Joint Parameter and Time-Delay Estimation for a Class of Nonlinear Time-Series Models

Huan Xu, Feng Ding, and Benoit Champagne, *Senior Member, IEEE*

***Abstract*—Nonlinear time-series modeling is fundamental to a wide variety of control and prediction problems. This letter focuses on the joint parameter and time-delay estimation for an extended version of the nonlinear exponential autoregressive (ExpAR) time-series model. To address the difficulties posed by the unknown time-delay and improve the estimation accuracy, we first employ the redundant rule to transform the ExpAR model into an augmented identification model. Then we invoke the multi-innovation theory to enhance data utilization and propose a new algorithm that combines stochastic gradient descent with discrete search for estimating the unknown model parameters and time-delay. The simulation results show that by properly adjusting the innovation length, the estimation accuracy of the proposed multi-innovation algorithm can significantly exceed that of the single-innovation algorithm.**

***Index Terms*—Multi-innovation theory, nonlinear time-series model, parameter and time-delay estimation, redundant rule.**

I. Introduction

ALTHOUGH linear time-series models provide powerful and convenient tools for the solution of a wide variety of problems in time-series prediction and control, they are not suitable for revealing the existence of some nonlinear phenomena [1]. In the past decades, the study of nonlinear time-series models has therefore attracted much interest [2]–[4]. An important member of this class is the so-called exponential autoregressive (ExpAR) model, which has the ability to capture the behaviors of nonlinear random vibrations, including amplitude dependent characteristic frequencies, jump phenomena and perturbed limit cycles [5], [6]. As introduced in [7], [8], given a time-series $\{y(t) : t \in \mathbb{Z}\}$, the nth-order ExpAR model can be described as

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

where $v(t)$ is a white noise process. This equation describes the classic ExpAR model which features a global nonlinear and local linear structure.

Parameter estimation is to establish the mathematic models based on observations and has been used in distributed networks [9], econometrics [10] and other areas. Several estimation methods have been developed for the ExpAR model, such as the maximum likelihood [11], [12] and the variable projection methods [13]. Most of these contributions do not consider the intrinsical time-delay which is inevitable in data transmission processes and renders the estimation problem more challenging [14], [15]. Recently, Chen *et al.* have derived a biased compensation recursive least squares threshold algorithm for a time-delay rational model [16]. Besides, Zhang *et al.* have investigated a redundant recursive hierarchical least squares algorithm for bilinear time-delay systems [17]. However, little attention has been given to jointly estimating the unknown time-delay and parameters of the *time-delay* ExpAR model.

The traditional identification methods cannot be directly applied to the time-delay ExpAR model because of the poor performance caused by the unknown time-delay. Multi-innovation theory is an effective tool to improve estimation accuracy and has been used for estimating system parameters without time-delay [18], [19]. The key idea is to expand the scalar innovation into a multi-dimensional innovation vector/matrix and to make full use of the measurement data [20]. Using the multi-innovation theory, Chaudhary *et al.* investigated a multi-innovation fractional least mean square adaptive algorithm algorithm for input nonlinear systems [21]. Wang *et al.* derived two multi-innovation stochastic gradient algorithms for bilinear systems [22]. Besides, Xia *et al.* developed a maximum likelihood multi-innovation extended stochastic gradient algorithm for multivariable systems [23].

In this letter, aiming to extend the methods in [24], [25] for estimating the unknown parameters of the classic ExpAR model, we employ the redundant rule and multi-innovation theory to derive joint parameter and time-delay estimation algorithms for the time-delay ExpAR model. The main contributions of this paper are as follows.

- Using the redundant rule, we transform the identification model into an augmented one and derive a threshold-free procedure for estimating the unknown time-delay, such that the empirical choice of the threshold can be avoided.

Manuscript received October 22, 2021; revised November 22, 2021; accepted December 3, 2021. Date of publication February 16, 2022; date of current version April 18, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61873111 and in part by 111 Project under Grant B12018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Luis Antonio Azpicueta-Ruiz. *(Corresponding author: Feng Ding.)*

Huan Xu is with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China, and also with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: hxu1992@126.com).

Feng Ding is with the Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, China (e-mail: fding@jiangnan.edu.cn).

Benoit Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: benoit.champagne@mcgill.ca).

Digital Object Identifier 10.1109/LSP.2022.3152108

1070-9908 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: McGill University. Downloaded on November 21,2022 at 05:30:11 UTC from IEEE Xplore. Restrictions apply.

- Applying the innovation extension, we propose a redundant rule-based multi-innovation stochastic gradient algorithm which exhibits higher data utilization and estimation accuracy than the single-innovation algorithm.

This letter is organized as follows. Section II describes the time-delay ExpAR model and the augmented identification problem. Two joint parameter and time-delay estimation algorithms are developed in Section III. The comparative performance of the algorithms is evaluated in Section IV. Finally, Section V provides some concluding remarks.

II. Problem Description

Given a time-series $\{y(t) : t \in \mathbb{Z}\}$, the time-delay ExpAR model is described as

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

where $v(t)$ is a white noise with zero mean and variance σ^2, n is the model order and $\tau \in \mathbb{N}$ is an unknown time-delay. It is assumed that the model order n is known and that the data $y(t)$ is measurable for $t \geq 0$; without loss of generality, the initial values are set to $y(t) = 0$ for $t < 0$.

Define the parameter vectors: $\boldsymbol{\theta} := [\boldsymbol{\alpha}^{\mathrm{T}}, \boldsymbol{\beta}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n}$, $\boldsymbol{\alpha} := [\alpha_1, \ldots, \alpha_n]^{\mathrm{T}} \in \mathbb{R}^n$, $\boldsymbol{\beta} := [\beta_1, \ldots, \beta_n]^{\mathrm{T}} \in \mathbb{R}^n$, and the information vectors:

$$\begin{aligned} \boldsymbol{\phi}(t; \gamma, \tau) &:= [\boldsymbol{\varphi}^{\mathrm{T}}(t; \tau), \mathrm{e}^{-\gamma y^2(t-1)} \boldsymbol{\varphi}^{\mathrm{T}}(t; \tau)]^{\mathrm{T}} \in \mathbb{R}^{2n}, \\ \boldsymbol{\varphi}(t; \tau) &:= [y(t-1-\tau), \ldots, y(t-n-\tau)]^{\mathrm{T}} \in \mathbb{R}^n. \end{aligned}$$

Then, the time-delay ExpAR model in (1) can be compactly written as the identification model

$$y(t) = \boldsymbol{\phi}^{\mathrm{T}}(t; \gamma, \tau)\boldsymbol{\theta} + v(t). \tag{2}$$

Note that the information vector $\boldsymbol{\phi}(t; \gamma, \tau)$ contains the variables $y(t-1-\tau)$, $y(t-2-\tau)$, …, $y(t-n-\tau)$ which cannot be determined due to the unknown time-delay τ. Hence, some existing algorithms derived in [24], [25] cannot be used for this time-delay ExpAR model. To address this difficulty, we set a maximum regression length $N \geq |n + \tau|$ and employ the redundant rule to derive an augmented identification model.

Define the augmented parameter vectors:

$$\begin{aligned} \boldsymbol{\theta}_{\mathrm{a}} &:= [\boldsymbol{\alpha}_{\mathrm{a}}^{\mathrm{T}}, \boldsymbol{\beta}_{\mathrm{a}}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2N}, \\ \boldsymbol{\alpha}_{\mathrm{a}} &:= [\zeta_1, \ldots, \zeta_\tau, \alpha_1, \ldots, \alpha_n, \zeta_{\tau+1}, \ldots, \zeta_{N-n}]^{\mathrm{T}} \in \mathbb{R}^N, \\ \boldsymbol{\beta}_{\mathrm{a}} &:= [\eta_1, \ldots, \eta_\tau, \beta_1, \ldots, \beta_n, \eta_{\tau+1}, \ldots, \eta_{N-n}]^{\mathrm{T}} \in \mathbb{R}^N, \end{aligned}$$

and the augmented information vectors:

$$\begin{aligned} \boldsymbol{\phi}_{\mathrm{a}}(t; \gamma) &:= [\boldsymbol{\varphi}_{\mathrm{a}}^{\mathrm{T}}(t), \mathrm{e}^{-\gamma y^2(t-1)} \boldsymbol{\varphi}_{\mathrm{a}}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2N}, \\ \boldsymbol{\varphi}_{\mathrm{a}}(t) &:= [y(t-1), y(t-2), \ldots, y(t-N)]^{\mathrm{T}} \in \mathbb{R}^N. \end{aligned}$$

Then, the identification model in (2) can be seen as a special case of the augmented identification model

$$y(t) = \boldsymbol{\phi}_{\mathrm{a}}^{\mathrm{T}}(t; \gamma)\boldsymbol{\theta}_{\mathrm{a}} + v(t). \tag{3}$$

Estimates of γ and $\boldsymbol{\theta}_{\mathrm{a}}$ can be computed by solving a suitable optimization problem, e.g.,

$$\min_{\gamma, \boldsymbol{\theta}_{\mathrm{a}}} \sum_{i=0}^{t} w_i \| y(i) - \boldsymbol{\phi}_{\mathrm{a}}^{\mathrm{T}}(i; \gamma)\boldsymbol{\theta}_{\mathrm{a}} \|^2, \tag{4}$$

where w_i is a non-negative weight parameter that can be set in different ways. For instance, the choice $w_i = \lambda^{t-i}$, where $0 < \lambda < 1$ is a forgetting factor, corresponds to an exponential weighted least-squares criterion.

Obviously, the dimension of the augmented parameter vector $\boldsymbol{\theta}_{\mathrm{a}}$ is larger than that of $\boldsymbol{\theta}$ because of the redundant parameters ζ_i and η_i $(i = 1, 2, \ldots, N-n)$ whose true values are all supposed to be zeros. These zeros form three distinct blocks in the vector $\boldsymbol{\theta}_{\mathrm{a}}$ which can be expanded as

$$\boldsymbol{\theta}_{\mathrm{a}} = [\underbrace{0, \ldots, 0}_{\tau}, \alpha_1, \ldots, \alpha_n, \underbrace{0, \ldots, 0}_{N-n}, \beta_1, \ldots, \beta_n, \underbrace{0, \ldots, 0}_{N-n-\tau}]^{\mathrm{T}}.$$

Compared with $\boldsymbol{\theta}$, $\boldsymbol{\theta}_{\mathrm{a}}$ involves both the unknown time-delay and parameters of the time-delay ExpAR model.

Let $\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t) := [\hat{\theta}_{\mathrm{a},1}(t), \ldots, \hat{\theta}_{\mathrm{a},2N}(t)]^{\mathrm{T}} \in \mathbb{R}^{2N}$ denote the estimation vector of $\boldsymbol{\theta}_{\mathrm{a}}$ at time t. Ideally, once $\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t)$ is obtained, the redundant parameter estimates should be zeros and the time-delay could then be determined from the structure of $\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t)$. In practice however, due to the presence of the non-zero noise term $v(t)$ in (1), the redundant parameter estimates will not be zeros and consequently, the time-delay can be misjudged. To address this problem, the redundant rule-based threshold (RR-TH) algorithms were presented in [16], [17] wherein a small threshold ϵ is applied to filter the parameter estimates. The key idea is to set $\hat{\theta}_{\mathrm{a},l}(t) = 0$ if $|\hat{\theta}_{\mathrm{a},l}(t)| < \epsilon$ for $l = 1, 2, \ldots, 2N$. Clearly, the choice of an adequate threshold ϵ is critical for the estimation accuracy. A larger ϵ may lead to the incorrect identification of model parameters as redundant parameters, while a smaller ϵ may lead to the opposite effects. In both cases, the time-delay and model parameters will be inaccurately estimated.

The objective of this letter is to present new estimation algorithms for the time-delay ExpAR model by exploiting the redundant rule and the multi-innovation theory.

III. Joint Parameter and Time-Delay Estimation Algorithms

Note that the identification model in (3) is highly nonlinear with respect to the parameter γ. Hence, Equation (4) is a complex nonlinear optimization problem and cannot be directly solved by using the least-squares or other related approaches [26].

A. Redundant Rule-Based Stochastic Gradient Algorithm

In this sub-section, we derive a basic redundant rule-based stochastic gradient (RR-SG) algorithm. For simplicity, define the parameter vector $\boldsymbol{\vartheta} := [\boldsymbol{\theta}_{\mathrm{a}}^{\mathrm{T}}, \gamma]^{\mathrm{T}} \in \mathbb{R}^{2N+1}$. Using the gradient search to solve the optimization problem in (4) gives the following recursive relations:

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mu(t)\boldsymbol{\psi}(t; \hat{\gamma}(t-1))e(t), \tag{5}$$

$$e(t) = y(t) - \boldsymbol{\phi}_{\mathrm{a}}^{\mathrm{T}}(t; \hat{\gamma}(t-1))\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t-1), \tag{6}$$

Authorized licensed use limited to: McGill University. Downloaded on November 21,2022 at 05:30:11 UTC from IEEE Xplore. Restrictions apply.

where $\mu(t)$ is the step-size, $e(t)$ is the innovation referring to [27], $\psi(t;\gamma)$ is the generalized information vector:

$$\boldsymbol{\psi}(t;\gamma) := [\boldsymbol{\phi}_{\mathrm{a}}^{\mathrm{T}}(t;\gamma), \hat{\boldsymbol{\theta}}_{\mathrm{a}}^{\mathrm{T}}(t-1)\boldsymbol{\phi}'_{\mathrm{a}}(t;\gamma)]^{\mathrm{T}} \in \mathbb{R}^{2N+1}, \quad (7)$$

and $\boldsymbol{\phi}'_{\mathrm{a}}(t;\gamma)$ is the derivative of $\boldsymbol{\phi}_{\mathrm{a}}(t;\gamma)$ with respect to γ:

$$\boldsymbol{\phi}'_{\mathrm{a}}(t;\gamma) := [\mathbf{0}, -y^2(t-1)\mathrm{e}^{-\gamma y^2(t-1)}\boldsymbol{\varphi}_{\mathrm{a}}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{2N}. \quad (8)$$

Several methods are available for the determination of the step-size $\mu(t)$. Referring to the previous work [24], we adopt the first-order Taylor expansion approximation and compute the optimal step-size through

$$\mu(t) := 1/r(t), \quad r(t) = r(t-1) + \|\boldsymbol{\psi}(t;\hat{\gamma}(t-1))\|^2. \quad (9)$$

Then, $\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t)$ and $\hat{\gamma}(t)$ can be updated by (5)–(9).

For the redundant rule-based threshold (RR-TH) estimation algorithms in [16], [17], an inappropriate threshold will lead to an incorrect time-delay estimate. Currently, no satisfactory method exists for the choice of an appropriate threshold, which must rely on an empirical and error prone approach. To overcome this critical limitation, we propose in this letter the following threshold-free procedure for estimating the unknown time-delay. Let L_e denote the length of the data sequence $y(t)$ on which the recursions in (5)–(9) are applied. Considering the structure of $\boldsymbol{\theta}_{\mathrm{a}}$, define the cost function

$$E(\tau) := \sum_{j=\tau+1}^{\tau+n} \left[\hat{\theta}^2_{\mathrm{a},j}(L_e) + \hat{\theta}^2_{\mathrm{a},j+N}(L_e)\right].$$

The time-delay estimate can be computed by

$$\hat{\tau} = \underset{\tau\in\{0,\ldots,\tau_{\max}\}}{\arg\max}\ E(\tau), \quad (10)$$

where $\tau_{\max} = N - n$. That is, the time-delay estimate maximizes the energy of the non-redundant parameter blocks of length n in the augmented identification model. Equations (5)–(10) form the RR-SG algorithm for the time-delay ExpAR model.

B. *Redundant Rule-Based Multi-Innovation Stochastic Gradient Algorithm*

Inspired by the multi-innovation theory in [18], [19], we expand the innovation $e(t)$ in (6) into a p-dimensional vector

$$\boldsymbol{E}(p,t) := \boldsymbol{Y}(p,t) - \boldsymbol{\Phi}_{\mathrm{a}}^{\mathrm{T}}(t;\hat{\gamma}(t-1))\hat{\boldsymbol{\theta}}_{\mathrm{a}}(t-1) \in \mathbb{R}^p, \quad (11)$$

where p is the innovation length, $\boldsymbol{Y}(p,t)$ is the stacked output vector consisting of $\{y(t-j)\}$ and $\boldsymbol{\Phi}_{\mathrm{a}}(t;\hat{\gamma}(t-1))$ is the stacked information matrix consisting of $\{\boldsymbol{\phi}_{\mathrm{a}}(t-j;\hat{\gamma}(t-1))\}$, $j = 0, 1, \ldots, p-1$.

To guarantee the compatibility of the matrix dimensions in various products, we expand $\psi(t;\hat{\gamma}(t-1))$ in (5) into the generalized stacked information matrix

$$\begin{aligned}\boldsymbol{\Psi}(p,t) := [&\boldsymbol{\psi}(t;\hat{\gamma}(t-1)), \boldsymbol{\psi}(t-1;\hat{\gamma}(t-1)), \ldots,\\ &\boldsymbol{\psi}(t-p+1;\hat{\gamma}(t-1))] \in \mathbb{R}^{(2N+1)\times p}. \quad (12)\end{aligned}$$

Then, the recursive relation in (5) can be expanded into

$$\hat{\boldsymbol{\vartheta}}(t) = \hat{\boldsymbol{\vartheta}}(t-1) + \mu(t)\boldsymbol{\Psi}(p,t)\boldsymbol{E}(p,t). \quad (13)$$

Equations (9)–(13) form the RR-MISG algorithm for the time-delay ExpAR model. Equation (13) can be seen as the discrete-time stochastic system of $\hat{\boldsymbol{\vartheta}}(t)$. In order to ensure the convergence of $\hat{\boldsymbol{\vartheta}}(t)$, it is required that all the eigenvalues are inside the unit circle [18].

Remark 1: Different from [16], [17] which employed the least-squares principle to derive several single-innovation RR-TH algorithms, this letter combines the gradient search with the multi-innovation theory and propose a novel RR-MISG algorithm using more innovations and observations.

Remark 2: Clearly, the RR-MISG algorithm has a larger computational complexity than the RR-SG algorithm because it involves the innovation extension. However, this extension enhances data utilization and improves the estimation accuracy for the RR-MISG algorithm.

Remark 3: In terms of system modeling, some techniques including the H_∞ filter and Lyapunov function have been used for linear and nonlinear systems [28], [29]. These techniques can be combined with the multi-innovation algorithm proposed in this letter to improve the data utilization and modeling accuracy.

IV. Examples

In this section, two examples are provided to test the effectiveness of the proposed algorithms. Example 1 serves to compare the estimation accuracy of the proposed algorithms with the RR-TH algorithm in [16], [17]. A more complex example highly related to a practical application in hydrology is presented in Example 2, to demonstrate the merits and comparative performance of the RR-SG and RR-MISG algorithms.

A. *Example 1*

Consider the time-delay ExpAR model with order $n = 2$, whose parameters and time-delay to be estimated are

$$\boldsymbol{\theta} = [\alpha_1, \alpha_2, \beta_1, \beta_2]^{\mathrm{T}} = [-0.10, 0.58, 1.23, 0.69]^{\mathrm{T}},$$

$\gamma = 2.00$, and $\tau = 1$. The selection of the model parameters should ensure that the model is stable. Assuming the maximum regression length $N = 4$, the augmented parameter vector is

$$\begin{aligned}\boldsymbol{\theta}_{\mathrm{a}} &= [\zeta_1, \alpha_1, \alpha_2, \zeta_2, \eta_1, \beta_1, \beta_2, \eta_2]^{\mathrm{T}}\\ &= [0, -0.10, 0.58, 0, 0, 1.23, 0.69, 0]^{\mathrm{T}}.\end{aligned}$$

In the simulations, $v(t)$ is taken as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$, and the data length is taken as $L_e = 6000$. The estimation errors $\delta := \|\hat{\boldsymbol{\vartheta}}(t) - \boldsymbol{\vartheta}\|/\|\boldsymbol{\vartheta}\| \times 100\%$ is used as a performance metric. For comparison purpose, we use the proposed RR-SG and RR-MISG algorithms, and the RR-TH algorithm in [16], [17] to jointly estimate the model parameters and time-delay. The RR-TH algorithm used here is based on the gradient search and is a single-innovation algorithm. Set the threshold of the RR-TH algorithm to $\epsilon_1 = 0.03$, $\epsilon_2 = 0.05$ and $\epsilon_3 = 0.11$.

Parameter estimates and their errors at time $t = L_e = 6000$ are shown in Tables I and II. RR-SG and RR-MISG parameter estimation errors δ versus t are shown in Fig. 1. From these tables and figure, we can see that: (1) the RR-TH algorithm cannot always set all the redundant parameters to zero for different thresholds and the time-delay estimate cannot always be

Authorized licensed use limited to: McGill University. Downloaded on November 21,2022 at 05:30:11 UTC from IEEE Xplore. Restrictions apply.

TABLE I
RR-TH AND RR-SG ESTIMATES AND ERRORS

Algorithms	ζ_1	α_1	α_2	ζ_2	η_1	β_1	β_2	η_2	γ	$\delta(\%)$
RR-TH (ϵ_1)	0.042	-0.128	0.517	0.000	0.101	1.100	0.771	0.041	1.476	22.35
RR-TH (ϵ_2)	0.000	-0.128	0.517	0.000	0.101	1.100	0.771	0.000	1.476	22.23
RR-TH (ϵ_3)	0.000	-0.128	0.517	0.000	0.000	1.100	0.771	0.000	1.476	21.87
RR-SG	0.000	-0.128	0.517	0.000	0.000	1.100	0.771	0.000	1.476	21.87
True Values	0.000	-0.100	0.580	0.000	0.000	1.230	0.690	0.000	2.000	

TABLE II
RR-SG ($p = 1$) AND RR-MISG ESTIMATES AND ERRORS

p	ζ_1	α_1	α_2	ζ_2	η_1	β_1	β_2	η_2	γ	$\delta(\%)$
1	0.000	-0.128	0.517	0.000	0.000	1.100	0.771	0.000	1.476	21.87
2	0.000	-0.115	0.553	0.000	0.000	1.167	0.725	0.000	1.741	10.74
4	0.000	-0.106	0.569	0.000	0.000	1.207	0.703	0.000	1.919	3.432
8	0.000	-0.102	0.574	0.000	0.000	1.219	0.697	0.000	1.987	0.791
True Values	0.000	-0.100	0.580	0.000	0.000	1.230	0.690	0.000	2.000	

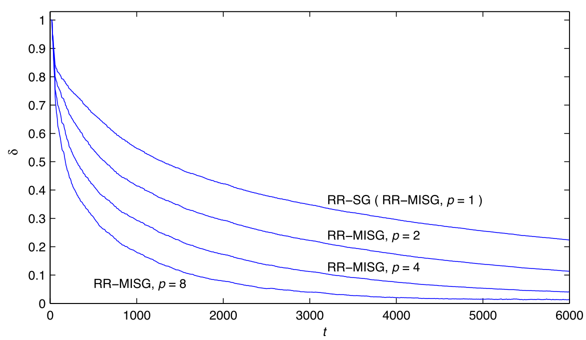

Fig. 1. RR-SG and RR-MISG estimation errors δ versus t.

effectively obtained; (2) the RR-SG and RR-MISG algorithms can give the true time-delay estimate and set all the redundant parameters to zero; (3) the RR-MISG algorithm provides higher estimation accuracy than the RR-SG algorithm at all time, and the estimation accuracy improves as the innovation length p increases.

B. Example 2

In hydrology sciences, it is well known that the river flows are affected by various factors, and the dynamics in their associated time-series exhibit complex and nonlinear behaviors [30]. In this example, the ExpAR model with the time-delay $\tau = 1$ and the proposed algorithms are used for fitting a set of daily mean water discharge time-series. To make this series more nearly symmetric and stabilize the variance, the zscore function in MatLab is used to normalize the original data. The transformed time-series $\{y(t) : t \in \mathbb{Z}\}$ is shown in Fig. 2, where the first $L_e = 4000$ data are used for the time-delay ExpAR model training, and the additional $L_r = 100$ data are used for the model validation. Employing the RR-SG algorithm and RR-MISG algorithm with $p = 6$ to predict this river discharge time-series, the fitting curves so obtained are shown in Fig. 3, while the errors between the actual and predictive outputs are shown in Fig. 4. Let $\hat{y}_1(t)$ and $\hat{y}_2(t)$ be the RR-SG and RR-MISG predictive outputs. To evaluate the prediction performance, we define the mean square error (MSE):

$$\mathrm{MSE}_i := \sqrt{\frac{1}{L_r}\sum_{t=L_e+1}^{L_e+L_r}[\hat{y}_i(t) - y(t)]^2}, \quad i = 1, 2.$$

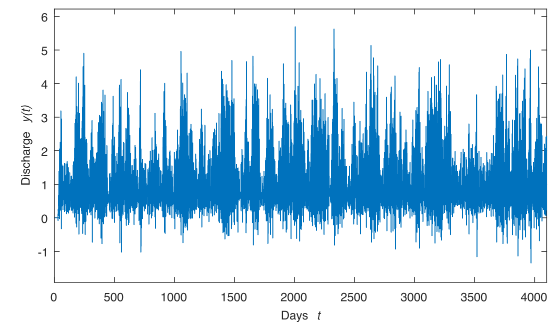

Fig. 2. Transformed daily mean water discharge time-series.

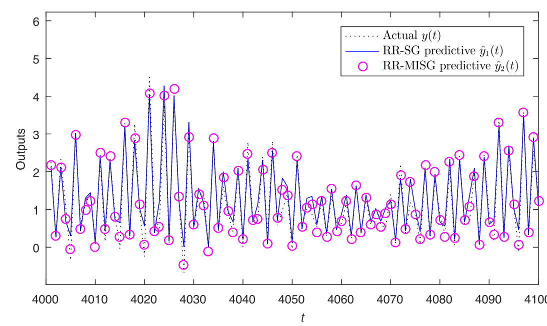

Fig. 3. Fitting results of the RR-SG and RR-MISG algorithms.

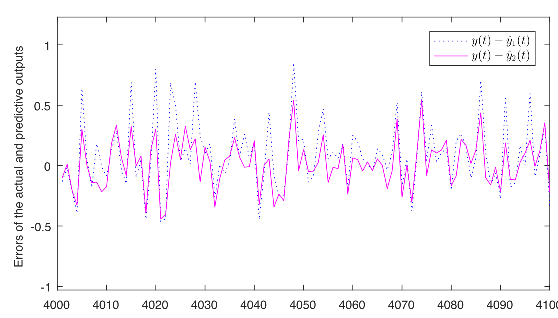

Fig. 4. Errors of the actual and predictive outputs.

Computing the MSEs of the RR-SG and RR-MISG algorithms gives $\mathrm{MSE}_1 = 0.2987$ and $\mathrm{MSE}_2 = 0.2033$, that is, -10.5 dB and -13.8 dB. This demonstrates that the smaller parameter estimation error of RR-MISG leads to the improved modeling accuracy, i.e., reduction of 3.3 dB in MSE. Figs. 3–4 show that the proposed RR-MISG algorithm has better prediction performance than the RR-SG algorithm.

V. CONCLUSION

This letter has studied the joint parameter and time-delay estimation problem for an extended version of the nonlinear ExpAR model. Two stochastic gradient estimation algorithms, referred to as RR-SG and RR-MISG, were proposed based on the redundant rule and multi-innovation theory. Compared with the existing threshold-based algorithms, the proposed algorithms avoid the effects of an empirical threshold and can effectively generate the time-delay estimate. Compared with the RR-SG algorithm, the RR-MISG provides improved estimation accuracy, with the improvement being more significant as the innovation length increases.

Authorized licensed use limited to: McGill University. Downloaded on November 21,2022 at 05:30:11 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] T. Ozaki, "The statistical analysis of perturbed limit cycle processes using nonlinear time-series models," *J. Time Ser. Anal.*, vol. 3, no. 1, pp. 29–41, Jan. 1982.

[2] H. Peng, T. Ozaki, V. Haggan-Ozaki, and Y. Toyoda, "A nonlinear exponential ARX model-based multivariable generalized predictive control strategy for thermal power plants," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 2, pp. 256–262, Mar. 2002.

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

[4] T. Chen and S. Y. Wang, "An efficient nonlinear dichotomous coordinate descent adaptive algorithm based on random Fourier features," *IEEE Signal Process. Lett.*, vol. 27, pp. 1804–1808, 2020.

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

[6] M. B. Priestley, "State-dependent models: A general approach to non-linear time series analysis," *J. Time Ser. Anal.*, vol. 1, no. 1, pp. 47–71, 1980.

[7] T. Ozaki, "Non-linear time-series models for non-linear random vibrations," *J. Appl. Probab.*, vol. 17, no. 1, pp. 84–93, 1980.

[8] V. Haggan and T. Ozaki, "Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time-series model," *Biometrika*, vol. 68, no. 1, pp. 189–196, Apr. 1981.

[9] R. Abdolee, B. Champagne, and A. H. Sayed, "Diffusion adaptation over multi-agent networks with wireless link impairments," *IEEE. Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1362–1376, Jun. 2016.

[10] M. V. Kulikova, J. V. Tsyganova, and G. Y. Kulikov, "UD-based pairwise and MIMO Kalman-like filtering for estimation of econometric model structures," *IEEE Trans. Autom. Control*, vol. 65, no. 10, pp. 4472–4479, Oct. 2020.

[11] Z. Shi, Y. Tamura, and T. Ozaki, "Monitoring the stability of BWR oscillation by nonlinear time-series modeling," *Ann. Nucl. Energy*, vol. 28, no. 10, pp. 953–966, Jul. 2001.

[12] S. Z. Mehreyan and A. Sayyareh, "Separated hypotheses testing for autoregressive models with non-negative residuals," *J. Statist. Comput. Simul.*, vol. 87, no. 4, pp. 689–711, Mar. 2017.

[13] G. Y. Chen, M. Gan, and G. L. Chen, "Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications," *Inf. Sci.*, vol. 438, pp. 46–57, Apr. 2018.

[14] B. Jelfs, S. Sun, K. Ghorbani, and C. Gilliam, "An adaptive all-pass filter for time-varying delay estimation," *IEEE Signal Process. Lett.*, vol. 28, pp. 628–632, 2021.

[15] F. Shang, B. Champagne, and I. Psaromiligkos, "A ML-based framework for joint TOA/AOA estimation of UWB pulses in dense multipath environments," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5305–5318, Oct. 2014.

[16] J. Chen, Q. M. Zhu, J. Li, and Y. J. Liu, "Biased compensation recursive least squares-based threshold algorithm for time-delay rational models via redundant rule," *Nonlinear Dyn.*, vol. 91, no. 2, pp. 797–807, Jan. 2018.

[17] X. Zhang, Q. Y. Liu, and T. Hayat, "Recursive identification of bilinear time-delay systems through the redundant rule," *J. Franklin Inst.*, vol. 357, no. 1, pp. 726–747, Jan. 2020.

[18] F. Ding and T. Chen, "Performance analysis of multi-innovation gradient type identification methods," *Automatica*, vol. 43, no. 1, pp. 1–14, Jan. 2007.

[19] F. Ding, X. P. Liu, and G. Liu, "Multiinnovation least-squares identification for system modeling," *IEEE Trans. Syst. Man Cybern. Part B-Cybern.*, vol. 40, no. 3, pp. 767–778, Jun. 2010.

[20] J. Pan, X. Jiang, X. K. Wan, and W. Ding, "A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems," *Int. J. Control Autom. Syst.*, vol. 15, no. 3, pp. 1189–1197, Jun. 2017.

[21] N. I. Chaudhary, M. A. Z. Raja, Y. G. He, Z. A. Khan, and J. A. T. Machado, "Design of multi innovation fractional LMS algorithm for parameter estimation of input nonlinear control autoregressive systems," *Appl. Math. Model.*, vol. 93, pp. 412–425, May 2021.

[22] L. J. Wang, Y. Ji, H. L. Yang, and L. Xu, "Decomposition-based multi-innovation gradient identification algorithms for a special bilinear system based on its input-output representation," *Int. J. Robust Nonlinear Control*, vol. 30, no. 9, pp. 3607–3623, Mar. 2020.

[23] H. F. Xia, Y. Ji, Y. J. Liu, and L. Xu, "Maximum likelihood-based multi-innovation stochastic gradient method for multivariable systems," *Int. J. Control Autom. Syst.*, vol. 17, no. 3, pp. 565–574, Mar. 2019.

[24] H. Xu, F. Ding, and E. F. Yang, "Modeling a nonlinear process using the exponential autoregressive time-series model," *Nonlinear Dyn.*, vol. 95, no. 3, pp. 2079–2092, Feb. 2019.

[25] H. Xu, F. Ding, and J. Sheng, "On some parameter estimation algorithms for the nonlinear exponential autoregressive model," *Int. J. Adapt. Control Signal Process.*, vol. 33, no. 6, pp. 999–1015, Jun. 2019.

[26] T. Zhang, S. Y. Wang, X. W. Huang, and L. Jia, "Kernel recursive least squares algorithm based on the Nyström method with k-means sampling," *IEEE Signal Process. Lett.*, vol. 27, pp. 361–365, 2020.

[27] L. Ljung, *System Identification: Theory for the User*. 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.

[28] X. Zhang, S. P. He, V. Stojanovic, X. L. Luan, and F. Liu, "Finite-time asynchronous dissipative filtering of conic-type nonlinear Markov jump systems," *Sci. China-Inf. Sci.*, vol. 64, no. 5, May 2021, Art. no. 152206.

[29] P. Cheng, M. Y. Chen, V. Stojanovic, and S. P. He, "Asynchronous fault detection filtering for piecewise homogenous Markov jump linear systems via a dual hidden Markov model," *Mech. Syst. Signal Proc.*, vol. 151, Apr. 2021, Art. no. 107353.

[30] E. Amiri, "Forecasting daily river flows using nonlinear time series models," *J. Hydrol.*, vol. 527, pp. 1054–1072, Aug. 2015.

Authorized licensed use limited to: McGill University. Downloaded on November 21,2022 at 05:30:11 UTC from IEEE Xplore. Restrictions apply.