Two-stage recursive identification algorithms for a class of nonlinear time series models with colored noise

Huan Xu[1], Feng Ding[1†], Min Gan[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, China*
[2]*College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350116, China*
[3]*Department of Design, Manufacturing and Engineering Management, University of Strathclyde, Glasgow G1 1XJ, Scotland, UK*

SUMMARY

This paper concentrates on the recursive identification algorithms for the exponential autoregressive model with moving average noise. Using the decomposition technique, we transform the original identification model into a linear and nonlinear sub-identification model and derive a two-stage least squares extended stochastic gradient algorithm. In order to improve the parameter estimation accuracy, we employ the multi-innovation identification theory and develop a two-stage least squares multi-innovation extended stochastic gradient algorithm. A simulation example is provided to test the effectiveness of the proposed algorithms. Copyright © 2020 John Wiley & Sons, Ltd.

1. INTRODUCTION

System modeling and parameter estimation are the basis of investigating natural sciences by using mathematical equations in theory [1, 2, 3, 4]. Time series modeling is necessary for the analysis and prediction of random processes in the field of economics, biology, power load scheduling and others [5]. During the past four decades, time series analysis and modeling have been deviating from the linear paradigm with the aim of revealing nonlinearities. The exponential autoregressive (ExpAR) family is an important kind of nonlinear time series models, and has shown the appropriateness in capturing certain nonlinear phenomena, such as amplitude-dependent frequency, jump phenomena and limit cycles [6, 7]. In the early stage, the classic ExpAR model is derived to model a series of actual ship rolling data [8]. Later, Ozaki derived a variant of the ExpAR model by using the Hermite type polynomials, such that a more sophisticated specification for the dynamics of the characteristic roots of autoregressive models can be obtained [9]; introducing a time-delay and a scalar parameter, Teräsvirta developed a different variant of the ExpAR model [10]. Recently, Chen et al summarized other generalized ExpAR models and discussed the stationary conditions of these models [11].

  System identification and parameter estimation are widely used to establish mathematical models of dynamical systems [12, 13], and are basic for adaptive control and fault diagnosis [14] and so on. In the field of system identification, much attention has been focused on various identification algorithms, including the least squares (LS) methods [15], the Kalman filter [16] and the subspace

---

*Correspondence to: Feng Ding, School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China.
†Email: fding@jiangnan.edu.cn

identification [17], etc. The LS methods are suitable for linear parameter system identification and can also deal with the identification problems for the nonlinear ExpAR model under certain special conditions. For example, fixing the nonlinear subsystem parameter at one of a grid of values [7] or imposing it on certain definition [18], the LS estimates of the parameters of linear subsystem can be derived for the ExpAR model. However, the LS method expenses large computational cost and cannot be directly used to estimate all the unknown parameters because the nonlinearity exists in the ExpAR family. Gradient search based method has showed huge superiority in nonlinear system identification. For instance, Xu developed several gradient algorithms for sine response signals using the moving window data [19, 20, 21].

Over the years, some efforts have been advanced to improve the performance of identification algorithms. For example, the decomposition technique has been applied to linear and nonlinear dynamical systems [22]. This paper combines the decomposition technique and the multi-innovation identification theory to derive new methods for nonlinear exponential time series models. The key of the decomposition based algorithms is to divide the original identification model into several sub-identification (Sub-ID) models, such that the parameters of different Sub-ID models can be estimated separately and the computational efficiency of the identification algorithms can be improved. In this aspect, a decomposition based least squares algorithm was derived for bilinear state-space systems [23]. In addition, the multi-innovation identification has been developed as an indispensable branch in the field of system identification. The innovation is known as the useful information that can improve the parameter estimation accuracy. Multi-innovation identification is on the basis of innovation expansion, in other words, multi-innovation identification is to expand the dimension of the innovation and to enhance the data utilization. Using the multi-innovation identification theory, a weighted multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems [24]. Other techniques such as the particle filter [25] can be applied to model some practical systems [26] and to solve some engineering problems [27, 28].

This paper completely differs from [29] which applied the decomposition to the linear models and presented a two-stage least squares algorithm while this paper applies the decomposition to nonlinear ExpAR models and presents a new two-stage multi-innovation algorithm. The previous works in [30, 31] are based on the assumptions with white noises. This paper extends the previous works from white noise to colored noises (i.e., correlated noises) by combining the hierarchical principle and the multi-innovation theory. By means of some mathematical tools [32, 33], the proposed two-stage recursive identification algorithms for the nonlinear ExpAR model with colored noise in this paper can be extended to study the identification problems of different systems and can be applied other literatures [34, 35, 36] such as signal modeling and communication networked systems and engineering application systems and so on. The main contributions of this paper are as follows.

- Using the decomposition technique, we divide the ExpARMA model into two subsystems, one of which contains the unknown parameter vector of the linear subsystem and the noise model, and the other contains the unknown parameter of the nonlinear part.
- Considering the different mapping relationships between the subsystem parameters and their corresponding outputs, we adopt the LS principle and gradient search to the optimization problems respectively and derive a two-stage least squares extended stochastic gradient algorithm for the ExpARMA model.
- Applying the multi-innovation identification theory, we develop a two-stage least squares multi-innovation extended stochastic gradient algorithm for the ExpARMA model, such that the parameter estimation accuracy can be improved.

In summary, the rest of this paper is organized as follows. Section 2 describes the identification problem for the ExpARMA model. A two-stage least squares extended stochastic gradient algorithm is derived in Section 3. A two-stage least squares multi-innovation extended stochastic gradient algorithm is developed 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

Let us introduce some notations first.

Symbols	Meaning
$\mathbf{0}$	The zero vector of appropriate size.
$\mathbf{1}_n$	An n-dimensional column vector whose entries are all 1.
I or $\boldsymbol{I}_n$	The identity matrix of appropriate sizes or $n \times n$.
$\boldsymbol{X}^{\mathrm{T}}$	The transpose of the vector or matrix $\boldsymbol{X}$.
$\|\boldsymbol{X}\|$	The norm of a column vector $\boldsymbol{X}$.
$A =: X$	A is defined as X.
$X := A$	A is defined as X.

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

$$y(t) = [\alpha_1 + \beta_1 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-1) + [\alpha_2 + \beta_2 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-2) + \cdots$$
$$+ [\alpha_n + \beta_n \mathrm{e}^{-\gamma y^2(t-1)}]y(t-n) + w(t), \tag{1}$$

where $w(t) := D(z)v(t) \in \mathbb{R}$ is a colored noise and can be regarded as the output of an MA model driven by a stochastic white noise $v(t)$, $D(z)$ 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 identification problem for the ExpAR model with AR or ARMA noise are different from the one which is focused on in this paper and remain to be investigated in the future.

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 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, such that $\mathrm{e}^{-\gamma y^2(t-1)}$ is different from both zero and one.

Assume that the orders m and n are known, the data $y(t)$ 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

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

where subscripts (roman) s and n represent the first letters of the words "system" and "noise", respectively. Then, Equation (1) can be written as

$$y(t) = \sum_{i=1}^{n} \alpha_i y(t-i) + \mathrm{e}^{-\gamma y^2(t-1)} \sum_{i=1}^{n} \beta_i y(t-i) + w(t)$$

$$
\begin{aligned}
&= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\alpha} + \mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\beta} + w(t) \\
&= \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\gamma, t)\boldsymbol{\theta} + w(t),
\end{aligned}
$$

$$
w(t) = \sum_{j=1}^{m} d_j v(t-j) + v(t) = \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} + v(t),
$$

and we obtain the identification model of the ExpARMA process:

$$
\begin{aligned}
y(t) &= \boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\gamma, t)\boldsymbol{\theta} + \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} + v(t) \\
&= \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta} + v(t).
\end{aligned}
\tag{2}
$$

The proposed algorithms in this paper are based on this identification model in (2). Many identification methods are derived based on the identification models of the systems [37, 38, 39, 40, 41] and can be used to estimate the parameters of other linear systems and nonlinear systems [42, 43, 44] and can be applied to fields such as chemical process control systems.

Note that $y(t)$ is highly nonlinear with respect to the parameter γ, the LS method cannot be used directly for the ExpARMA model. The objective of this paper is to study new recursive identification algorithms for the ExpARMA model by using the decomposition technique and the multi-innovation identification theory.

3. THE TWO-STAGE LEAST SQUARES EXTENDED STOCHASTIC GRADIENT ALGORITHM

Considering $y(t)$ is nonlinear with respect to the parameter γ, we decompose the identification model in (2) into two sub-identification (Sub-ID) models by applying the decomposition technique [45]. In this section, we apply the LS principle and gradient search to these two Sub-ID models, respectively, and derive a two-stage least squares extended stochastic gradient algorithm for the ExpARMA model.

3.1. The decomposition of the ExpARMA model

Define the information item $\psi(\boldsymbol{\beta}, t)$ and the intermediate variable $y_1(t)$ as

$$
\begin{aligned}
\psi(\boldsymbol{\beta}, t) &:= \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\beta} \in \mathbb{R}, \\
y_1(t) &:= y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\alpha} - \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} \in \mathbb{R}.
\end{aligned}
$$

From (2), we can see that the ExpARMA model is decomposed into the following fictitious subsystems or Sub-ID models,

$$
S_1 : y(t) = \boldsymbol{\phi}^{\mathrm{T}}(\gamma, t)\boldsymbol{\vartheta} + v(t),
\tag{3}
$$

$$
S_2 : y_1(t) = \psi(\boldsymbol{\beta}, t)\mathrm{e}^{-\gamma y^2(t-1)} + v(t).
\tag{4}
$$

The parameter sets $\boldsymbol{\vartheta}$ and γ in Sub-ID models (3) and (4) contain all the unknown parameters to be estimated. The parameter γ in $\boldsymbol{\phi}(\gamma, t)$ and the parameter vector $\boldsymbol{\beta}$ in $\psi(\boldsymbol{\beta}, t)$ are the associate terms between these two Sub-ID models.

3.2. The least squares and extended stochastic gradient parameter estimation for subsystems

After decomposition, a linear and nonlinear Sub-ID model arises which can be solved by different optimization methods. For simplicity, we first derive the parameter estimation sub-algorithms for $\boldsymbol{\vartheta}$ and γ, respectively. Since these two sub-algorithms contain the unknown associate terms, an interactive identification algorithm is then proposed for the ExpARMA model. The details are as follows.

According to the Sub-ID model in (3), we define a quadratic criterion function

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

Let $\hat{\boldsymbol{\vartheta}}(t)$ denote the estimate of $\boldsymbol{\vartheta}$ at time t. Minimizing $J_1(\boldsymbol{\vartheta})$ by the LS principle, we obtain the least squares sub-algorithm for computing $\hat{\boldsymbol{\vartheta}}(t)$:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \boldsymbol{P}(t)\boldsymbol{\phi}(\gamma,t)[y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma,t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{5}$$

$$\boldsymbol{P}^{-1}(t) = \boldsymbol{P}^{-1}(t-1) + \boldsymbol{\phi}(\gamma,t)\boldsymbol{\phi}^{\mathrm{T}}(\gamma,t), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}, \tag{6}$$

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

$$\boldsymbol{\phi}_{\mathrm{s}}(\gamma,t) = [\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{8}$$

$$\boldsymbol{\varphi}(t) = [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}}, \tag{9}$$

$$\boldsymbol{\phi}_{\mathrm{n}}(t) = [v(t-1), v(t-2), \cdots, v(t-m)]^{\mathrm{T}}. \tag{10}$$

The difficulty of identification is that the information vectors $\boldsymbol{\phi}(\gamma,t)$ and $\boldsymbol{\phi}_{\mathrm{n}}(t)$ in (7) and (10) contain the unmeasurable noise $v(t-j)$. The solution is to apply the hierarchical identification principle, that is, to replace $v(t-j)$ with the estimate $\hat{v}_1(t-j)$ and to construct the estimation vectors $\hat{\boldsymbol{\phi}}(\gamma,t)$ and $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ using $\hat{v}_1(t-j)$. From (2), we have

$$v(t) = y(t) - \boldsymbol{\phi}^{\mathrm{T}}(\gamma,t)\boldsymbol{\vartheta}. \tag{11}$$

Replacing $y(t)$, $\boldsymbol{\phi}(\gamma,t)$ and $\boldsymbol{\vartheta}$ in (11) with $y(t-j)$, $\hat{\boldsymbol{\phi}}(\gamma,t-j)$ and $\hat{\boldsymbol{\vartheta}}(t)$, the estimate of $v(t-j)$ can be computed by

$$\hat{v}_1(t-j) = y(t-j) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\gamma,t-j)\hat{\boldsymbol{\vartheta}}(t). \tag{12}$$

Then, the estimation vectors $\hat{\boldsymbol{\phi}}(\gamma,t)$ and $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ can be constructed as

$$\hat{\boldsymbol{\phi}}(\gamma,t) = [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\gamma,t), \hat{\boldsymbol{\phi}}_{\mathrm{n}}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{13}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t) = [\hat{v}_1(t-1), \hat{v}_1(t-2), \cdots, \hat{v}_1(t-m)]^{\mathrm{T}}. \tag{14}$$

In order to avoid the large computational cost of matrix inversion, we apply the matrix inversion formula

$$(\boldsymbol{A} + \boldsymbol{BC})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{I} + \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B})\boldsymbol{C}\boldsymbol{A}^{-1}$$

to (6). Then, Equations (5)–(6) can be written as

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \boldsymbol{L}(t)[y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\gamma,t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{15}$$

$$\boldsymbol{L}(t) = \boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\gamma,t)[1 + \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\gamma,t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\gamma,t)]^{-1}, \tag{16}$$

$$\boldsymbol{P}(t) = \boldsymbol{P}(t-1) - \boldsymbol{L}(t)\boldsymbol{\phi}^{\mathrm{T}}(\gamma,t)\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}, \tag{17}$$

where $\boldsymbol{P}(t) \in \mathbb{R}^{(2n+m)\times(2n+m)}$ is the covariance matrix and $\boldsymbol{L}(t) \in \mathbb{R}^{2n+m}$ is the gain vector. Equations (8)–(9) and (12)–(17) construct the recursive least squares (RLS) sub-algorithm for computing $\hat{\boldsymbol{\vartheta}}(t)$. The following derives the sub-algorithm for computing $\hat{\gamma}(t)$.

From the Sub-ID model in (4), we define the criterion function

$$J_2(\gamma) := \frac{1}{2}[y_1(t) - \psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}]^2.$$

Note that $J_2(\gamma)$ is highly nonlinear with respect to the unknown parameter γ, we cannot directly use the linear LS optimization to obtain the estimate $\hat{\gamma}(t)$. Instead, the gradient search is adopted to minimize $J_2(\gamma)$. Computing the gradient of $J_2(\gamma)$ gives

$$\mathrm{grad}[J_2(\gamma)] = \frac{\partial J_2(\gamma)}{\partial \gamma} = y^2(t-1)\psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}[y_1(t) - \psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}]$$

$$= y^2(t-1)\psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}[y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\alpha} - \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} - \psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}]$$

$$= -\boldsymbol{\vartheta}^{\mathrm{T}}\boldsymbol{\phi}'(\gamma,t)[y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\boldsymbol{\alpha} - \boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{d} - \psi(\boldsymbol{\beta},t)\mathrm{e}^{-\gamma y^2(t-1)}]$$

$$= -\boldsymbol{\vartheta}^\mathrm{T} \boldsymbol{\phi}'(\gamma, t)[y(t) - \boldsymbol{\phi}^\mathrm{T}(\gamma, t)\boldsymbol{\vartheta}],$$

where $\boldsymbol{\phi}'(\gamma, t)$ is the derivative of $\boldsymbol{\phi}(\gamma, t)$ with respect to γ and can be computed by

$$\boldsymbol{\phi}'(\gamma, t) = \frac{\partial \boldsymbol{\phi}(\gamma, t)}{\partial \gamma} = [\mathbf{0}_n, -y^2(t-1)\mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}^\mathrm{T}(t), \mathbf{0}_m]^\mathrm{T}. \tag{18}$$

Let $\mu(t)$ represent the step-size. Using the gradient search to minimize $J_2(\gamma)$, we obtain the following recursive relations for computing $\hat{\gamma}(t)$:

$$\begin{aligned}
\hat{\gamma}(t) &= \hat{\gamma}(t-1) - \mu(t)\mathrm{grad}[J_2(\hat{\gamma}(t-1))] \\
&= \hat{\gamma}(t-1) + \mu(t)\boldsymbol{\vartheta}^\mathrm{T}\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)[y(t) - \boldsymbol{\phi}^\mathrm{T}(\hat{\gamma}(t-1), t)\boldsymbol{\vartheta}].
\end{aligned} \tag{19}$$

Referring to the gradient-based identification algorithms for the ExpAR model in [31], we can simply take the step-size $\mu(t)$ as

$$\mu(t) := \frac{1}{r(t)}, \tag{20}$$

$$r(t) = r(t-1) + \|\boldsymbol{\vartheta}^\mathrm{T}\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2. \tag{21}$$

However, a difficulty arises that the information vector $\boldsymbol{\phi}(\hat{\gamma}(t-1), t)$ in (19) contains the unknown noise $v(t-j)$. Similar to the RLS sub-algorithm for computing $\hat{\boldsymbol{\vartheta}}(t)$, we replace the unmeasurable noise $v(t-j)$ with the estimate $\hat{v}_2(t-j)$ and construct the estimation vectors $\hat{\boldsymbol{\phi}}_\mathrm{n}(t)$ and $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)$ using $\hat{v}_2(t-j)$. Replacing $y(t)$ and $\boldsymbol{\phi}(\gamma, t)$ in (11) with $y(t-j)$ and $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t-j)$, the estimate of $v(t-j)$ can be computed by

$$\hat{v}_2(t-j) = y(t-j) - \hat{\boldsymbol{\phi}}^\mathrm{T}(\hat{\gamma}(t-1), t-j)\boldsymbol{\vartheta}. \tag{22}$$

Then, we obtain the estimation vectors

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t) = [\boldsymbol{\phi}_\mathrm{s}^\mathrm{T}(\hat{\gamma}(t-1), t), \hat{\boldsymbol{\phi}}_\mathrm{n}(t)]^\mathrm{T}, \tag{23}$$

$$\hat{\boldsymbol{\phi}}_\mathrm{n}(t) = [\hat{v}_2(t-1), \hat{v}_2(t-2), \cdots, \hat{v}_2(t-m)]^\mathrm{T}. \tag{24}$$

According to (20)–(24), Equation (19) can be written as

$$\hat{\gamma}(t) = \hat{\gamma}(t-1) + \frac{1}{r(t)}\boldsymbol{\vartheta}^\mathrm{T}\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)[y(t) - \hat{\boldsymbol{\phi}}^\mathrm{T}(\hat{\gamma}(t-1), t)\boldsymbol{\vartheta}]. \tag{25}$$

Combining (8)–(9), (18) and (21)–(25), we obtain the extended stochastic gradient (ESG) sub-algorithm for computing $\hat{\gamma}(t)$:

$$\hat{\gamma}(t) = \hat{\gamma}(t-1) + \frac{1}{r(t)}\boldsymbol{\vartheta}^\mathrm{T}\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)[y(t) - \hat{\boldsymbol{\phi}}^\mathrm{T}(\hat{\gamma}(t-1), t)\boldsymbol{\vartheta}], \tag{26}$$

$$r(t) = r(t-1) + \|\boldsymbol{\vartheta}^\mathrm{T}\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2, \quad r(0) = 1, \tag{27}$$

$$\boldsymbol{\phi}'(\hat{\gamma}(t-1), t) = [\mathbf{0}_n, -y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^\mathrm{T}(t), \mathbf{0}_m]^\mathrm{T}, \tag{28}$$

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t) = [\boldsymbol{\phi}_\mathrm{s}^\mathrm{T}(\hat{\gamma}(t-1), t), \hat{\boldsymbol{\phi}}_\mathrm{n}(t)]^\mathrm{T}, \tag{29}$$

$$\boldsymbol{\phi}_\mathrm{s}(\hat{\gamma}(t-1), t) = [\boldsymbol{\varphi}^\mathrm{T}(t), \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^\mathrm{T}(t)]^\mathrm{T}, \tag{30}$$

$$\boldsymbol{\varphi}(t) = [y(t-1), y(t-2), \cdots, y(t-n)]^\mathrm{T}, \tag{31}$$

$$\hat{\boldsymbol{\phi}}_\mathrm{n}(t) = [\hat{v}_2(t-1), \hat{v}_2(t-2), \cdots, \hat{v}_2(t-m)]^\mathrm{T}, \tag{32}$$

$$\hat{v}_2(t) = y(t) - \hat{\boldsymbol{\phi}}^\mathrm{T}(\hat{\gamma}(t-1), t)\boldsymbol{\vartheta}. \tag{33}$$

Although the RLS and ESG sub-algorithm is derived for computing $\hat{\boldsymbol{\vartheta}}(t)$ and $\hat{\gamma}(t)$, respectively, they are technically impossible to be complemented. Since the right-hand sides of the RLS sub-algorithm contain the unknown parameter γ and that of the ESG sub-algorithm contain the unknown parameter vector $\boldsymbol{\vartheta}$, which are the associate terms.

3.3. The coordination for the subsystem parameter estimation

In order to address the above problem, we replace unknown parameters with their corresponding estimates by the hierarchical identification principle [46]. Specifically speaking, the parameter γ in the RLS sub-algorithm is replaced with $\hat{\gamma}(t-1)$ and the parameter vector $\boldsymbol{\vartheta}$ in the ESG sub-algorithm is replaced with $\hat{\boldsymbol{\vartheta}}(t-1)$. Then, the estimates of $\boldsymbol{\vartheta}$ and γ at time t can be computed by an iterative way:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \boldsymbol{L}(t)e(t), \tag{34}$$

$$e(t) = y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{35}$$

$$\boldsymbol{L}(t) = \boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)[1 + \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)]^{-1}, \tag{36}$$

$$\boldsymbol{P}(t) = \boldsymbol{P}(t-1) - \boldsymbol{L}(t)\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}, \tag{37}$$

$$\hat{\gamma}(t) = \hat{\gamma}(t-1) + \frac{1}{r(t)}\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)e(t), \tag{38}$$

$$r(t) = r(t-1) + \|\hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)\|^2, \quad r(0) = 1, \tag{39}$$

$$\boldsymbol{\phi}'(\hat{\gamma}(t-1), t) = [\boldsymbol{0}_n, -y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t), \boldsymbol{0}_m]^{\mathrm{T}}, \tag{40}$$

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t) = [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\hat{\gamma}(t-1), t), \hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)]^{\mathrm{T}}, \tag{41}$$

$$\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1), t) = [\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{42}$$

$$\boldsymbol{\varphi}(t) = [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}}, \tag{43}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t) = [\hat{v}(t-1), \hat{v}(t-2), \cdots, \hat{v}(t-m)]^{\mathrm{T}}, \tag{44}$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t), \tag{45}$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\alpha}_1(t), \cdots, \hat{\alpha}_n(t), \hat{\beta}_1(t), \cdots, \hat{\beta}_n(t), \hat{d}_1(t), \cdots, \hat{d}_m(t)]^{\mathrm{T}}. \tag{46}$$

Equations (34)–(46) form the two-stage least squares extended stochastic gradient (2S-LS-ESG) algorithm for the ExpARMA model.

Algorithm 1 The pseudo-code of implementing the 2S-LS-ESG algorithm

---

Require: $\{y(t), t = 1, 2, \cdots, L_e\}$, m, n, ε

Ensure: $\hat{\boldsymbol{\vartheta}}(t)$ and $\hat{\gamma}(t)$

  1: Initialization: $t = 1$, $\hat{\boldsymbol{\vartheta}}(0) = \boldsymbol{1}_{2n+m}/p_0$, $\hat{\gamma}(0) = 1/p_0$, $\boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}$, $\hat{v}(t-j) = 1/p_0$ ($j = 1, 2, \cdots, m$), $p_0 = 10^6$.

  2: **for** $t = 1 : L_e$ **do**

  3:      Form the information vectors $\boldsymbol{\varphi}(t)$, $\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1), t)$, $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ and $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)$ by (41)–(44).

  4:      Compute the gain vector $\boldsymbol{L}(t)$, covariance matrix $\boldsymbol{P}(t)$ and innovation $e(t)$ by (35)–(37).

  5:      Update the parameter estimation vector $\hat{\boldsymbol{\vartheta}}(t)$ by (34).

  6:      Form the derivative vector $\boldsymbol{\phi}'(\hat{\gamma}(t-1), t)$ by (40).

  7:      Compute the reciprocal of the step-size $r(t)$ by (39).

  8:      Update the parameter estimate $\hat{\gamma}(t)$ by (38).

  9:      Compute the noise estimate $\hat{v}(t)$ by (45).

10:      **If** $\|\hat{\boldsymbol{\vartheta}}(t) - \hat{\boldsymbol{\vartheta}}(t-1)\| + \|\hat{\gamma}(t) - \hat{\gamma}(t-1)\| > \varepsilon$

11:      $t = t + 1$;

12:      **else**

13:      Read out $\hat{\alpha}_i(t)$, $\hat{\beta}_i(t)$ and $\hat{d}_j(t)$ from $\hat{\boldsymbol{\vartheta}}(t)$ in (46), break;

14:      **end**

15: **end for**

The pseudo-code of implementing the 2S-LS-ESG algorithm is shown in Algorithm 1. The proposed two-stage recursive identification algorithms for a class of nonlinear time series models with colored noise in this paper can combine some mathematical tools [47, 48, 49] to study the parameter estimation problems of different systems with colored noises and can be applied other

literatures [50, 51] such as signal modeling and communication networked systems [52, 53] and engineering application systems [54, 55] and so on. Moreover, the quantity $e(t) := y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1) \in \mathbb{R}$ in (35) is a scalar and called the innovation. Applying the multi-innovation identification theory, we derive a two-stage least squares multi-innovation extended stochastic gradient (2S-LS-MIESG) algorithm for the ExpARMA model in the next section.

4. TWO-STAGE LEAST SQUARES MULTI-INNOVATION EXTENDED STOCHASTIC GRADIENT ALGORITHM

In order to improve the parameter estimation accuracy of the 2S-LS-ESG algorithm, we expand the scalar innovation into an innovation vector and derive a 2S-LS-MIESG algorithm by using the multi-innovation identification theory. The details are as follows.

Let p represent the innovation length and expand $e(t)$ in (35) into the innovation vector

$$
\boldsymbol{E}(p,t) = \begin{bmatrix} y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1) \\ y(t-1) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t-1)\hat{\boldsymbol{\vartheta}}(t-1) \\ \vdots \\ y(t-p+1) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t-p+1)\hat{\boldsymbol{\vartheta}}(t-1) \end{bmatrix} \in \mathbb{R}^p. \tag{47}
$$

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

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

Then, the innovation vector $\boldsymbol{E}(p,t)$ in (47) can be written as

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

For simplicity, we define the generalized information item

$$
\psi_0(t) := \hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1), t) \in \mathbb{R}.
$$

In order to make the matrix multiplication dimension compatible, we expand the generalized information item $\psi_0(t)$ into the generalized stacked information vector

$$
\boldsymbol{\Psi}(p,t) := [\psi_0(t), \psi_0(t-1), \cdots, \psi_0(t-p+1)]^{\mathrm{T}} \in \mathbb{R}^p.
$$

When the innovation length $p = 1$, we have

$$
\boldsymbol{E}(p,t) = e(t), \quad \boldsymbol{Y}(p,t) = y(t), \quad \hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1), t) = \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t), \quad \boldsymbol{\Psi}(p,t) = \psi_0(t).
$$

Thus, the 2S-LS-MIESG algorithm for the ExpARMA model can be derived as follows,

$$
\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \boldsymbol{L}(t)[y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\hat{\boldsymbol{\vartheta}}(t-1)], \tag{48}
$$

$$
\boldsymbol{L}(t) = \boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)[1 + \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1), t)]^{-1}, \tag{49}
$$

$$
\boldsymbol{P}(t) = \boldsymbol{P}(t-1) - \boldsymbol{L}(t)\hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1), t)\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}, \tag{50}
$$

$$
\hat{\gamma}(t) = \hat{\gamma}(t-1) + \frac{1}{r(t)}\boldsymbol{\Psi}^{\mathrm{T}}(p,t)\boldsymbol{E}(p,t), \tag{51}
$$

$$\boldsymbol{E}(p,t) = \boldsymbol{Y}(p,t) - \hat{\boldsymbol{\Phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t-1), \tag{52}$$

$$\boldsymbol{Y}(p,t) = [y(t), y(t-1), \cdots, y(t-p+1)]^{\mathrm{T}}, \tag{53}$$

$$\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t) = [\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t), \hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t-1), \cdots, \hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t-p+1)], \tag{54}$$

$$\boldsymbol{\Psi}(p,t) = [\psi_0(t), \psi_0(t-1), \cdots, \psi_0(t-p+1)]^{\mathrm{T}}, \tag{55}$$

$$r(t) = r(t-1) + \|\psi_0(t)\|^2, \quad r(0) = 1, \tag{56}$$

$$\psi_0(t) = \hat{\boldsymbol{\vartheta}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'(\hat{\gamma}(t-1),t), \tag{57}$$

$$\boldsymbol{\phi}'(\hat{\gamma}(t-1),t) = [\boldsymbol{0}_n, -y^2(t-1)\mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t), \boldsymbol{0}_m]^{\mathrm{T}}, \tag{58}$$

$$\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t) = [\boldsymbol{\phi}_{\mathrm{s}}^{\mathrm{T}}(\hat{\gamma}(t-1),t), \hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)]^{\mathrm{T}}, \tag{59}$$

$$\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t) = [\boldsymbol{\varphi}^{\mathrm{T}}(t), \mathrm{e}^{-\hat{\gamma}(t-1)y^2(t-1)}\boldsymbol{\varphi}^{\mathrm{T}}(t)]^{\mathrm{T}}, \tag{60}$$

$$\boldsymbol{\varphi}(t) = [y(t-1), y(t-2), \cdots, y(t-n)]^{\mathrm{T}}, \tag{61}$$

$$\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t) = [\hat{v}(t-1), \hat{v}(t-2), \cdots, \hat{v}(t-m)]^{\mathrm{T}}, \tag{62}$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\phi}}^{\mathrm{T}}(\hat{\gamma}(t-1),t)\hat{\boldsymbol{\vartheta}}(t), \tag{63}$$

$$\hat{\boldsymbol{\vartheta}}(t) = [\hat{\alpha}_1(t), \cdots, \hat{\alpha}_n(t), \hat{\beta}_1(t), \cdots, \hat{\beta}_n(t), \hat{d}_1(t), \cdots, \hat{d}_m(t)]^{\mathrm{T}}. \tag{64}$$

When $p = 1$, the 2S-LS-MIESG algorithm reduces to the 2S-LS-ESG algorithm. Similar to the 2S-LS-ESG algorithm in (34)–(46), the 2S-LS-MIESG algorithm computes the parameter estimates $\hat{\boldsymbol{\vartheta}}(t)$ and $\hat{\gamma}(t)$ by an interactive way. Moreover, the recursion computation of 2S-LS-MIESG algorithm involves more innovations and data. The proposed two-stage least squares multi-innovation extended stochastic gradient algorithm in this paper can joint some mathematical strategies [56, 57] to explore new identification methods of linear and nonlinear stochastic systems [58, 59, 60, 61] and can applied to other fields [62, 63, 64, 65].

The pseudo-code of implementing the 2S-LS-MIESG algorithm is shown in Algorithm 2.

Algorithm 2 The pseudo-code of implementing the 2S-LS-MIESG algorithm

---

Require: $\{y(t), t = 1, 2, \cdots, L_e\}, m, n, \varepsilon, p$

Ensure: $\hat{\boldsymbol{\vartheta}}(t)$ and $\hat{\gamma}(t)$

1: Initialization: $t = 1$, $\hat{\boldsymbol{\vartheta}}(0) = \boldsymbol{1}_{2n+m}/p_0$, $\hat{\gamma}(0) = 1/p_0$, $\boldsymbol{P}(0) = p_0\boldsymbol{I}_{2n+m}$, $\hat{v}(t-j) = 1/p_0$ ($j = 1, 2, \cdots, m$), $p_0 = 10^6$.

2: **for** $t = 1 : L_e$ **do**

3:     Form the information vectors $\boldsymbol{\varphi}(t)$, $\boldsymbol{\phi}_{\mathrm{s}}(\hat{\gamma}(t-1),t)$, $\hat{\boldsymbol{\phi}}_{\mathrm{n}}(t)$ and $\hat{\boldsymbol{\phi}}(\hat{\gamma}(t-1),t)$ by (59)–(62).

4:     Compute the gain vector $\boldsymbol{L}(t)$ and the covariance matrix $\boldsymbol{P}(t)$ by (49)–(50).

5:     Update the parameter estimation vector $\hat{\boldsymbol{\vartheta}}(t)$ by (48).

6:     Compute the derivative $\boldsymbol{\phi}'(\hat{\gamma}(t-1),t)$, information item $\psi_0(t)$ and $r(t)$ by (56)–(58).

7:     Form the stacked vectors and matrix $\boldsymbol{Y}(p,t)$, $\boldsymbol{\Psi}(p,t)$ and $\hat{\boldsymbol{\Phi}}(\hat{\gamma}(t-1),t)$ by (53)–(55).

8:     Compute the innovation vector $\boldsymbol{E}(p,t)$ by (52).

9:     Update the parameter estimate $\hat{\gamma}(t)$ by (51).

10:    Compute the noise estimate $\hat{v}(t)$ by (63).

11:    **If** $\|\hat{\boldsymbol{\vartheta}}(t) - \hat{\boldsymbol{\vartheta}}(t-1)\| + \|\hat{\gamma}(t) - \hat{\gamma}(t-1)\| > \varepsilon$

12:    $t = t + 1$;

13:    **else**

14:    Read out $\hat{\alpha}_i(t)$, $\hat{\beta}_i(t)$ and $\hat{d}_j(t)$ from $\hat{\boldsymbol{\vartheta}}(t)$ in (64), break;

15:    **end**

16: **end for**

On the basis of the multi-innovation identification, the scalar innovation $e(t)$ in (35) is expanded into the innovation vector $\boldsymbol{E}(p,t)$ in (52). This way, the 2S-LS-MIESG algorithm involves not only the current measurement data and innovation, but also the preceding $(p-1)$ measurement data and innovations at each recursion. Moreover, $(p-1)$ data and innovations are repeatedly used in two

adjacent recursions of the 2S-LS-MIESG algorithm. These are the reasons why the 2S-LS-MIESG algorithm has improved parameter estimation accuracy.

Remark 1. For the convergence analysis of the proposed 2S-LS-MIESG algorithm, we assume that the noise sequence $\{v(t)\}$ satisfies

(A1) $\mathrm{E}[v(t)] = 0, \quad \mathrm{E}[v(t)v(t+i)] = 0, \quad i \neq 0,$

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

That is, $\{v(t)\}$ is a white noise sequence with zero mean and variance σ^2.

Theorem. For the identification model in (2) and the 2S-LS-MIESG algorithm in (48)–(64), if assumptions (A1) and (A2) hold, and there exist two constants a, b and an integer t_0 such that when $t \geqslant t_0$, the following persistent excitation condition holds,

$$(A3) \quad a\boldsymbol{I}_{2n+m} \leq \frac{1}{t}\sum_{j=1}^{t} \phi(\hat{\gamma}(j-1), j)\phi^{\mathrm{T}}(\hat{\gamma}(j-1), j) \leq b\boldsymbol{I}_{2n+m}, \text{ a.s.,}$$

and there exist an integer N and a positive constant c independent of t such that the following persistent excitation condition holds,

$$(A4) \quad \sum_{j=0}^{N-1} \frac{\boldsymbol{\Psi}^{\mathrm{T}}(p, t+j)\boldsymbol{\Psi}(p, t+j)}{r(t+j)} \geqslant c, \text{ a.s.}$$

Then the parameter estimation errors $\|\hat{\boldsymbol{\vartheta}}(t) - \boldsymbol{\vartheta}\|$ and $\|\hat{\gamma}(t) - \gamma\|$ given by the 2S-LS-MIESG algorithm converge to zero.

The proofs of Theorem given in this paper can be done in a similar way to those in [66].

Remark 2. In respect of fitting 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 [7, 11, 18]. Other literatures converted the ExpAR model into a nonlinear state-space model through the model transformation technique and derived the extended Kalman filter methods [67]. Compared with the existing identification algorithms in these publications, the multi-innovation identification algorithm proposed in this paper relaxes the conditions on the parameter γ, does not involve model transformation, and can identify all the unknown parameters by an interactive way.

5. EXAMPLE

Consider the following ExpARMA process

$$\begin{aligned} y(t) &= [\alpha_1 + \beta_1 \mathrm{e}^{-\gamma y^2(t-1)}]y(t-1) + \cdots + [\alpha_n + \beta_n \mathrm{e}^{-\gamma y^2(t-1)}]y(t-n) + D(z)v(t) \\ &= [1.06 - 1.20\mathrm{e}^{-0.13y^2(t-1)}]y(t-1) + [-0.59 + 2.03\mathrm{e}^{-0.13y^2(t-1)}]y(t-2) \\ &\quad + 0.03v(t-1) + 0.04v(t-2) + v(t). \end{aligned}$$

The parameters to be estimated are

$$\boldsymbol{\vartheta} = [\alpha_1, \alpha_2, \beta_1, \beta_2, d_1, d_2]^{\mathrm{T}} = [1.06, -0.59, -1.20, 2.03, 0.03, 0.04]^{\mathrm{T}}, \quad \gamma = 0.13.$$

For convenience, we define the parameter vector $\boldsymbol{\Theta} := [\boldsymbol{\vartheta}^{\mathrm{T}}, \gamma]^{\mathrm{T}} \in \mathbb{R}^{2n+m+1}$. 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 $L_e = 5000$. The measurement data $y(t)$ is shown in Figure 1.

To exhibit the advantage of the proposed multi-innovation identification algorithm, taking the variance $\sigma^2 = 0.20^2$, we adopt the 2S-LS-MIESG algorithm with the innovation length $p = 2$ and $p = 3$ to estimate the parameters of this ExpARMA model, respectively. The parameter estimation errors $\delta := \|\hat{\boldsymbol{\Theta}}(t) - \boldsymbol{\Theta}\|/\|\boldsymbol{\Theta}\| \times 100\%$ versus t are shown in Figure 2, the parameter estimates of the 2S-LS-MIESG with $p = 3$ versus t are shown in Figures 3–4.
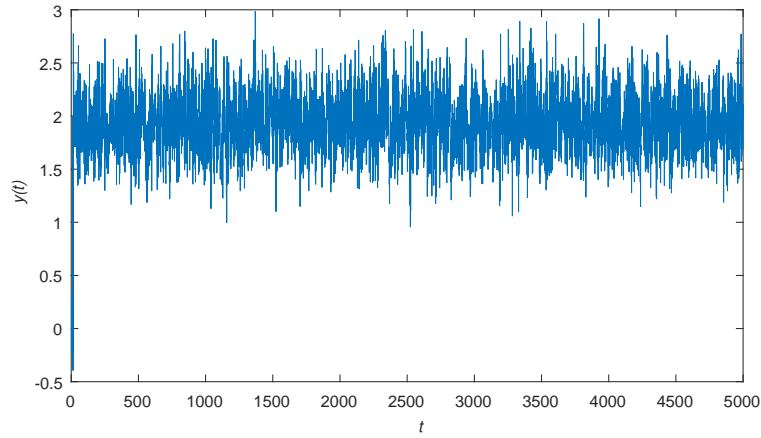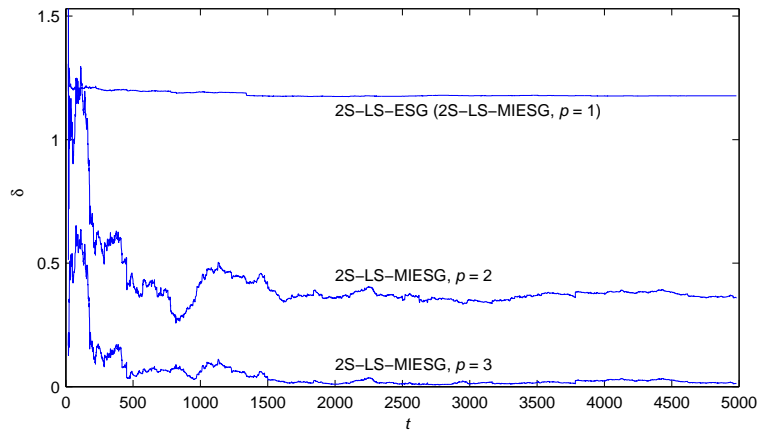
Figure 1. The measurement data

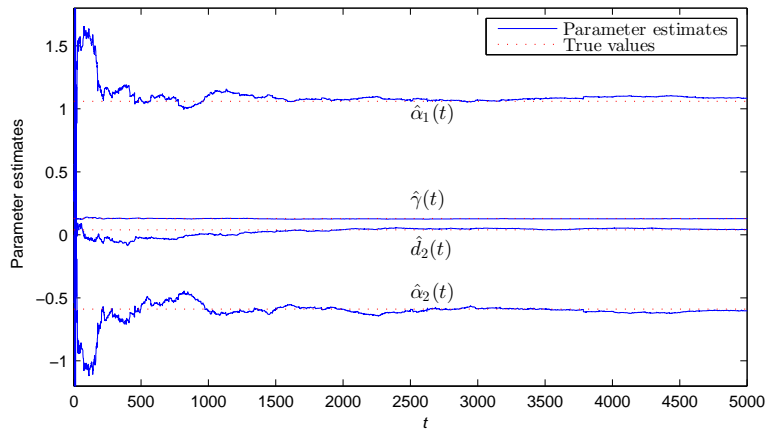Figure 2. The LS-ESG and 2S-LS-MIESG estimation errors δ versus t ($\sigma^2 = 0.20^2$)

Figure 3. The 2S-LS-MIESG parameter estimates $\hat{\alpha}_1(t)$, $\hat{\alpha}_2(t)$, $\hat{d}_2(t)$ and $\hat{\gamma}(t)$ versus t ($\sigma^2 = 0.20^2$, $p = 3$)
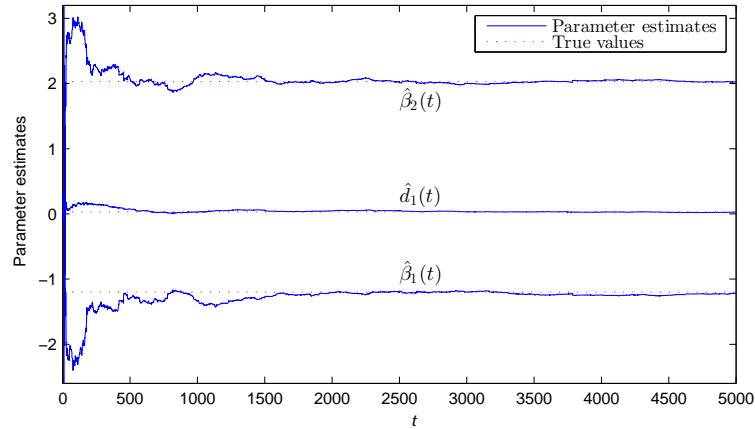
Figure 4. The 2S-LS-MIESG parameter estimates $\hat{\beta}_1(t)$, $\hat{\beta}_2(t)$ and $\hat{d}_1(t)$ versus t ($\sigma^2 = 0.20^2$, $p = 3$)

To show the influence of the noise level on the proposed 2S-LS-MIESG algorithm, we fix the innovation length $p = 3$ and adopt the 2S-LS-MIESG algorithm to estimate the parameters of this ExpARMA model with the noise variance $\sigma^2 = 0.20^2$, $\sigma^2 = 0.30^2$ and $\sigma^2 = 0.40^2$, respectively. The parameter estimation errors versus t are shown in Tables I–III and Figure 5.

Table I. The 2S-LS-MIESG parameter estimates and errors ($p = 3$, $\sigma^2 = 0.20^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	1.60530	-1.07061	-2.25956	2.93696	0.12183	-0.02871	0.14076	59.38394
200	1.13814	-0.69759	-1.42494	2.28159	0.15808	-0.03735	0.13235	14.77609
500	1.06689	-0.56187	-1.29102	2.05322	0.07383	-0.02357	0.12937	4.70709
1000	1.10952	-0.60223	-1.33956	2.09667	0.02311	0.00648	0.12942	6.26669
2000	1.07329	-0.59043	-1.21047	2.01511	0.04445	0.04073	0.12682	1.01587
3000	1.06941	-0.58618	-1.19981	2.00300	0.03263	0.05160	0.12730	1.17930
4000	1.09596	-0.61016	-1.24723	2.04562	0.02789	0.05149	0.12696	2.47511
5000	1.08396	-0.60258	-1.21841	2.02425	0.02649	0.04249	0.12822	1.26365
True values	1.06000	-0.59000	-1.20000	2.03000	0.03000	0.04000	0.13000	

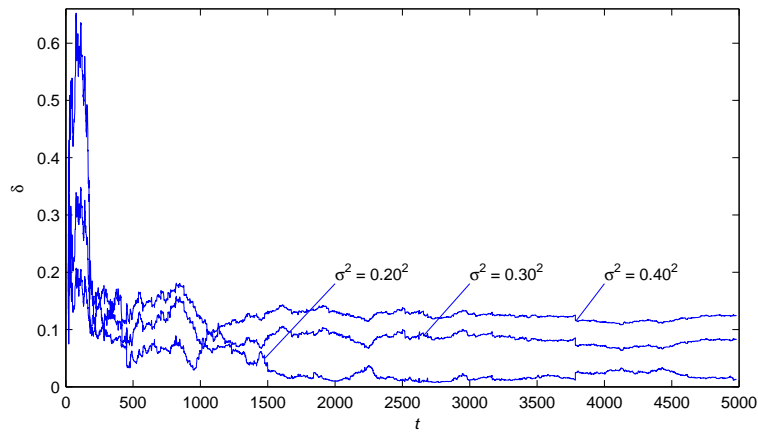Table II. The 2S-LS-MIESG parameter estimates and errors ($p = 3$, $\sigma^2 = 0.30^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	1.24663	-0.75993	-1.78566	2.56814	0.17118	-0.01361	0.16332	31.95780
200	0.99980	-0.53834	-1.29207	2.15873	0.19016	-0.03156	0.15153	9.42004
500	0.96184	-0.40049	-1.21491	1.92145	0.07242	-0.03106	0.14658	9.57691
1000	0.99035	-0.42692	-1.23242	1.93447	0.02224	0.00186	0.14641	7.84580
2000	0.98281	-0.43616	-1.12248	1.86780	0.03359	0.04118	0.14346	9.38690
3000	0.98311	-0.44436	-1.12191	1.87978	0.03380	0.04962	0.14330	8.91695
4000	1.00237	-0.46350	-1.15747	1.91412	0.02949	0.05190	0.14296	7.03171
5000	0.99256	-0.45500	-1.12852	1.88823	0.02645	0.04356	0.14524	8.27074
True values	1.06000	-0.59000	-1.20000	2.03000	0.03000	0.04000	0.13000	

From Tables I–III and Figures 2–5, we can draw the following conclusions.

- The parameter estimation accuracy of the proposed 2S-LS-MIESG algorithm becomes higher with the innovation length p increasing – see Figure 2.

Table III. The 2S-LS-MIESG parameter estimates and errors ($p = 3$, $\sigma^2 = 0.40^2$)

t	α_1	α_2	β_1	β_2	d_1	d_2	γ	δ (%)
100	1.04840	-0.53654	-1.56017	2.32251	0.19381	-0.01743	0.18108	18.86669
200	0.89303	-0.38614	-1.19312	2.01674	0.19630	-0.03236	0.16842	12.14561
500	0.90058	-0.32039	-1.16662	1.86941	0.07940	-0.02951	0.16152	13.74720
1000	0.93157	-0.34354	-1.18071	1.86772	0.02276	0.00373	0.16088	12.27536
2000	0.94466	-0.37072	-1.08303	1.81686	0.02743	0.04542	0.15750	13.11461
3000	0.94154	-0.37444	-1.08823	1.83283	0.03721	0.05165	0.15611	12.64386
4000	0.95673	-0.38820	-1.11799	1.85824	0.03134	0.05385	0.15594	11.20137
5000	0.95070	-0.38034	-1.09176	1.82938	0.02460	0.04638	0.15939	12.42111
True values	1.06000	-0.59000	-1.20000	2.03000	0.03000	0.04000	0.13000	

Figure 5. The 2S-LS-MIESG estimation errors δ versus t ($p = 3$)

- With an appropriate innovation length, the 2S-LS-MIESG parameter estimates approach to their corresponding true values and the 2S-LS-MIESG algorithm is effective to identify the ExpARMA process – see Table I and Figures 3–4.
- The parameter estimation errors of the 2S-LS-MIESG algorithm become smaller with the decreasing of noise levels – see Tables I–III and Figure 5.

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 2S-LS-ESG algorithm and the 2S-LS-MIESG algorithm with $\sigma^2 = 0.20^2$ and $p = 3$. The 2S-LS-ESG predicted data $\hat{y}_1(t)$, the 2S-LS-MIESG predicted data $\hat{y}_2(t)$ and the measurement data $y(t)$ are plotted in Figures 6–7. To evaluate the prediction performance, we define and compute the mean square errors (MSEs):

$$MSE_1 := \left[\frac{1}{L_r}\sum_{t=L_e+1}^{L_e+L_r}[\hat{y}_1(t) - y(t)]^2\right]^{1/2} = 0.22896,$$

$$MSE_2 := \left[\frac{1}{L_r}\sum_{t=L_e+1}^{L_e+L_r}[\hat{y}_2(t) - y(t)]^2\right]^{1/2} = 0.20919.$$

From Figures 6–7, we can see that the 2S-LS-MIESG predicted data $\hat{y}_2(t)$ is close to the measurement data $y(t)$, which means the estimated model can capture the dynamics of this ExpARMA model.
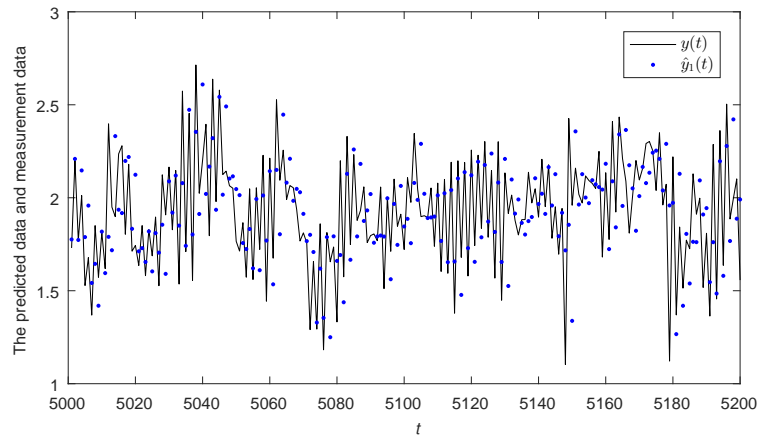
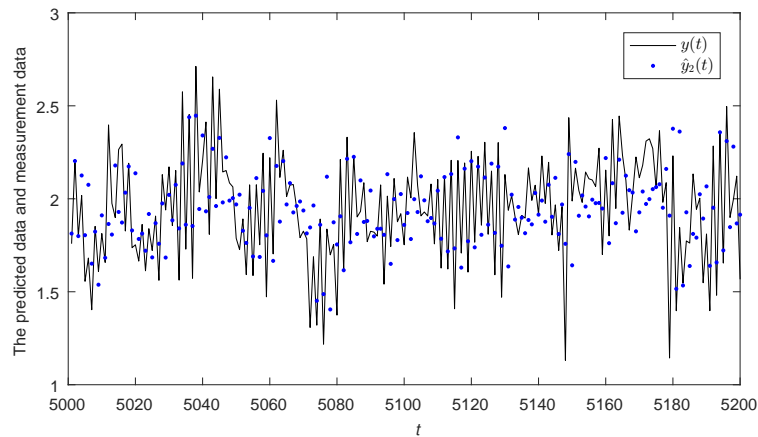Figure 6. The 2S-LS-ESG predicted data and measurement data

Figure 7. The 2S-LS-MIESG predicted data and measurement data

6. CONCLUSIONS

This paper studies the identification problems for the ExpARMA model. Using the decomposition technique and the multi-innovation identification theory, we derive a two-stage least squares extended stochastic gradient (2S-LS-ESG) algorithm and a two-stage least squares multi-innovation extended stochastic gradient (2S-LS-MIESG) algorithm. The simulation results indicate that the proposed 2S-LS-MIESG algorithm can effectively deal with the parameter estimation for the ExpARMA model and has improved parameter estimation accuracy. The methods in this paper can be applied to system modeling [68, 69, 70] and other fields such as information processing [71, 72, 73, 74, 75] and communication systems and so on.

ORCID

Feng Ding, https://orcid.org/0000-0002-2721-2025

REFERENCES

1. Ding F. *System Identification - New Theory and Methods*. Science Press: Beijing, 2013.
2. Ding F. *System Identification - Auxiliary Model Identification Idea and Methods*. Science Press: Beijing, 2017.
3. Ding F. *System Identification - Iterative Search Principle and Identification Methods*. Science Press: Beijing, 2018.
4. Ding F. *System Identification - Multi-Innovation Identification Theory and Methods*. Science Press: Beijing, 2016.
5. Chen GY, Gan M, Chen CLP, et al. A regularized variable projection algorithm for separable nonlinear least squares problems. *IEEE Transactions on Automatic Control*. 2019;64(2):526-537.
6. Ozaki T. Non-linear time series models for non-linear random vibrations. *Journal of Applied Probability*. 1980;17(1):84-93.
7. Haggan V, Ozaki T. Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika*. 1981;68(1):189-196.
8. Ozaki T, Oda H. Nonlinear time series model identification by Akaike's information criterion. *IFAC Proceedings Volumes*. 1977;10(12):83-91.
9. Ozaki T. The statistical analysis of perturbed limit cycle processes using nonlinear time series models. *Journal of Time Series Analysis*. 1982;3(1):29-41.
10. Teräsvirta T. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*. 1994;89(425):208-218.
11. Chen GY, Gan M, Chen GL. Generalized exponential autoregressive models for nonlinear time series: stationarity, estimation and applications. *Information Sciences*. 2018;438:46-57.
12. Pan J, Ma H, Zhang X, et al. Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises. *IET Signal Processing*. 2020;14. https://doi.org/10.1049/iet-spr.2019.0481.
13. Yu CP, Verhaegen M. Subspace identification of individual systems operating in a network. *IEEE Transactions on Automatic Control*. 2018;63(4):1120-1125.
14. Zhou ZP, Liu XF. State and fault estimation of sandwich systems with hysteresis. *International Journal of Robust and Nonlinear Control*. 2018;28(13):3974-3986.
15. Gan M, Chen CLP, Chen GY, et al. On some separated algorithms for separable nonlinear squares problems. *IEEE Transactions on Cybernetics*. 2018;48(10):2866-2874.
16. Zhang X, Ding F, Xu L, Yang EF. Highly computationally efficient state filter based on the delta operator. *International Journal of Adaptive Control and Signal Processing*. 2019;33(6):875-889.
17. Yu CP, Ljung L, Wills A, Verhaegen M. Constrained subspace method for the identification of structured state-space models. *IEEE Transactions on Automatic Control*. 2020;65. https://doi.org/10.1109/TAC.2019.2957703.
18. Shi Z, Tamura Y, Ozaki T. Monitoring the stability of BWR oscillation by nonlinear time series modeling. *Annals of Nuclear Energy*. 2001;28:953-966.
19. Xu L. The parameter estimation algorithms based on the dynamical response measurement data. *Advances in Mechanical Engineering*. 2017;9(11):1-12. https://doi.org/10.1177/1687814017730003.
20. Xu L, Xiong WL, Alsaedi A, Hayat T. Hierarchical parameter estimation for the frequency response based on the dynamical window data. *International Journal of Control Automation and Systems*. 2018;16(4):1756-1764.
21. Xu L, Song GL. A recursive parameter estimation algorithm for modeling signals with multi-frequencies. *Circuits Systems and Signal Processing*. 2020;39(8):4198-4224.
22. Zhang X, Ding F, Xu L. Recursive parameter estimation methods and convergence analysis for a special class of nonlinear systems. *International Journal of Robust and Nonlinear Control*. 2020;30(4):1373-1393.
23. Wang LJ, Ji Y, Wan LJ, Bu N. Hierarchical recursive generalized extended least squares estimation algorithms for a class of nonlinear stochastic systems with colored noise. *Journal of the Franklin Institute*. 2019;356(16):10102-10122.
24. Ding J, Cao ZX, Chen JZ, Jiang GP. Weighted parameter estimation for Hammerstein nonlinear ARX systems. *Circuits Systems and Signal Processing*. 2020;39(4):2178-2192.
25. Ding J, Chen JZ, Lin JX, Jiang GP. Particle filtering-based recursive identification for controlled auto-regressive systems with quantised output. *IET Control Theory and Applications*. 2019;13(14):2181-2187.
26. Zhang GZ, Zhang XX, Cheng HT, Tang J. Ladder-Wise calculation method for z-coordinate of transformer PD source based on planar layout UHF antenna sensors. *IEEJ Transactions on Electrical and Electronic Engineering*. 2020;15(3):340-345.
27. Cao Y, Wang Z. Liu F, et al. Bio-inspired speed curve optimization and sliding mode tracking control for subway trains. *IEEE Transactions on Vehicular Technology*. 2019;68(7):6331-6342.
28. Cao Y, Sun YK, Xie G, et al. 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):7544-7551.
29. Duan HH, Jia J, Ding RF. Two-stage recursive least squares parameter estimation algorithm for output error models. *Mathematical and Computer Modelling*. 2012;55(3-4):1151-1159.
30. Xu H, Wan LJ, Ding F, Alsaedi A, Hayat T. Fitting the exponential autoregressive model through recursive search. *Journal of the Franklin Institute*. 2019;356(11):5801-5818.
31. Xu H, Ding F, Yang EF. Modeling a nonlinear process using the exponential autoregressive time series model. *Nonlinear Dynamics*. 2019;95(3):2079-2092.
32. Yang F, Zhang Y, Liu X, Li XX. The quasi-boundary value method for identifying the initial value of the space-time an fractional diffusion equation. *Acta Mathematica Scientia*. 2020;40B(3):641-658.

33. Yang F, Zhang P, Li XX, Ma XY. Tikhonov regularization method for identifying the space-dependent source for time-fractional diffusion equation on a columnar symmetric domain. *Advances in Difference Equations*. 2020;2020(1):128.
34. Wang L, Liu H, Dai LV, Liu YW. Novel method for identifying fault location of mixed lines. *Energies*. 2018;11(6), Article Number: 1529.
35. Liu H, Zou QX, Zhang ZP. Energy disaggregation of appliances consumptions using ham approach. *IEEE Access*. 2019;7:185977-185990.
36. Zhao XL, Lin ZY, 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), Article Number: 3325.
37. Ding J, Chen JZ, Lin JX, Wan LJ. Particle filtering based parameter estimation for systems with output-error type model structures. *Journal of the Franklin Institute*. 2019;356(10):5521-5540.
38. Gu Y, Liu J, Li X, Chou Y, Ji Y. State space model identification of multirate processes with time-delay using the expectation maximization. *Journal of the Franklin Institute*. 2019;356(3):1623-1639.
39. Gu Y, Chou Y, Liu J, Ji Y. Moving horizon estimation for multirate systems with time-varying time-delays. *Journal of the Franklin Institute*. 2019;356(4):2325-2345.
40. Ji Y, Zhang C, Kang Z, Yu T. Parameter estimation for block-oriented nonlinear systems using the key term separation. *International Journal of Robust and Nonlinear Control*. 2020;30(9):3727-3752. https://doi.org/10.1002/rnc.4961
41. Ji Y, Jiang XK, Wan LJ. Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems. *Journal of the Franklin Institute*. 2020;357(8):5019-5032.
42. Wang LJ, Ji Y, Yang HL, Xu L. Decomposition-based multiinnovation gradient identification algorithms for a special bilinear system based on its input-output representation. *International Journal of Robust and Nonlinear Control*. 2020;30(9):3607-3623.
43. Wan LJ, Liu XM, Ding F, Chen CP. Decomposition least-squares-based iterative identification algorithms for multivariable equation-error autoregressive moving average systems. *Mathematics*. 2019;7(7), Article Number: 609.
44. Wan LJ, Ding F, Liu XM, Chen CP. A new iterative least squares parameter estimation approach for equation-error autoregressive systems. *International Journal of Control Automation and Systems*. 2020;18(3):780-790.
45. Li MH, Liu XM. 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*. 2020;18(6):1581-1592.
46. Guo CQ, Wang LJ, Deng F. The auxiliary model based hierarchical estimation algorithms for bilinear stochastic systems with colored noises. *International Journal of Control Automation and Systems*. 2020;18(3):650-660.
47. Yang F, Pu Q, Li XX. The fractional Tikhonov regularization methods for identifying the initial value problem for a time-fractional diffusion equation. *Journal of Computational and Applied Mathematics*. 2020;380(15):112998.
48. Yin CC, Wen YZ. An extension of Paulsen-Gjessing's risk model with stochastic return on investments. *Insurance Mathematics & Economics*. 2013;52(3):469-476.
49. Yin CC, Wen YZ, Y.X. Zhao. On the optimal dividend problem for a spectrally positive levy process. *Astin Bulletin*. 2014;44(3):635-651.
50. Chang YF, Zhai GS, Fu B, Xiong LL. Quadratic stabilization of switched uncertain linear systems: a convex combination approach. *CAA Journal of Automatica Sinica*. 2019;6(5):1116-1126.
51. Chang YF, Sun CJ, Qiu, Y. Effective notch stress method for fatigue assessment of sheet alloy material and bi-material welded joints. *Thin-Walled Structures*. 2020;151:106745.
52. Ma FY, Yin YK, Pang SP, Liu JX, Chen W. A data-driven based framework of model optimization and neural network modeling for microbial fuel cells. *IEEE Access*. 2019;7:162036-162049.
53. Wang LJ, Guo J, Xu C, et al. Hybrid model predictive control strategy of supercapacitor energy storage system based on double active bridge. *Energies*. 2019;12(11):2134.
54. Ma FY, Yin YK, Li M. Start-up process modelling of sediment microbial fuel cells based on data driven. *Mathematical Problems in Engineering*. 2019;2019, Article Number: 7403732.
55. Ma FY, Fu CC, Yang J, Yang QZ. Control strategy for adaptive active energy harvesting in sediment microbial fuel cells. *Journal of Energy Engineering*. 2020;146(1) Article Number: 04019034.
56. Yin CC, Yuen KC. Exact joint laws associated with spectrally negative Levy processes and applications to insurance risk theory. *Frontiers of Mathematics in China*. 2014;9(6):1453-1471.
57. Yin CC, Yuen KC. Optimal dividend problems for a jump-diffusion model with capital injections and proportional transaction costs. *Journal of Industrial and Management Optimization*. 2015;11(4):1247-1262.
58. Wu MH, Chen R, Tong Y. Shadow elimination algorithm using color and texture features. *Computational Intelligence and Neuroscience*. 2020;2020:2075781.
59. Geng L, Xiao RB. Control and backbone identification for the resilient recovery of a supply network utilizing outer synchronization. *Applied Sciences*. 2020;10(1): 313
60. Tang L, Liu GJ, Yang M, Li FY, Ye FP, Li CY. Joint design and torque feedback experiment of rehabilitation robot. *Advances in Mechanical Engineering* 2020;12:1687814020924498
61. Wu TZ, Ye FC, Su YH, Wang YB, Riffat S. Coordinated control strategy of DC microgrid with hybrid energy storage system to smooth power output fluctuation. *International Journal of Low-Carbon Technologies*. 2020;15(1):46-54.
62. Wu MH, Yue HH, Wang J, et al. Object detection based on RGC mask R-CNN, *IET Image Processing*. 2020;14(8):1502-1508.
63. Wang LJ, Feng BY, Wang Y, et al. Bidirectional short-circuit current blocker for DC microgrid based on solid-state circuit breaker. *Electronics*. 2020;9(2):306.
64. Zhou YM, Mei SJ, Feng JJ, Sun DW, Mei F, Xu JX, Cao XA. Effects of PEDOT:PSS:GO composite hole transport layer on the luminescence of perovskite light-emitting diodes. *RSC Advances*. 2020;10(44):26381-26387.

65. He ZQ, Chen WB, Xu X, Harn L, Wan ML. A reliable and efficient PUF-based cryptographic key generator using bit self-tests. *Electronics Letters*. 2020;56(16):803-806.
66. F. Ding, System Identification - Performances Analysis for Identification Methods, Science Press, Beijing, 2014.
67. Ghosh H, Gurung B, Gupta P. Fitting EXPAR models through the extended Kalman filter. *The Indian Journal of Stats*. 2015;77:27-44.
68. Xu L, Chen L, Xiong WL. Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. *Nonlinear Dynamics*. 2015;79(3):2155-2163.
69. Xu L. The damping iterative parameter identification method for dynamical systems based on the sine signal measurement. *Signal Processing*. 2016;120:660-667.
70. Zhan XS, Zhang WH, Wu J, Yan H. Performance analysis of NCSs under channel noise and bandwidth constraints. *IEEE Access*. 2020;8:20279-20288.
71. 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). Article Number: 013130. https://doi.org/10.1063/1.5085397
72. Yang GC, Chen ZJ, Li Y, Su ZD. Rapid relocation method for mobile robot based on improved ORB-SLAM2 algorithm.*Remote Sensing*. 2019;11(2):149.
73. Su ZD, Li Y, Yang GC. Dietary composition perception algorithm using social robot audition for Mandarin Chinese. *IEEE Access*. 2020;8:8768-8782.
74. Cao Y, Ma LC, Xiao S, et al. Standard analysis for transfer delay in CTCS-3. *Chinese Journal of Electronics*. 2017;26(5):1057-1063.
75. Hu X, Zhan XS, Wu J, Yan, H. Performance of SIMO networked time-delay systems with encoding-decoding and quantization constraints. *IEEE Access*. 2020;8:55125-55134.