Biased compensation recursive least squares based threshold algorithm for time-delay rational models via redundant rule[☆]

Jing Chen[*,a], Quanmin Zhu[b], Juan Li[c], Yanjun Liu[a]

[a] *School of Science, Jiangnan University, Wuxi 214122, PR China*
[b] *Department of Engineering Design and Mathematics, University of the West of England, Bristol BS16 1QY, UK*
[c] *College of Mechanical and Electrical Engineering, Qingdao Agricultural University, Qingdao 266109, PR China*

Abstract

This paper develops a biased compensation recursive least squares based threshold (BCRLS-TH) algorithm for a time-delay rational model. The time-delay rational model is first transformed into an augmented model by using the redundant rule, and then a RLS algorithm is proposed to estimate the parameters of the augmented model. Since the output of the augmented model is correlated with the noise, a biased compensation method is derived to eliminate the bias of the parameter estimates. Furthermore, based on the structures of the augmented model parameter vector and the rational model parameter vector, the unknown time-delay can be computed by using a threshold given in prior. A simulated example is used to illustrate the efficiency of the proposed algorithm.

Key words: Parameter estimation, Recursive least squares, Biased compensation, Time-delay, Rational model

1. Introduction

The rational model is a special kind of nonlinear systems which is defined as the ratio of two polynomial expressions [1]. This type of nonlinear models are widely existed in modern society, for example, in life science, in chemical engineering and in economic systems [2, 3]. Compared with the polynomial nonlinear model, the rational model provides a very concise and parsimonious representation for complex nonlinear systems and has excellent extrapolation properties [4, 5, 6]. Though the rational model has many advantages over the polynomial model and is widely used in many fields, surprisingly, there is only scattered work reported in the literature on identification of rational models. The difficulty in identifying the rational models is that the denominator polynomial terms make the rational model be nonlinear in both the parameters and the regression terms. Thus the traditional linear system and polynomial nonlinear system identification algorithms cannot be directly used for rational models [7, 8].

Among various identification algorithms, the off-line algorithms are perhaps the most widely used algorithms for rational models. For example, Zhu provided an implicit LS algorithm for rational models, the proposed off-line algorithm is efficient in dealing with the parameter estimation problems associated with nonlinear in the parameters models [9]. Mu et al studied a globally consistent nonlinear LS estimator for identification of a nonlinear rational system, where the proposed off-line algorithm is the first globally convergent algorithm for the nonlinear rational systems [10]. It has been noted that the off-line algorithms usually use the existed data to update the parameters in real time, while the new data are not involved. On the other hand, the on-line algorithms have less computational efforts and can update the parameters with new data. Therefore, the extension of on-line algorithms to rational model identification is becoming a hot and promising spot of present research. Recently, Zhu proposed an error back propagation parameter estimation algorithm for a class of rational models, by combining an orthogonal correlation test method, the model structure and the associate parameters can be estimated simultaneously [11]. Zhu et al also proposed an enhanced linear Kalman filter

[☆]This work was supported by the National Natural Science Foundation of China (No. 61403165), the Natural Science Foundation of Jiangsu Province (No. BK20131109) and the Natural Science Foundation for Colleges and Universities in Jiangsu Province (No. 16KJB120006).

[*]Corresponding author: School of Science, Jiangnan University, Wuxi 214122, PR China
Email addresses: chenjing1981929@126.com (Jing Chen), quan.zhu@uwe.ac.uk (Quanmin Zhu), lijuan291@sina.com (Juan Li), yanjunliu_1983@126.com (Yanjun Liu)

algorithm for parameter estimation of nonlinear rational models, in which the proposed algorithm is an on-line algorithm [12]. However, all the rational models in above literature are non-time-delay systems, when the rational models have unknown time-delays, those methods mentioned in above literature are invalid.

Time-delay systems often exist in engineering practice [13, 14, 15]. For example, in the communication network, due to the network congestion, the signals which are transmitted over a communication channel often have time delays [16, 17]; in chemical processes, some variables such as chemical component concentrations are often measured through laboratory analysis, which would introduce time delays [18, 19]. Recently, there exist many identification algorithms for time-delay systems [20, 21, 22]. Zhao et al developed a variational Bayesian (VB) approach for ARX models with a Markov chain time-varying time-delays, where the unknown parameters and the varying time-delays can be estimated iteratively [23]. Liu et al studied a compressed sensing (CS) recovery algorithm for MISO-FIR systems with unknown time-delays [24]. Both the VB and the CS recovery algorithms are off-line algorithms. In order to estimate the time-delay systems based on on-line algorithms, Ma et al provided a Kalman filter-based least squares iterative and a recursive least squares algorithms for time-delay Hammerstein systems with the assumption that the time-delay is known in prior [25]. Chen et al proposed a recursive least squares based redundant parameter method for time-delay systems, in which the time-delay is unknown [26]. To the best of our knowledge, there is no work reported in the literature on identification of time-delay rational models. Thus the focus of this paper is to develop an identification algorithm for such models.

In this paper, a BCRLS-TH algorithm is proposed for a time-delay rational model via redundant rule. The main objective is to apply the redundant rule to transform the rational model into an augmented model, whose parameters are estimated by using a BCRLS algorithm. Since the estimated parameter vector contains redundant parameters, a threshold is introduced to pick out these redundant parameters. Then based on the structures of the augmented model parameter vector and the rational model parameter vector, the unknown time-delay can be obtained. The main contributions are summarized as follows.

1. Study a biased compensation on-line algorithm for a time-delay rational model.
2. Apply the redundant rule to transform the rational model into an augmented model, by which the time-delay effect can be ignored.
3. Propose a threshold to divide the augmented model parameter estimates into two parts: the redundant model parameter estimates and the rational model parameter estimates, then the unknown time-delay can be obtained based on the two parts.

Briefly, the rest of this paper is organized as follows. Section 2 introduces the rational model. Section 3 develops a biased compensation least squares based threshold (BCLS-TH) algorithm. Section 4 studies a BCRLS-TH algorithm. Section 5 provides an illustrative example. Finally, concluding remarks are given in Section 6.

2. The rational model

Consider the following rational model,

$$y(t) = \frac{\mathbf{a}(t-\tau)}{\mathbf{b}(t-\tau)} + e(t), \tag{1}$$

where $y(t)$ is the output, $e(t)$ a stochastic white noise with zero mean and variance σ^2, τ is an unknown integer time-delay, and $\mathbf{a}(t-\tau)$ and $\mathbf{b}(t-\tau)$ are expressed as

$$\begin{aligned}\mathbf{a}(t-\tau) &= \boldsymbol{\varphi}^{\mathrm{T}}(t-\tau)\boldsymbol{\theta}_a,\\ \mathbf{b}(t-\tau) &= \boldsymbol{\psi}^{\mathrm{T}}(t-\tau)\boldsymbol{\theta}_b.\end{aligned}$$

The information vectors $\boldsymbol{\varphi}(t-\tau)$ and $\boldsymbol{\psi}(t-\tau)$ are the products of past inputs $\{u(t-\tau-1), u(t-\tau-2), \cdots\}$ and past outputs $\{y(t-\tau-1), y(t-\tau-2), \cdots\}$, such as $y(t-\tau-1)u(t-\tau-1)$, $u(t-\tau-2)$ and $u^2(t-\tau-1)y(t-1-\tau)$, and the structures of $\boldsymbol{\varphi}(t-\tau)$ and $\boldsymbol{\psi}(t-\tau)$ are known in prior, $\boldsymbol{\theta}_a$ and $\boldsymbol{\theta}_b$ are the unknown parameters to be estimated and can be expressed as

$$\boldsymbol{\theta}_a = [a_1, a_2, \cdots, a_n]^{\mathrm{T}}, \tag{2}$$

$$\boldsymbol{\theta}_b = [b_1, b_2, \cdots, b_m]^{\mathrm{T}}. \tag{3}$$

Without loss of generality, assume $b_1 = 1$, then one can get

$$Y(t) = \boldsymbol{\varphi}^{\mathrm{T}}(t-\tau)\boldsymbol{\theta}_a + y(t)[-\boldsymbol{\psi}^{\mathrm{T}}(t-\tau)\boldsymbol{\theta}_b + b_1\psi_1(t-\tau)] + \mathbf{b}(t-\tau)e(t), \tag{4}$$

in which $Y(t) = y(t)b_1\psi_1(t-\tau)$, and $\psi_1(t-\tau)$ is the first element of the vector $\boldsymbol{\psi}(t-\tau)$. Rewriting Equation (4) as

$$Y(t) = \boldsymbol{\phi}^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t), \tag{5}$$

where

$$\begin{aligned}
\boldsymbol{\theta} &= [a_1, a_2, \cdots, a_n, b_2, \cdots, b_m]^{\mathrm{T}} \in \mathbb{R}^{n+m-1},\\
\boldsymbol{\phi}(t) &= [\varphi_1(t-\tau), \varphi_2(t-\tau), \cdots, \varphi_n(t-\tau), -y(t)\psi_2(t-\tau), \cdots, -y(t)\psi_m(t-\tau)]^{\mathrm{T}} \in \mathbb{R}^{n+m-1},\\
v(t) &= \mathbf{b}(t-\tau)e(t).
\end{aligned}$$

Since the time-delay τ is unknown, those elements in $\boldsymbol{\phi}(t)$ are also unknown, then the LS proposed in [10] and the enhanced linear Kalman filter algorithm proposed in [12] cannot be applied for this time-delay rational model. To overcome this difficulty is using the redundant rule proposed in [26].

Define

$$\begin{aligned}
\boldsymbol{\phi}_p(t) := [&\varphi_1(t), \cdots, \varphi_n(t), -y(t)\psi_2(t), \cdots, -y(t)\psi_m(t), \cdots, \varphi_1(t-\tau), \cdots, \varphi_n(t-\tau), -y(t)\psi_2(t-\tau), \cdots,\\
&-y(t)\psi_m(t-\tau), \cdots, \varphi_1(t-M+1), \cdots, \varphi_n(t-M+1), -y(t)\psi_2(t-M+1), \cdots,\\
&-y(t)\psi_m(t-M+1)]^{\mathrm{T}} \in \mathbb{R}^{M(n+m-1)},\\
\boldsymbol{\theta}_p := [&l_{1,1}, \cdots, l_{1,n}, l_{1,n+1}, \cdots, l_{1,n+m-1}, \cdots, a_1, \cdots, a_n, b_2, \cdots, b_m, \cdots, l_{M,1} \cdots, l_{M,n}, l_{M,n+1}, \cdots,\\
&l_{M,n+m-1}]^{\mathrm{T}} \in \mathbb{R}^{M(n+m-1)}.
\end{aligned} \tag{6}$$

Clearly, $l_{s,j}, s = 1, \cdots, \tau, \tau+2, \cdots, M, j = 1, \cdots, n+m-1$ are the redundant parameters and the true values of these parameters are equal to zero, $a_1, \cdots, a_n, b_2, \cdots, b_m$ are the true parameters, and all the terms in the information vector $\boldsymbol{\phi}_p(t)$ are known. Then Equation (5) can be transformed into the following augmented model,

$$Y(t) = \boldsymbol{\phi}_p^{\mathrm{T}}(t)\boldsymbol{\theta}_p + v(t). \tag{7}$$

It is observed from Equations (6) and (7) that the augmented model parameter vector contains two parts, one is the redundant part which can be expressed as

$$\boldsymbol{\theta}_{rp} = [l_{1,1}, \cdots, l_{1,n}, l_{1,n+1}, \cdots, l_{1,n+m-1}, \cdots, l_{M,1}, \cdots, l_{M,n}, l_{M,n+1}, \cdots, l_{M,n+m-1}]^{\mathrm{T}} \in \mathbb{R}^{(M-1)(n+m-1)}, \tag{8}$$

and the other is the true part (rational model parameter vector) which can be written as

$$\boldsymbol{\theta} = [a_1, a_2, \cdots, a_n, b_2, \cdots, b_m,]^{\mathrm{T}} \in \mathbb{R}^{n+m-1}. \tag{9}$$

The idea of the redundant rule is first to pick out the redundant part from the augmented model parameter vector $\boldsymbol{\theta}_p$ by using a threshold given in prior, and then compute the unknown time-delay based on the structures of $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_p$. For example, if the unknown time-delay is 2, the order of the true parameter is $n+m-1 = 5$. Assume $M = 4$, when apply the LS based threshold method to estimate the unknown parameter vector $\boldsymbol{\theta}_p$. The estimate of the augmented model parameter vector is $\hat{\boldsymbol{\theta}}_p = [\mathbf{0_5}, \mathbf{0_5}, \hat{\boldsymbol{\theta}}, \mathbf{0_5}]^{\mathrm{T}}$ ($\mathbf{0}_5 = [0, 0, 0, 0, 0]$), which means the position of the first non-zero value is 11. Thus we can get $\tau = 10/5 = 2$.

3. The BCLS-TH algorithm

Collect t input, output and noise data respectively, and define

$$\begin{aligned}
\boldsymbol{Y}(t) &:= [Y(t), Y(t-1), \cdots, Y(1)]^{\mathrm{T}},\\
\boldsymbol{\Phi}_p(t) &:= [\boldsymbol{\phi}_p(t), \boldsymbol{\phi}_p(t-1), \cdots, \boldsymbol{\phi}_p(1)]^{\mathrm{T}},\\
\boldsymbol{V}(t) &:= [v(t), v(t-1), \cdots, v(1)]^{\mathrm{T}}.
\end{aligned}$$

Then it follows that

$$\boldsymbol{Y}(t) = \boldsymbol{\Phi}_p(t)\boldsymbol{\theta}_p + \boldsymbol{V}(t). \tag{10}$$

Based on the LS algorithm, one can get

$$\hat{\boldsymbol{\theta}}_{LS}(t) = [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{Y}(t). \tag{11}$$

Define

$$\boldsymbol{P}(t) := [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}.$$

Substituting Equation (10) into Equation (11) gets

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{LS}(t) &= [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{Y}(t)\\ &= [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)(\boldsymbol{\Phi}_p(t)\boldsymbol{\theta}_p + \boldsymbol{V}(t))\\ &= \boldsymbol{\theta}_p + \boldsymbol{P}(t)[\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}(t)].\end{aligned} \tag{12}$$

Rewrite

$$\begin{aligned}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}(t) &= [\boldsymbol{\phi}_p(t), \boldsymbol{\phi}_p(t-1), \cdots, \boldsymbol{\phi}_p(1)]\begin{bmatrix} v(t)\\ v(t-1)\\ \vdots\\ v(1)\end{bmatrix}\\ &= [\boldsymbol{\phi}_p(t), \boldsymbol{\phi}(t-1), \cdots, \boldsymbol{\phi}(1)]\begin{bmatrix} \mathbf{b}(t-\tau)e(t)\\ \mathbf{b}(t-1-\tau)e(t-1)\\ \vdots\\ \mathbf{b}(1-\tau)e(1)\end{bmatrix}\\ &= \Big[\sum_{i=1}^{t}\mathbf{b}(i-\tau)\varphi_1(i)e(i), \cdots, \sum_{i=1}^{t}\mathbf{b}(i-\tau)\varphi_n(i)e(i), -\sum_{i=1}^{t}\mathbf{b}(i-\tau)\psi_2(i)y(i)e(i), \cdots,\\ &\quad -\sum_{i=1}^{t}\mathbf{b}(i-\tau)\psi_m(i)y(i)e(i), \cdots, \sum_{i=1}^{t}\mathbf{b}(i-\tau)\varphi_1(i-M+1)e(i), \cdots, \sum_{i=1}^{t}\mathbf{b}(i-\tau)\varphi_n(i-M+1)e(i),\\ &\quad -\sum_{i=1}^{t}\mathbf{b}(i-\tau)\psi_2(i-M+1)y(i)e(i), \cdots, -\sum_{i=1}^{t}\mathbf{b}(i-\tau)\psi_m(i-M+1)y(i)e(i)\Big]^{\mathrm{T}}.\end{aligned} \tag{13}$$

Since $y(t)$ is correlated with the noise $e(t)$, Equation (12) means that the proposed LS algorithm is a biased algorithm. It follows that an unbiased estimate $\hat{\boldsymbol{\theta}}_u(t)$ can be expressed as,

$$\hat{\boldsymbol{\theta}}_u(t) = \hat{\boldsymbol{\theta}}_{LS}(t) - \boldsymbol{P}(t)[\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}(t)]. \tag{14}$$

Unfortunately, $\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}(t)$ cannot be obtained because of the unknown polynomial $\mathbf{b}(i-\tau)$ and unknown noise $e(i)$ in Equation (13). That is to say, in order to obtain the unbiased estimate $\hat{\boldsymbol{\theta}}_u(t)$, one should first obtain the estimates of $\mathbf{b}(i-\tau)$ and $e(i)$. At time $t-1$, the unknown noise $e(i), i = 1, 2, \cdots, t$ can be estimated by using the parameter estimate $\hat{\boldsymbol{\theta}}_u(t-1)$, e.g.,

$$\hat{e}(i) = y(i) - \frac{\mathbf{a}(i)}{\mathbf{b}(i)}|_{\hat{\boldsymbol{\theta}}_u(t-1)}, \tag{15}$$

in which $(x)|_{\hat{\boldsymbol{\theta}}_u(t-1)}$ means the estimate of x by using $\hat{\boldsymbol{\theta}}_u(t-1)$. Unlike the work in [9, 10, 11], the colored noise $v(t)$ in this paper contains unknown time-delay τ, which leads $\mathbf{b}(i-\tau)$ also be unknown. Thus, to get the unbiased parameter estimate is more challenging in this paper. One will use the redundant rule to transform the rational model into an augmented model. Then Equation (1) can be expressed as

$$y(t) = \frac{\mathbf{a}_0(t) + \mathbf{a}_1(t-1) + \cdots + \mathbf{a}(t-\tau) + \cdots + \mathbf{a}_{M-1}(t-M+1)}{\mathbf{b}_0(t) + \mathbf{b}_1(t-1) + \cdots + \mathbf{b}(t-\tau) + \cdots + \mathbf{b}_{M-1}(t-M+1)} + e(t), \tag{16}$$

where

$$\begin{aligned}&\mathbf{a}_p(t-p) = \varphi_1(t-p)l_{2,1} + \varphi_2(t-p)l_{2,2} + \cdots + \varphi_n(t-p)l_{2,n}, p = 0, 1, \cdots, \tau-1, \tau+1, \cdots, M-1,\\ &\mathbf{a}(t-\tau) = \varphi_1(t-\tau)a_1 + \varphi_2(t-\tau)a_2 + \cdots + \varphi_n(t-\tau)a_n, \quad \text{true part},\\ &\mathbf{b}_p(t-p) = \psi_2(t-p)l_{1,n+1} + \psi_3(t-p)l_{1,n+2} + \cdots + \psi_m(t-p)l_{1,n+m-1},\\ &\mathbf{b}(t-\tau) = 1 + \psi_2(t-\tau)b_2 + \psi_3(t-\tau)b_3 + \cdots + \psi_m(t-\tau)b_m, \quad \text{true part}.\end{aligned} \tag{17}$$

Equation (7) can be enhanced as

$$Y(t) = \boldsymbol{\phi}_p^{\mathrm{T}}(t)\boldsymbol{\theta}_p + v_e(t), \tag{18}$$

in which

$$v_e(t) = [\mathbf{b}_0(t) + \mathbf{b}_1(t-1) + \cdots + \mathbf{b}(t-\tau) + \cdots + \mathbf{b}_{M-1}(t-M+1)]e(t).$$

Clearly, the true values of $l_{s,j}, s = 1, 2, \cdots, \tau, \tau+2, \cdots, M, j = 1, \cdots, n+m-1$ are all equal 0, thus Equation (18) is equivalent to Equation (7). Since $e(i)$ is a white noise with zero mean and is independent of the past inputs and outputs, Equation (13) can be converted as follows,

$$\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}_e(t) = [\boldsymbol{\Psi}_0^{\mathrm{T}}(t), \cdots, \boldsymbol{\Psi}_q^{\mathrm{T}}(t), \cdots, \boldsymbol{\Psi}_{M-1}^{\mathrm{T}}(t)]^{\mathrm{T}},\ q = 0, 1, \cdots, M-1, \tag{19}$$

where

$$\boldsymbol{\Psi}_q(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\sum_{i=1}^{t}(\mathbf{b}_0(i) + \mathbf{b}_1(i-1) + \cdots + \mathbf{b}(i-\tau) + \cdots + \mathbf{b}_{M-1}(i-M+1))\psi_2(i-q)y(i)\hat{e}(i) \\ \vdots \\ -\sum_{i=1}^{t}(\mathbf{b}_0(i) + \mathbf{b}_1(i-1) + \cdots + \mathbf{b}(i-\tau) + \cdots + \mathbf{b}_{M-1}(i-M+1))\psi_m(i-q)y(i)\hat{e}(i) \end{bmatrix} \in \mathbb{R}^{n+m-1},$$

and the unbiased parameter estimate $\hat{\boldsymbol{\theta}}_u(t)$ can be computed as

$$\hat{\boldsymbol{\theta}}_u(t) = \hat{\boldsymbol{\theta}}_{LS}(t) - \boldsymbol{P}(t)\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{V}_e(t). \tag{20}$$

From Equations (19) and (20), we can see that at time t, $\mathbf{b}_0(i)+\mathbf{b}_1(i-1)+\cdots+\mathbf{b}(i-\tau)+\cdots+\mathbf{b}_{M-1}(i-M+1), i = 1, \cdots, t$, can be estimated by using the unbiased parameter estimate $\hat{\boldsymbol{\theta}}_u(t-1)$ without the knowledge of the unknown time-delay τ. Assume that at time t, we have get the unbiased estimate $\hat{\boldsymbol{\theta}}_u(t)$,

$$\begin{aligned}\hat{\boldsymbol{\theta}}_u(t) = [&l_{1,1}(t), \cdots, l_{1,n}(t), l_{1,n+1}(t), \cdots, l_{1,n+m-1}(t), \cdots, a_1(t), \cdots, a_n(t), b_2(t), \cdots, b_m(t), \cdots, \\ &l_{M,1}(t), \cdots, l_{M,n}(t), l_{M,n+1}(t), \cdots, l_{M,n+m-1}(t)]^{\mathrm{T}}.\end{aligned} \tag{21}$$

Since $Y(t)$ only depends on $\varphi_1(t-\tau), \varphi_2(t-\tau), \cdots, \varphi_n(t-\tau), -y(t)\psi_2(t-\tau), \cdots, -y(t)\psi_m(t-\tau)$, once the parameter vector has been estimated, all the redundant parameters $l_{s,j}$ are equal zero, while the rational model parameter estimates can be estimated as

$$\hat{\boldsymbol{\theta}}(t) = [a_1(t), a_2(t), \cdots, a_n(t), b_2(t), \cdots, b_m(t)]^{\mathrm{T}} \in \mathbb{R}^{n+m-1}.$$

Furthermore, according to the position of $\hat{\boldsymbol{\theta}}$ in $\hat{\boldsymbol{\theta}}_u$, we can get the unknown time-delay.

Then the following BCLS-TH algorithm can be summarized as follows,

$$\begin{aligned}\hat{\boldsymbol{\theta}}_u^k(t) &= [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{Y}(t) - \boldsymbol{P}(t)\hat{\rho}^k(t), \\ \hat{\rho}^k(t) &= = [\hat{\boldsymbol{\Psi}}_{0,k}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\Psi}}_{q,k}^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\Psi}}_{M-1,k}^{\mathrm{T}}(t)]^{\mathrm{T}},\ q = 0, 1, \cdots, M-1, \\ \hat{\boldsymbol{\Psi}}_{q,k}(t) &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\sum_{i=1}^{t}(\hat{\mathbf{b}}_0^k(i) + \hat{\mathbf{b}}_1^k(i-1) + \cdots + \hat{\mathbf{b}}^k(i-\tau) + \cdots + \hat{\mathbf{b}}_{M-1}^k(i-M+1))\psi_2(i-q)y(i)\hat{e}^k(i) \\ \vdots \\ -\sum_{i=1}^{t}(\hat{\mathbf{b}}_0^k(i) + \hat{\mathbf{b}}_1^k(i-1) + \cdots + \hat{\mathbf{b}}^k(i-\tau) + \cdots + \hat{\mathbf{b}}_{M-1}^k(i-M+1))\psi_m(i-q)y(i)\hat{e}^k(i) \end{bmatrix},\end{aligned}$$

$$\hat{e}^k(i) = y(i) - \frac{\hat{\mathbf{a}}^k(i-\tau)}{\hat{\mathbf{b}}^k(i-\tau)}, \quad i = 1, \cdots, t,$$
$$\hat{\mathbf{a}}_r^k(i) = [\mathbf{a}_r(i)]|_{\hat{\boldsymbol{\theta}}_u^{k-1}(t)}, \quad r = 0, 1, \cdots, M-1,$$
$$\hat{\mathbf{b}}_r^k(i) = [\mathbf{b}_r(i)]|_{\hat{\boldsymbol{\theta}}_u^{k-1}(t)},$$

where k represents the count of BCLS-TH iteration.

However, the BCLS-TH algorithm is an off-line algorithm which has harder computational efforts and cannot update the parameters when new data becomes available. In order to overcome this difficulty, a BCRLS-TH algorithm is introduced in the next Section.

4. The BCRLS-TH algorithm

Define

$$\boldsymbol{P}(t) := [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}.$$

From Equation (11), one can get

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{LS}(t) &= [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{Y}(t) \\
&= \boldsymbol{P}(t)\begin{bmatrix} \boldsymbol{\Phi}_p(t-1) \\ \boldsymbol{\phi}_p^{\mathrm{T}}(t) \end{bmatrix}^{\mathrm{T}}\begin{bmatrix} \boldsymbol{Y}(t-1) \\ Y(t) \end{bmatrix} \\
&= \boldsymbol{P}(t)[\boldsymbol{P}^{-1}(t-1)\boldsymbol{P}(t-1)\boldsymbol{\Phi}_p^{\mathrm{T}}(t-1)\boldsymbol{Y}(t-1) + \boldsymbol{\phi}_p(t)Y(t)] \\
&= \boldsymbol{P}(t)[\boldsymbol{P}^{-1}(t-1)\hat{\boldsymbol{\theta}}_{LS}(t-1) + \boldsymbol{\phi}_p(t)Y(t)] \\
&= \hat{\boldsymbol{\theta}}_{LS}(t-1) + \boldsymbol{P}(t)\boldsymbol{\phi}_p(t)[Y(t) - \boldsymbol{\phi}_p^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{LS}(t-1)].
\end{aligned} \tag{22}$$

Then the BCRLS-TH algorithm can be summarized as follows,

$$\hat{\boldsymbol{\theta}}_u(t) = \hat{\boldsymbol{\theta}}_{LS}(t) - \boldsymbol{P}(t)\hat{\rho}(t), \tag{23}$$
$$\hat{\boldsymbol{\theta}}_{LS}(t) = \hat{\boldsymbol{\theta}}_{LS}(t-1) + \boldsymbol{P}(t)\boldsymbol{\phi}_p(t)[Y(t) - \boldsymbol{\phi}_p^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{LS}(t-1)], \tag{24}$$
$$\boldsymbol{P}(t) = [\boldsymbol{\Phi}_p^{\mathrm{T}}(t)\boldsymbol{\Phi}_p(t)]^{-1}, \tag{25}$$
$$\boldsymbol{\Phi}_p(t) = [\boldsymbol{\phi}_p(t), \boldsymbol{\phi}_p(t-1), \cdots, \boldsymbol{\phi}_p(1)]^{\mathrm{T}}, \tag{26}$$
$$\begin{aligned}
\boldsymbol{\phi}_p(t) = [&\varphi_1(t), \cdots, \varphi_n(t), -y(t)\psi_2(t), \cdots, -y(t)\psi_m(t), \cdots, \varphi_1(t-\tau), \cdots, \varphi_n(t-\tau), \\
&-y(t)\psi_2(t-\tau), \cdots, -y(t)\psi_m(t-\tau), \cdots, \varphi_1(t-M+1), \cdots, \varphi_n(t-M+1), \\
&-y(t)\psi_2(t-M+1), \cdots, -y(t)\psi_m(t-M+1)]^{\mathrm{T}},
\end{aligned} \tag{27}$$
$$\hat{e}(i) = y(i) - \frac{\hat{\mathbf{a}}(i-\tau)}{\hat{\mathbf{b}}(i-\tau)}, \quad i = 1, \cdots, t, \tag{28}$$

$$\hat{\rho}(t) = = [\hat{\boldsymbol{\Psi}}_0^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\Psi}}_q^{\mathrm{T}}(t), \cdots, \hat{\boldsymbol{\Psi}}_{M-1}^{\mathrm{T}}(t)]^{\mathrm{T}}, \ q = 0, 1, \cdots, M-1, \tag{29}$$

$$\hat{\boldsymbol{\Psi}}_q(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\sum_{i=1}^{t}(\hat{\mathbf{b}}_0(i) + \hat{\mathbf{b}}_1(i-1) + \cdots + \hat{\mathbf{b}}(i-\tau) + \cdots + \hat{\mathbf{b}}_{M-1}(i-M+1))\psi_2(i-q)y(i)\hat{e}(i) \\ \vdots \\ -\sum_{i=1}^{t}(\hat{\mathbf{b}}_0(i) + \hat{\mathbf{b}}_1(i-1) + \cdots + \hat{\mathbf{b}}(i-\tau) + \cdots + \hat{\mathbf{b}}_{M-1}(i-M+1))\psi_m(i-q)y(i)\hat{e}(i) \end{bmatrix}, \tag{30}$$

$$\hat{\mathbf{a}}_r(i) = [\mathbf{a}_j(i)]|_{\hat{\boldsymbol{\theta}}_u(t-1)}, \quad r = 0, 1, \cdots, M-1, \tag{31}$$
$$\hat{\mathbf{b}}_r(i) = [\mathbf{b}_j(i)]|_{\hat{\boldsymbol{\theta}}_u(t-1)}. \tag{32}$$

It observes from Equation (31) that we should use the threshold to compute the parameter $\hat{\boldsymbol{\theta}}(t)$ at each time t based on $\hat{\boldsymbol{\theta}}_u(t)$, which leads to heavy computational efforts. In order to overcome this difficulty, one can apply the following equation to compute the noise $e(i)$

$$\hat{e}(i) = y(i) - \frac{\hat{\mathbf{a}}_0(i) + \hat{\mathbf{a}}_1(i-1) + \cdots + \hat{\mathbf{a}}(i-\tau) + \cdots + \hat{\mathbf{a}}_{M-1}(t-M+1)}{\hat{\mathbf{b}}_0(i) + \hat{\mathbf{b}}_1(i-1) + \cdots + \hat{b}(i-\tau) + \cdots + \hat{\mathbf{b}}_{M-1}(t-M+1)}\Big|_{\hat{\boldsymbol{\theta}}_{u(t-1)}}, i = 1, \cdots, t. \tag{33}$$

Remark 1: In the BCLS-TH algorithm, we use the same input and output data to update the parameter vector at each iteration k; while in the BCRLS-TH algorithm, we update the parameter vector at each time t based on the new input and output data $u(t), y(t)$ and the used data $u(t-1), \cdots, u(1), y(t-1), \cdots, y(1)$, thus the BCRLS-TH algorithm can update the parameters in real time when new data becomes available.

The steps of computing the parameter estimation vector $\hat{\boldsymbol{\theta}}(t)$ by using the BCRLS-TH algorithm are listed in the following.

1. Let $\hat{\boldsymbol{\theta}}_{LS}(0) = \mathbf{1}/p_0$, $\hat{\boldsymbol{\theta}}_u(0) = \mathbf{1}/p_0$ and $\hat{\boldsymbol{P}}(0) = p_0\boldsymbol{I}$ with $\mathbf{1}$ being a column vector whose entries are all unity, $\boldsymbol{I}$ be an identity matrix of appropriate size and $p_0 = 10^6$.
2. Let $t = 1$, $y(h) = 0, u(h) = 0, e(h) = 0$, $h \leqslant 0$, and give a small positive number ε and a small positive threshold ϵ.
3. Form $\boldsymbol{\phi}_p(t)$ by (27).
4. Form $\boldsymbol{\Phi}(t)$ by (26) and $\boldsymbol{P}(t)$ by (25).
5. Compute $\hat{\mathbf{a}}_r(i)$ and $\hat{\mathbf{b}}_r(i)$, $r = 0, 1, \cdots, M-1$, $i = 1, 2, \cdots, t$ by (31) and (32), respectively.
6. Compute $\hat{e}(i)$ by (33), $i = 1, 2, \cdots, t$.
7. Compute $\hat{\boldsymbol{\Psi}}_q(t)$ by (30), $q = 0, \cdots, M-1$.
8. Compute $\hat{\rho}(t)$ by (29).
9. Update the parameter estimation vector $\hat{\boldsymbol{\theta}}_{LS}(t)$ by (24).
10. Compute $\hat{\boldsymbol{\theta}}_u(t)$ by (23).
11. Compare $\hat{\boldsymbol{\theta}}_u(t)$ and $\hat{\boldsymbol{\theta}}_u(t-1)$: if $\|\hat{\boldsymbol{\theta}}_u(t) - \hat{\boldsymbol{\theta}}_u(t-1)\| \leqslant \varepsilon$, then obtain the $\hat{\boldsymbol{\theta}}_u(t)$ and go to next step; otherwise, increase t by 1 and go to step 3.
12. Assume each element of $\hat{\boldsymbol{\theta}}_u(t)$ is $\hat{\theta}_u^s(t), s = 1, \cdots, Mn + Mm - M$, and compare $\hat{\theta}_u^s(t)$ with ϵ: if $|\hat{\theta}_u^s(t)| < \epsilon$, then pick it out and reduce the order of the parameter vector $\hat{\boldsymbol{\theta}}_u(t)$ by 1; otherwise, keep the order of the parameter vector $\hat{\boldsymbol{\theta}}_u(t)$ unchanged.
13. Obtain $\hat{\boldsymbol{\theta}}(t)$ based on step 12.
14. Compute the time-delay τ based on the structures of $\hat{\boldsymbol{\theta}}_u(t)$ and $\hat{\boldsymbol{\theta}}(t)$.

Remark 2: In application, systems are often disturbed by noises, which leads to the estimates of the redundant parameters not be equal to zero. In order to get the time-delay τ, a threshold ϵ should be given in prior. If the absolute value of a parameter estimate is smaller than ϵ, one can regard the parameter as a redundant parameter and pick it out. However, the choice of the threshold is difficult and challenging. A big ϵ may mistake some true parameters for redundant parameters, while a small ϵ may lead some redundant parameters not be picked out from the parameter vector [26].

5. Example

Consider a time-delay rational model with $\tau = 1$,

$$\begin{aligned} y(t) &= \frac{0.2y(t-\tau-1) + 0.1y(t-\tau-1)u(t-\tau-1) + u(t-\tau-1)}{1 + y^2(t-\tau-1) + y^2(t-\tau-2)} + e(t) \\ &= \frac{0.2y(t-2) + 0.1y(t-2)u(t-2) + u(t-2)}{1 + y^2(t-2) + y^2(t-3)} + e(t). \end{aligned} \tag{34}$$

Then one can get

$$\begin{aligned} y(t) = {} & 0.2y(t-2) + 0.1y(t-2)u(t-2) + u(t-2) - \\ & y(t)y^2(t-2) - y(t)y^2(t-3) + (1 + y^2(t-2) + y^2(t-3))e(t). \end{aligned}$$

Assume $M = 3$, then one can get the augmented model,

$$\begin{aligned}
y(t) =& l_{1,1}y(t-1) + l_{1,2}y(t-1)u(t-1) + l_{1,3}u(t-1) - l_{1,4}y(t)y^2(t-1) - l_{1,5}y(t)y^2(t-2) + a_1y(t-2) + \\
& a_2y(t-2)u(t-2) + a_3u(t-2) - b_2y(t)y^2(t-2) - b_3y(t)y^2(t-3) + l_{3,1}y(t-3) + l_{3,2}y(t-3)u(t-3) + \\
& l_{3,3}u(t-3) - l_{3,4}y(t)y^2(t-3) - l_{3,5}y(t)y^2(t-4) + (1 + y^2(t-2) + y^2(t-3))e(t), \\
\boldsymbol{\theta}_p =& [l_{1,1}, l_{1,2}, l_{1,3}, l_{1,4}, l_{1,5}, a_1, a_2, a_3, b_2, b_3, l_{3,1}, l_{3,2}, l_{3,3}, l_{3,4}, l_{3,5}]^{\mathrm{T}} \\
=& [0, 0, 0, 0, 0, 0.2, 0.1, 1, 1, 1, 0, 0, 0, 0, 0]^{\mathrm{T}}, \\
\boldsymbol{\phi}_p(t) =& [y(t-1), y(t-1)u(t-1), u(t-1), -y(t)y^2(t-1), -y(t)y^2(t-2), y(t-2), y(t-2)u(t-2), \\
& u(t-2), -y(t)y^2(t-2), -y(t)y^2(t-3), y(t-3), y(t-3)u(t-3), u(t-3), -y(t)y^2(t-3), \\
& -y(t)y^2(t-4)]^{\mathrm{T}},
\end{aligned}$$

the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance, and $\{e(t)\}$ is taken as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$.

First, Apply the RLS-TH algorithm to estimate the parameters of the time-delay rational model. The estimation errors $\tau := \|\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta}_p\|/\|\boldsymbol{\theta}_p\|$ versus t are shown in Figure 1. The parameter estimates and the estimation errors are shown in Table 1.

Next, we utilize the BCRLS-TH algorithm to estimate the parameters of the time-delay rational model. The estimation errors $\tau := \|\hat{\boldsymbol{\theta}}_u - \boldsymbol{\theta}_p\|/\|\boldsymbol{\theta}_p\|$ versus t are shown in Figure 1. The parameter estimates and the estimation errors are shown in Table 2. The rational parameters chosen by using different thresholds are given in Table 3.

Table 1: The RLS-TH algorithm estimates and errors

t	100	200	300	500	1000	True Values
$l_{1,1}$	0.08848	0.06943	0.04820	0.03636	0.00185	0.00000
$l_{1,2}$	-0.03102	-0.00915	-0.00954	-0.01238	-0.00259	0.00000
$l_{1,3}$	-0.00522	-0.00157	0.00097	-0.00164	0.00022	0.00000
$l_{1,4}$	-0.18408	-0.09982	-0.08178	-0.07129	-0.06369	0.00000
$l_{1,5}$	0.00930	0.05052	0.06805	0.08018	0.10145	0.00000
a_1	0.16799	0.19545	0.16866	0.16988	0.18015	0.20000
a_2	0.09244	0.08471	0.09027	0.08531	0.09092	0.10000
a_3	0.83042	0.94614	0.97555	0.98914	1.01213	1.00000
b_2	0.42528	0.48197	0.50607	0.52275	0.55200	1.00000
b_3	0.58081	0.67992	0.66609	0.66855	0.68706	1.00000
$l_{3,1}$	0.01885	-0.00200	0.01311	0.00071	-0.00902	0.00000
$l_{3,2}$	-0.00733	-0.00930	-0.00394	0.00328	-0.00315	0.00000
$l_{3,3}$	-0.02876	-0.02979	-0.01610	-0.01608	0.00062	0.00000
$l_{3,4}$	0.02054	0.08661	0.07739	0.07903	0.09137	0.00000
$l_{3,5}$	-0.15168	-0.12362	-0.10028	-0.08935	-0.08286	0.00000
τ (%)	44.47579	36.89364	35.63763	34.65814	32.83981	

Assume $\epsilon = 0.08$ and $t = 3000$. According to Tables 1-3, we can get that the absolute values of the estimates $\hat{l}_{1,1}$, $\hat{l}_{1,2}$, $\hat{l}_{1,3}$, $\hat{l}_{1,4}$, $\hat{l}_{3,1}$, $\hat{l}_{3,2}$ and $\hat{l}_{3,3}$ by using the RLS-TH algorithm are all smaller than 0.08, then the estimated parameter vector by using the RLS-TH algorithm can be expressed as

$$\hat{\boldsymbol{\theta}}_{LS} = [0, 0, 0, 0, 0.10145, 0.18015, 0.09092, 1.01213, 0.55200, 0.68706, 0, 0, 0, 0.09137, -0.08286]^{\mathrm{T}},$$

the rational model parameter vector estimate can be simplified as

$$\hat{\boldsymbol{\theta}} = [0.10145, 0.18015, 0.09092, 1.01213, 0.55200, 0.68706, 0.09137, -0.08286]^{\mathrm{T}}.$$

On the other hand, the estimated parameter vector by using the BCRLS-TH algorithm can be expressed as

$$\hat{\boldsymbol{\theta}}_u = [0, 0, 0, 0, 0, 0.19957, 0.09970, 0.99331, 0.99471, 0.98899, 0, 0, 0, 0, 0]^{\mathrm{T}},$$

Table 2: The BCRLS-TH algorithm estimates and errors

t	100	200	300	500	1000	True Values
$l_{1,1}$	0.00110	0.02066	0.01726	0.00991	0.00903	0.00000
$l_{1,2}$	0.00402	0.00002	0.00436	0.00280	0.00043	0.00000
$l_{1,3}$	0.00265	0.00694	0.00118	-0.00066	-0.00070	0.00000
$l_{1,4}$	-0.05420	-0.03162	-0.02235	-0.01499	-0.00660	0.00000
$l_{1,5}$	-0.14496	-0.06149	-0.03979	-0.02434	0.00831	0.00000
a_1	0.13838	0.16652	0.18185	0.18866	0.19957	0.20000
a_2	0.04614	0.09432	0.08869	0.08724	0.09970	0.10000
a_3	0.77605	0.87758	0.92000	0.95004	0.99331	1.00000
b_2	0.69149	0.85661	0.89955	0.93012	0.99471	1.00000
b_3	0.71927	0.82511	0.89030	0.92699	0.98899	1.00000
$l_{3,1}$	0.02893	0.00348	-0.00359	-0.00169	0.00101	0.00000
$l_{3,2}$	-0.00144	0.00890	0.00796	0.00392	0.00065	0.00000
$l_{3,3}$	-0.00003	-0.01920	-0.01780	-0.00991	-0.00676	0.00000
$l_{3,4}$	-0.11378	-0.05922	-0.02562	-0.00670	0.02525	0.00000
$l_{3,5}$	-0.05571	-0.03133	-0.02151	-0.01634	-0.00970	0.00000
τ (%)	29.84687	15.93955	10.38526	6.85983	1.95706	

Table 3: The rational parameters chosen by different thresholds

	The RLS-TH estimates	The BCRLS-TH estimates
ϵ	Rational model parameters	Rational model parameters
0.1	$l_{1,5}, a_1, a_3, b_2, b_3$	a_1, a_3, b_2, b_3
0.09	$l_{1,5}, a_1, a_2, a_3, b_2, b_3, l_{3,4}$	a_1, a_2, a_3, b_2, b_3 (True)
0.08	$l_{1,5}, a_1, a_2, a_3, b_2, b_3, l_{3,4}, l_{3,5}$	a_1, a_2, a_3, b_2, b_3 (True)
0.03	$l_{1,4}, l_{1,5}, a_1, a_2, a_3, b_2, b_3, l_{3,4}, l_{3,5}$	a_1, a_2, a_3, b_2, b_3 (True)
0.02	$l_{1,4}, l_{1,5}, a_1, a_2, a_3, b_2, b_3, l_{3,4}, l_{3,5}$	$a_1, a_2, a_3, b_2, b_3, l_{3,4}$

and then the rational model parameter vector estimate can be simplified as

$$\hat{\boldsymbol{\theta}} = [0.19957, 0.09970, 0.99331, 0.99471, 0.98899]^{\mathrm{T}}.$$

Clearly, in the RLS-TH algorithm, some redundant parameters cannot be picked out from $\hat{\boldsymbol{\theta}}_{LS}$ ($l_{1,5}$, $l_{3,4}$ and $l_{3,5}$), while in the BCRLS-TH, all the redundant parameters can be picked out from $\hat{\boldsymbol{\theta}}_u$ and then the unknown time-delay can be obtained based on the structures of $\hat{\boldsymbol{\theta}}_u$ and $\hat{\boldsymbol{\theta}}$. Furthermore, from Table 1, we can conclude that there is no ϵ which can be used to pick out the redundant parameters from $\hat{\boldsymbol{\theta}}_{LS}$ because of the estimates $l_{1,5}$ and a_2. A big ϵ will mistake a_2 for the redundant parameter. On the other hand, a small ϵ will lead to the redundant parameter $l_{1,5}$ not be picked out.

Ultimately, we can draw the following conclusions.

1. Figure 1 shows that the BCRLS-TH algorithm is more effective than the RLS-TH algorithm.
2. Tables 1 and 3 declare that when apply the RLS-TH algorithm to estimate the parameters, there will be no threshold which can be used to pick out all the redundant parameters from $\hat{\boldsymbol{\theta}}_{LS}$.
3. Values in Tables 2 and 3 witness that the threshold can be easily chosen in the BCRLS-TH algorithm, e.g.,the threshold can be chosen in $[0.03, 0.09]$ and the unknown time-delay is 1.

6. Conclusions

A BCRLS-TH algorithm is proposed for time-delay rational models in this paper. By using the redundant rule, the rational model can be transformed into an augmented model which consists of two parts: one is the redundant part and the other is the true part. Then an unbiased parameter estimate can be obtained by

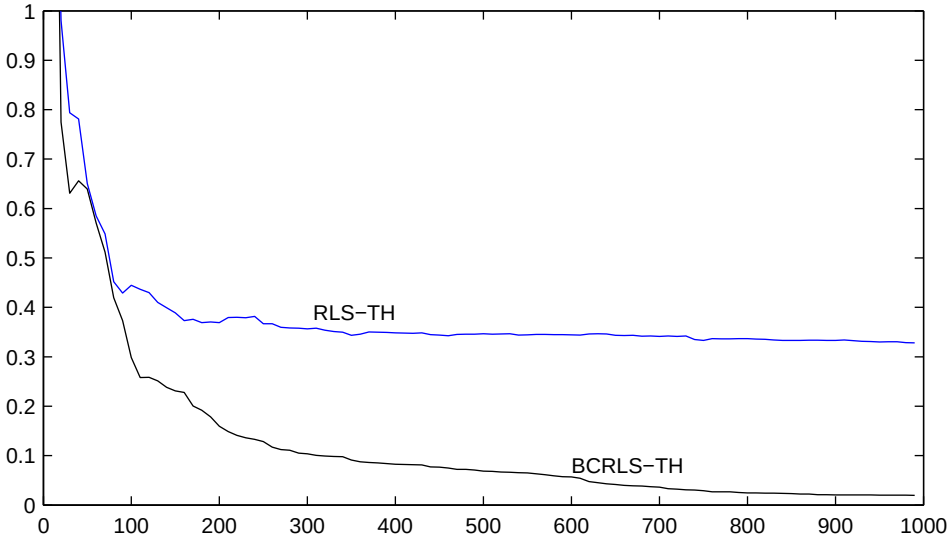

Figure 1: The parameter estimation errors τ versus t

the BCRLS-TH algorithm, and the unknown time-delay can be computed based on the two structures of the augmented model parameter vector and the rational model parameter vector. Interesting extensions of this work can also be pursued in some real rational model estimation [9, 27] where time-delays are likely to occur.

To the best of our knowledge, this is the first on-line algorithm proposed for time-delay rational models. As a new direction, there are some potential topics associated with this work. For example, if the structures of the denominator and numerator polynomials are unknown, how to estimate the parameters and the unknown time-delay? Another topic is can this method be extended to rational models with varying time-delays? These topics will remain as challenging issues in future.

References

[1] S.A. Billings, Q.M. Zhu, Rational model identification using extended least squares algorithm, International Journal of Control 54 (3) (1991) 529-546.

[2] D.I. Kamenski, S.D. Dimitrov, Parameter estimation in differential equations by application of rational functions, Computers & Chemical Engineering 17 (1993) 643-651.

[3] E. Klipp, R. Herwig, A. Kowald, Systems biology in practice: concepts, implementation and application, Weinheim, Germany: Wiley-VCH, (2005).

[4] J.X. Ma, F. Ding, E.F. Yang, Data filtering-based least squares iterative algorithm for Hammerstein nonlinear systems by using the model decomposition, Nonlinear Dynamics 83 (4) (2016) 1895-1908.

[5] M.H. Li, X.M. Liu, F. Ding, Least-squares-based iterative and gradient-based iterative estimation algorithms for bilinear systems, Nonlinear Dynamics 89 (1) (2017) 197-211.

[6] Q.M. Zhu, Y. Wang, D. Zhao, et al., Review of rational (total) nonlinear dynamic system modelling, identification, and control, International Journal of Systems Science 46 (12) (2015) 2122-2133.

[7] Y.W. Mao, F. Ding, Multi-innovation stochastic gradient identification for Hammerstein controlled autoregressive autoregressive systems based on the filtering technique, Nonlinear Dynamics 79 (3) (2015) 1745-1755.

[8] F.Y. Chen, F. Ding, J.H. Li, Maximum likelihood gradient-based iterative estimation algorithm for a class of input nonlinear controlled autoregressive ARMA systems, Nonlinear Dynamics 79 (2) (2015) 927-936.

[9] Q.M. Zhu, An implicit least squares algorithm for nonlinear rational model parameter estimation, Applied Mathematical Modelling 29 (7) (2005) 673-689.

[10] B.Q. Mu, E.W. Bai, W.X. Zheng, et al., A globally consistent nonlinear least squares estimator for identification of nonlinear rational systems, Automatica 77 (2017) 322-335.

[11] Q.M. Zhu, A back propagation algorithm to estimate the parameters of nonlinear dynamic rational models, Applied Mathematical Modelling 27 (3) (2003) 169-187.

[12] Q.M. Zhu, D.L. Yu, D.Y. Zhao, An enhanced linear Kalman filter (EnLKF) algorithm for parameter estimation of nonlinear rational models, International Journal of Systems Science 48 (3) (2017) 451-461.

[13] H.L. Yang, L. Shu, S.M. Zhong, Pinning lag synchronization of complex dynamical networks with known state time-delay and unknown channel time-delay, Nonlinear Dynamics (2017). doi: 10.1007/s11071-017-3552-8

[14] H. Li, Y. Shi, Distributed receding horizon control of large-scale nonlinear systems: Handling communication delays and disturbances, Automatica 50 (2014) 1264-1271.

[15] S. Bedoui, M. Ltaief, K. Abderrahim, New results on discrete time delay systems, International Journal of Automation and Computing 9 (6) (2012) 570-577.

[16] Y. Shi, B. Yu, Robust mixed H_2/H_∞ control of networked control systems with random time delays in both forward and backward communication links, Automatica 47 (4) (2011) 754-760.

[17] H. Li, Y. Shi, Robust H_∞ filtering for nonlinear stochastic systems with uncertainties and random delays modeled by Markov chains, Automatica 48 (1) (2012) 159-166.

[18] W.L. Xiong, X.Q. Yang, B. Huang, et al., Multiple-based linear parameter varying time-delay system identification with missing output data using an Expectation-Maximization algorithm, Industrial & Engineering Chemistry Research 53 (2014) 11074-11083.

[19] Y. Guo, B. Huang, State estimation incorporating infrequent, delayed and integral measurements, Automatica 58 (2015) 32-38.

[20] D. Angelosante, G.B. Giannakis, N.D. Sidiropoulos, Estimating multiple frequency-hopping signal parameters via sparse linear regression, IEEE Transactions on Signal Processing 58 (10) (2010) 5044-5056.

[21] F. Ding, X.H. Wang, L. Mao, et al., Joint state and multi-innovation parameter estimation for time-delay linear systems and its convergence based on the Kalman filtering, Digital Signal Processing 62 (2017) 211-223.

[22] I. Peñarrocha, R. Sanchis, P. Albertos, Estimation in multisensor networked systems with scarce measurements and time varying delays, Systems & Control Letters 61 (4) (2012) 555-562.

[23] Y.J. Zhao, A. Fatehi, B. Huang, Robust estimation of ARX models with time-varying time delays using variational Bayesian approach, IEEE Transactions on Cybernetics (2017). doi: 10.1109/TCYB.2016.2646059

[24] Y.J. Liu, T.Y. Tao, A CS recovery algorithm for model and time delay identification of MISO-FIR systems, Algorithms 8 (2015) 743-753.

[25] J.X. Ma, F. Ding, W.L. Xiong, et al., Combined state and parameter estimation for Hammerstein systems with time-delay using the Kalman filtering, International Journal of Adaptive Control and Signal Processing (2017). doi: 10.1002/acs.2752

[26] J. Chen, J.X. Ma, Y.J. Liu, et al., Identification methods for time-delay systems based on the redundant rules, Signal Processing 137 (2017) 192-198.

[27] D.M. Bates, D.G. Watts, Nonlinear regression analysis and its applications, Hoboken, NJ: John Wiley & Sons, (2007).