-MI-ESG algorithm with an appropriate innovation length and forgetting factor. The following conclusions can be drawn from Figures 1–4 and Tables 1–7.

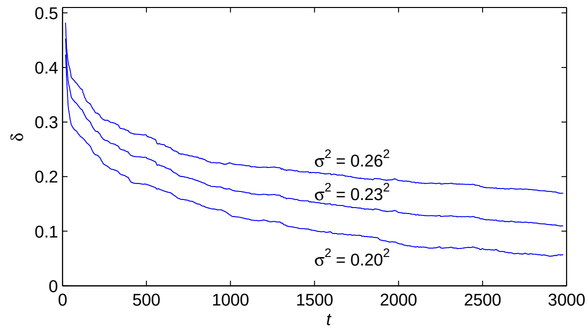

Fig. 3: FF-MI-ESG errors δ against t ($p = 15$, $\lambda = 0.98$)

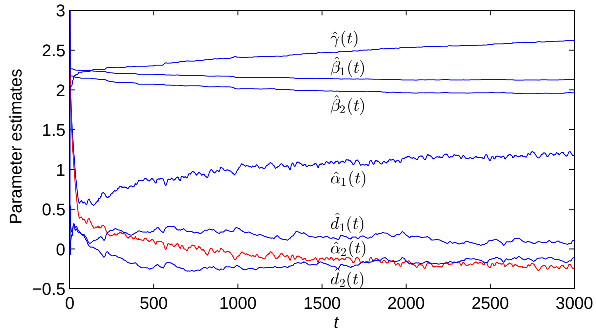

Fig. 4: FF-MI-ESG estimates against t ($p = 15$, $\lambda = 0.98$)

From Figure 1 and Tables 1–3, we find that when $\lambda = 0.98$ and $\sigma^2 = 0.20^2$, the FF-MI-ESG algorithm with $p = 10$ and $p = 15$ has higher parameter estimation accuracy than the FF-ESG algorithm when $p = 1$, and the parameter estimation accuracy becomes higher with the innovation length p increasing.

From Figure 2 and Tables 3–5, we find that when $p = 15$ and $\sigma^2 = 0.20^2$, the FF-MI-ESG algorithm with $\lambda = 0.98$ and $\lambda = 0.99$ has higher parameter estimation accuracy than the MI-ESG algorithm when $\lambda = 1.00$, and the estimation errors of the FF-MI-ESG algorithm become smaller with the decreasing of the forgetting factors.

From Figure 3, Table 3 and Tables 6–7, we find that when $p = 15$ and $\lambda = 0.98$, the parameter estimation errors of the FF-MI-ESG algorithm become smaller with the decreasing of the noise levels.

From Figure 4, we find that under the appropriate forgetting factor and multi-innovation length, the parameter estimates of the FF-MI-ESG algorithm tend to the corresponding true values.

For the model validation, we use the remaining $L_r = 200$ data from $t = L_e + 1$ to $t = L_e + L_r$ and the estimated model obtained by the FF-MI-ESG algorithm with $\lambda = 0.98$ and $p = 15$. The predicted data $\hat{y}(t)$ and the measurement data $y(t)$ are plotted in Figure 5. To evaluate the prediction performance, we define and compute the mean square error (MSE) as

$$MSE := \left[\frac{1}{L_r}\sum_{t=L_e+1}^{L_e+L_r}[\hat{y}(t) - y(t)]^2\right]^{1/2} = 0.21018.$$

From Figure 5, we can see that the predicted data $\hat{y}(t)$ is close to the measurement data $y(t)$, which means the estimated model can capture the dynamics of this ExpARMA model.

Remark 4: In terms of identifying the nonlinear ExpAR model, many literatures suppose that the parameter of the nonlinear part γ is known a priori or imposed on certain conditions, and use the LS estimator to identify the parameters of the linear part [2, 25, 26]. In addition, colored noise was hardly taken into account. Compared with the recursive identification algorithms in these publications, the multi-innovation identification algorithm proposed in this paper

Table 1 FF-ESG parameter estimates and errors ($p = 1$, $\lambda = 0.98$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.23511	0.26681	0.22467	0.25507	0.05204	-0.01804	0.01015	93.45056
200	0.24729	0.27229	0.23630	0.26018	0.05808	-0.02088	0.00584	93.27890
500	0.25797	0.26534	0.24616	0.25358	0.07158	-0.03250	0.02913	92.71970
1000	0.27867	0.24902	0.26516	0.23920	0.10365	-0.04529	0.04968	92.09531
2000	0.32386	0.23368	0.30319	0.22539	0.15347	-0.07135	0.05074	91.46356
3000	0.34902	0.20326	0.32136	0.19726	0.20343	-0.08325	0.05829	91.20649
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

Table 2 FF-MI-ESG parameter estimates and errors ($p = 10$, $\lambda = 0.98$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.60229	0.31075	1.06646	0.92473	0.13237	-0.01259	1.40848	50.82531
200	0.69296	0.28233	1.05284	0.90249	0.18121	-0.08157	1.43258	49.97128
500	0.77458	0.13912	1.03700	0.87213	0.22854	-0.22789	1.51029	48.03716
1000	0.89646	-0.01545	1.01651	0.83556	0.30344	-0.24284	1.63614	45.77020
2000	1.04824	-0.05567	1.02746	0.84350	0.22686	-0.23467	1.78769	42.53559
3000	1.08229	-0.14975	1.05618	0.87543	0.15308	-0.13462	1.90325	39.69558
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

Table 3 FF-MI-ESG parameter estimates and errors ($p = 15$, $\lambda = 0.98$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.55589	0.31925	2.24138	2.14677	0.13644	0.10226	2.22887	27.65778
200	0.71193	0.29624	2.22847	2.12799	0.11448	-0.09881	2.25693	23.95071
500	0.86134	0.09741	2.19618	2.07137	0.22886	-0.23127	2.30289	18.58408
1000	0.99530	-0.09519	2.15640	2.01246	0.26082	-0.21296	2.41099	12.91608
2000	1.15592	-0.15372	2.12482	1.96271	0.17554	-0.15008	2.52964	7.75066
3000	1.17948	-0.23927	2.12790	1.96237	0.10568	-0.11382	2.62146	5.49173
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

Table 4 MI-ESG parameter estimates and errors ($\lambda = 1.00$, $p = 15$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.56564	0.38885	2.23932	2.14695	0.24692	0.24258	2.21467	29.69835
200	0.56407	0.35165	2.23742	2.14416	0.18196	0.14363	2.22996	28.19570
500	0.61867	0.34532	2.23360	2.13678	0.14191	0.04498	2.23204	26.69020
1000	0.62121	0.29184	2.22818	2.12899	0.13780	-0.00736	2.24878	25.46024
2000	0.65855	0.27320	2.22326	2.12159	0.15265	-0.04628	2.26252	24.34341
3000	0.68441	0.26388	2.22105	2.11832	0.15695	-0.07340	2.27010	23.67712
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

Table 5 FF-MI-ESG parameter estimates and errors ($\lambda = 0.99$, $p = 15$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.53333	0.33275	2.24056	2.14728	0.18527	0.17154	2.22441	28.67996
200	0.62984	0.33615	2.23511	2.13915	0.12865	0.02364	2.24480	26.18731
500	0.74715	0.21697	2.21634	2.10550	0.16899	-0.13742	2.26999	22.04848
1000	0.84964	0.02584	2.18560	2.06078	0.26499	-0.21714	2.35584	17.18589
2000	1.02851	-0.02874	2.15867	2.01963	0.27475	-0.20701	2.44950	12.75436
3000	1.06581	-0.12723	2.15354	2.00985	0.19872	-0.18773	2.52776	9.60038
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

Table 6 FF-MI-ESG parameter estimates and errors ($\sigma^2 = 0.23^2$, $p = 15$, $\lambda = 0.98$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.58532	0.26730	2.17817	2.05554	0.16031	0.11100	1.78116	32.76890
200	0.76307	0.21254	2.14495	2.00777	0.18812	-0.08588	1.84323	28.27856
500	0.93646	0.01647	2.08632	1.90471	0.25662	-0.21953	1.91140	23.45756
1000	1.06574	-0.20754	2.00440	1.78370	0.27584	-0.16086	2.08296	17.67674
2000	1.22838	-0.22229	1.94411	1.69432	0.15671	-0.11107	2.24358	13.45933
3000	1.23096	-0.29613	1.94083	1.68448	0.09262	-0.08036	2.36405	10.87852
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

relaxes the conditions on the parameter γ and estimates all the unknown parameters simultaneously.

6 Conclusions

This paper studies the parameter estimation problems of the nonlinear ExpARMA model. Using the gradient search and the multi-innovation identification theory, we derive an extended stochastic gradient (ESG) algorithm and a multi-innovation extended stochastic

Table 7 FF-MI-ESG parameter estimates and errors ($\sigma^2 = 0.26^2$, $p = 15$, $\lambda = 0.98$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	0.61384	0.22498	2.11402	1.96108	0.17009	0.11774	1.53205	36.47859
200	0.80507	0.15333	2.06676	1.89252	0.22326	-0.08515	1.62147	31.63197
500	0.98578	-0.03228	1.99436	1.76350	0.25283	-0.21275	1.70108	27.54511
1000	1.10716	-0.26658	1.88923	1.60417	0.25302	-0.13730	1.90613	22.46679
2000	1.25851	-0.25212	1.81583	1.50057	0.13280	-0.09038	2.08634	19.30154
3000	1.25026	-0.31685	1.81291	1.49080	0.08494	-0.06697	2.21894	16.87283
True values	1.23000	-0.26000	2.00000	1.86000	0.11000	-0.15000	2.76000	

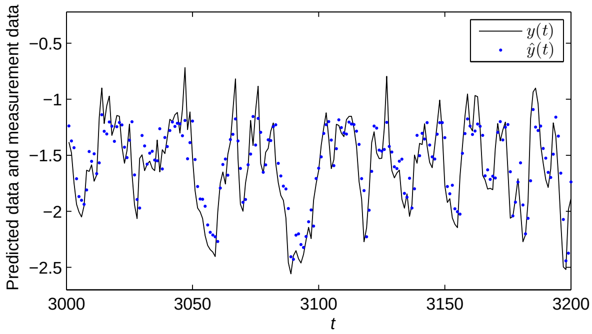

Fig. 5: Predicted data and measurement data

gradient (MI-ESG) algorithm to identify the unknown parameters. Introducing a forgetting factor into the MI-ESG algorithm, we obtain the forgetting factor multi-innovation extended stochastic gradient (FF-MI-ESG) algorithm. The simulation results indicate that with an appropriate innovation length and forgetting factor, the FF-MI-ESG algorithm has improved parameter estimation accuracy and convergence rate than the FF-ESG algorithm, and can be effectively used to identify the nonlinear ExpARMA model. The proposed multi-innovation extended stochastic gradient (MI-ESG) algorithm for the ExpARMA model can combine other estimation algorithms [55–59] to explore new identification methods of linear and nonlinear systems [60–64] and can be applied to other fields such as information processing and communication [65–69].

7 Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant no. 61873111), the 111 Project (B12018), and the National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-26).

8 References

1 Chen, G.Y., Gan, M., Chen, C.L.P., *et al.*: 'A regularized variable projection algorithm for separable nonlinear least squares problems', *IEEE Transactions on Automatic Control*, 2019, **64**, (2), pp. 526-537

2 Chen, G.Y., Gan, M., Chen, G.L.: 'Generalized exponential autoregressive models for nonlinear time series: stationarity, estimation and applications', *Information Sciences*, 2018, **438**, pp. 46-57

3 Ozaki, T.: 'Non-linear time series models for non-linear random vibrations', *Journal of Applied Probability*, 1980, **17**, (1), pp. 84-93

4 Merzougui, M., Dridi, H., Chadli, A.: 'Test for periodicity in restrictive EXPAR models', *Communications in Statistics — Theory and Methods*, 2016, **45**, (9), pp. 2770-2783

5 Ozaki, T.: 'The statistical analysis of perturbed limit cycle processes using nonlinear time series models', *Journal of Time Series Analysis*, 1982, **3**, (1), pp. 29-41

6 Teräsvirta, T.: 'Specification, estimation, and evaluation of smooth transition autoregressive models', *Journal of the American Statistical Association*, 1994, **89**, (425), pp. 208-218

7 Ma, H., Pan, J., Lv, L., *et al.*: 'Recursive algorithms for multivariable output-error-like ARMA systems', *Mathematics*, 2019, **7**, (6), p. 558, Article Number: 558, https://doi.org/10.3390/math7060558

8 Liu, X.D., Yu, H.S., Yu, J.P., Zhao, L.: 'Combined speed and current terminal sliding mode control with nonlinear disturbance observer for PMSM drive', *IEEE Access*, 2018, **6**, pp. 29594-29601

9 Liu, X.D., Yu, H.S., Yu, J.P., Zhao, L.: 'A novel speed control method based on port-controlled Hamiltonian and disturbance observer for PMSM drives', *IEEE Access*, 2019, **7**, pp. 111115-111123

10 Cao, Y., Wang, Z., Liu, F., Li, P., Xie, G.: 'Bio-inspired speed curve optimization and sliding mode tracking control for subway trains', *IEEE Transactions on Vehicular Technology*, 2019, **68**, (7), pp. 6331-6342

11 Cao, Y., Sun, Y.K., Guo, X., Wen, T.: '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*, 2019, **68**, (8), pp. 7544-7551

12 Cao, Y., Lu, H., Wen, T.: 'A safety computer system based on multi-sensor data processing', *Sensors*, 2019, **19**, (4), p. 818, Article Number: 818, https://doi.org/10.3390/s19040818

13 Cao, Y., Zhang, Y., Wen, T., Li, P.: 'Research on dynamic nonlinear input prediction of fault diagnosis based on fractional differential operator equation in high-speed train control system', *Chaos*, 2019, **29**, (1), p. 013130, Article Number: 013130, https://doi.org/10.1063/1.5085397

14 Cao, Y., Li, P., Zhang, Y.: 'Parallel processing algorithm for railway signal fault diagnosis data based on cloud computing,' *Future Generation Computer Systems*, 2018, **88**, pp. 279-283.

15 Zhang, X., Ding, F., Xu, L., *et al.*: 'State filtering-based least squares parameter estimation for bilinear systems using the hierarchical identification principle', *IET Control Theory and Applications*, 2018, **12**, (12), pp. 1704-1713

16 Chen, G.Y., Gan, M., Ding, F., *et al.*: 'Modified Gram-Schmidt method-based variable projection algorithm for separable nonlinear models', *IEEE Transactions on Neural Networks and Learning Systems*, 2019, **30**, (8), pp. 2410-2418

17 Pan, J., Li, W., Zhang, H.P.: '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*, 2018, **16**, (6), pp. 2878-2887

18 Zhou, Z.P., Liu, X.F.: 'State and fault estimation of sandwich systems with hysteresis', *International Journal of Robust and Nonlinear Control*, 2018, **28**, (13), pp. 3974-3986

19 Gu, Y., Chou, Y.X., Liu, J.C., *et al.*: 'Moving horizon estimation for multirate systems with time-varying time-delays', *Journal of the Franklin Institute*, 2019, **356**, (4), pp. 2325-2345

20 Zhang, X., Ding, F., Xu, L., *et al.*: 'Highly computationally efficient state filter based on the delta operator', *International Journal of Adaptive Control and Signal Processing*, 2019, **33**, (6), pp. 875-889

21 Zhang, X., Ding, F., Xu, L., *et al.*: 'A hierarchical approach for joint parameter and state estimation of a bilinear system with autoregressive noise', *Mathematics*, 2019, **7**, (4), p. 356, Article Number: 356, https://doi.org/10.3390/math7040356

22 Li, M.H., Liu, X.M.: 'The least squares based iterative algorithms for parameter estimation of a bilinear system with autoregressive noise using the data filtering technique', *Signal Processing*, 2018, **147**, pp. 23-34

23 Li, M.H., Liu, X.M.: 'Auxiliary model based least squares iterative algorithms for parameter estimation of bilinear systems using interval-varying measurements', *IEEE Access*, 2018, **6**, pp. 21518-21529

24 Li, M.H., Liu, X.M., Ding, F.: '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*, 2019, **33**, (7), pp. 1189-1211

25 Haggan, V., Ozaki, T.: 'Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model', *Biometrika*, 1981, **68**, (1), pp. 189-196

26 Shi, Z., Tamura, Y., Ozaki, T.: 'Monitoring the stability of BWR oscillation by nonlinear time series modeling', *Annals of Nuclear Energy*, 2001, **28**, pp. 953-966

27 Xu, H., Ding, F., Sheng, J.: 'On some parameter estimation algorithms for the nonlinear exponential autoregressive model', *International Journal of Adaptive Control and Signal Processing*, 2019, **33**, (6), pp. 999-1015

28 Xu, H., Ding, F., Yang, E.F.: 'Modeling a nonlinear process using the exponential autoregressive time series model', *Nonlinear Dynamics*, 2019, **95**, (3), pp. 2079-2092

29 Ding, J., Chen, J.Z., Lin, J.X., *et al.*: 'Particle filtering based parameter estimation for systems with output-error type model structures', *Journal of the Franklin Institute*, 2019, **356**, (10), pp. 5521-5540

30 Ding, J., Chen, J.Z., Lin, J.X., Jiang, G.P.: 'Particle filtering-based recursive identification for controlled auto-regressive systems with quantised output', *IET Control Theory and Applications*, 2019, **13**, (14), pp. 2181-2187.

31 Xu, L.: 'The parameter estimation algorithms based on the dynamical response measurement data', *Advances in Mechanical Engineering*, 2017, **9**, (11), pp. 1-12. https://doi.org/10.1177/1687814017730003

32 Xu, L., Xiong, W.L., Alsaedi, A., *et al.*: 'Hierarchical parameter estimation for the frequency response based on the dynamical window data', *International Journal of Control Automation and Systems*, 2018, **16**, (4), pp. 1756-1764

33 Zhang, X., Xu, L., Ding, F., *et al.*: 'Combined state and parameter estimation for a bilinear state space system with moving average noise', *Journal of the Franklin Institute*, 2018, **355**, (6), pp. 3079-3103

34 Ding, F., Pan, J., Alsaedi, A., *et al.*: 'Gradient-based iterative parameter estimation algorithms for dynamical systems from observation data', *Mathematics*, 2019, **7**, (5), p. 428, Article Number: 428, https://doi.org/10.3390/math7050428
35 Xu, L., Ding, F.: 'Iterative parameter estimation for signal models based on measured data', *Circuits Systems and Signal Processing*, 2018, **37**, (7), pp. 3046-3069
36 Ding, F., Xu, L., Alsaadi, F.E., *et al.*: 'Iterative parameter identification for pseudo-linear systems with ARMA noise using the filtering technique', *IET Control Theory and Applications*, 2018, **12**, (7), pp. 892-899
37 Hafezi, Z., Arefi, M.M.: 'Recursive generalized extended least squares and RML algorithms for identification of bilinear systems with ARMA noise', *ISA Transactions*, 2019, **88**, pp. 50-61
38 Kazemi, M., Arefi, M.M.: 'A fast iterative recursive least squares algorithm for Wiener model identification of highly nonlinear systems', *ISA Transactions*, 2017, **67**, pp. 382-388
39 Kazemi, M., Arefi, M.M., Alipouri, Y.: 'Wiener model based GMVC design considering sensor noise and delay', *ISA Transactions*, 2019, **88**, pp. 73-81
40 Kazemi, M., Arefi, M.M.: 'Nonlinear generalized minimum variance control and control performance assessment of nonlinear systems based on a Wiener model', *Transactions of the Institute of Measurement and Control*, 2018, **40**, (5), pp. 1538-1553
41 Ding, F., Liu, X.P., Liu, G.: 'Identification methods for Hammerstein nonlinear systems', *Digital Signal Processing*, 2011, **21**, (2), pp. 215-238
42 Wan, L.J., Ding, F.: 'Decomposition-based gradient iterative identification algorithms for multivariable systems using the multi-innovation theory', *Circuits Systems and Signal Processing*, 2019, **38**, (7), pp. 2971-2991
43 Xu, L., Ding, F., Gu, Y., *et al.*: 'A multi-innovation state and parameter estimation algorithm for a state space system with d-step state-delay', *Signal Processing*, 2017, **140**, pp. 97-103
44 Xu, H., Wan, L.J., Ding, F., *et al.*: 'Fitting the exponential autoregressive model through recursive search', *Journal of the Franklin Institute*, 2019, **356**, (11), pp. 5801-5818
45 Pan, J., Jiang, X., Wan, X.K., *et al.*: 'A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems', *International Journal of Control Automation and Systems*, 2017, **15**, (3), pp. 1189-1197
46 Gong, P.C., Wang, W.Q., Wan, X.R.: 'Adaptive weight matrix design and parameter estimation via sparse modeling for MIMO radar', *Signal Processing*, 2017, **139**, pp. 1–11
47 Gong, P.C., Wang, W.Q., Li, F.C., Cheung, H.: 'Sparsity-aware transmit beamspace design for FDA-MIMO radar', *Signal Processing*, 2018, **144**, pp. 99–103
48 Wan, X.K., Li, Y., Xia, C., *et al.*: 'A T-wave alternans assessment method based on least squares curve fitting technique', *Measurement*, 2016, **86**, pp. 93–100
49 Zhao, N.: 'Joint Optimization of cooperative spectrum sensing and resource allocation in multi-channel cognitive radio sensor networks', *Circuits Systems and Signal Processing*, 2016, **35**, (7), pp. 2563–2583
50 Zhao, N., Wu, M.H., Chen, J.J.: 'Android-based mobile educational platform for speech signal processing', *International Journal of Electrical Engineering Education* , 2017, **54**, (1), pp. 3–16
51 Zhao, N., Liang, Y., Pei, Y.: 'Dynamic contract incentive mechanism for cooperative wireless networks', *IEEE Transactions on Vehicular Technology*, 2018, **67**, (11), pp. 10970–10982
52 Zhao, X.L., Liu, F., Fu, B., Na, F.: 'Reliability analysis of hybrid multi-carrier energy systems based on entropy-based Markov model', *Proceedings of the Institution of Mechanical Engineers Part O-Journal of Risk and Reliability*, 2016, **230**, (6), pp. 561–569
53 Zhao, X.L., Lin, Z.Y., Fu, B., He, L., Na, F.: 'Research on automatic generation control with wind power participation based on predictive optimal 2-degree-of-freedom PID strategy for multi-area interconnected power system', *Energies*, 2018, **11**, (12), p. 3325, Article Number: 3325, http://dx.doi.org/10.3390/en11123325
54 Wang, L., Liu, H., Dai, L.V,, Liu, Y.W.: 'Novel method for identifying fault location of mixed lines', *Energies*, 2018, **11**, (6), p. 1529, Article Number: 1529, https://doi.org/10.3390/en11061529
55 Wan, L.J., Liu, X.M., Ding, F., Chen, C.P.: 'Decomposition least-squares-based iterative identification algorithms for multivariable equation-error autoregressive moving average systems. *Mathematics*, 2019, **7**, (7), p. 609, Article Number: 609, https://doi.org/10.3390/math7070609
56 Zhang, X., Ding, F., Yang, E.F.: 'State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors', *International Journal of Adaptive Control and Signal Processing*, 2019, **33**, (7), pp. 1157-1173
57 Liu, S.Y., Ding, F., Xu, L., et al.: 'Hierarchical principle-based iterative parameter estimation algorithm for dual-frequency signals', *Circuits Systems and Signal Processing*, 2019, **38**, (7), pp. 3251-3268
58 Yang, F., Sun, Y.R., Li, X.X., Huang, C.Y.: 'The quasi-boundary value method for identifying the initial value of heat equation on a columnar symmetric domain', *Numerical Algorithms*, 2019, **82**, (2), pp. 623-639
59 Yang, F., Wang, N, Li, X.X., Huang, C.Y.: 'A quasi-boundary regularization method for identifying the initial value of time-fractional diffusion equation on spherically symmetric domain', *Journal of Inverse and Ill-posed Problems*, 2019, **27**, (5), p. 609, https://doi.org/10.1515/jiip-2018-0050
60 Fu, B., Ouyang, C.X., Li, C.S., Wang, J.W., Gul, E.: 'An improved mixed integer linear programming approach based on symmetry diminishing for unit commitment of hybrid power system', *Energies*, 2019, **12**, (5), p. 833, Article Number: 833, https://doi.org/10.3390/en12050833
61 Shi, W.X., Liu, N., Zhou, Y.M., Cao, X.A.: 'Effects of postannealing on the characteristics and reliability of polyfluorene organic light-emitting diodes', *IEEE Transactions on Electron Devices*, 2019, **66**, (2), pp. 1057–1062
62 Liu, N., Mei, S., Sun, D., Shi, W., Feng, J., Zhou, Y.M., Mei, F., Xu, J., Jiang, Y., Cao, X.A.: 'Effects of charge transport materials on blue fluorescent organic light-emitting diodes with a host-dopant system', *Micromachines*, 2019, **10**, (5), p. 344, Article Number: 344, https://doi.org/10.3390/mi10050344
63 Wu, T.Z., Shi, X., Liao, L., Zhou, C.J., Zhou, H., Su, Y.H.: 'A capacity configuration control strategy to alleviate power fluctuation of hybrid energy storage system based on improved particle swarm optimization', *Energies*, 2019, **12**, (4), p. 642, Article Number: 642, https://doi.org/10.3390/en12040642
64 Zhao, X.L., Lin, Z.Y., Fu, B., He, L., Li, C.S.: 'Research on the predictive optimal PID plus second order derivative method for AGC of power system with high penetration of photovoltaic and wind power', *Journal of Electrical Engineering and Technology*, 2019, **14**, (3), pp. 1075–1086
65 Ma, F.Y., Yin, Y.K., Li, M.: 'Start-up process modelling of sediment microbial fuel cells based on data driven', *Mathematical Problems in Engineering*, 2019, **2019**, p. 10, Article Number: 7403732, https://doi.org/10.1155/2019/7403732
66 Sun, Z.Y., Zhang, D., Meng, Q., Chen, C.C.: 'Feedback stabilization of time-delay nonlinear systems with continuous time-varying output function', *International Journal of Systems Science*, 2019, **50**, (2), pp. 244-255
67 Wei, W., Xue, W.C., Li, D.H.: 'On disturbance rejection in magnetic levitation', *Control Engineering Practice*, 2019, **82**, pp. 24-35.
68 Liu, J.C. Liu, Gu, Y., Chou, Y.X.: 'Seismic data reconstruction via complex shearlet transform and block coordinate relaxation', *Journal of Seismic Exploration*, 2019, **28**, (4), pp. 307-332
69 Zhan, X.S., Cheng, L.L., Wu, J., Yan, H.C.: 'Modified tracking performance limitation of networked time-delay systems with two-channel constraints', *Journal of the Franklin Institute*, 2019, **356**, (12), pp. 6401-6418