

A bias compensation on-line algorithm for dual-rate rational models [☆]

Jing Chen^{*,a}, Quanmin Zhu^b, 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*

Abstract

In dual-rate rational systems, some output data are missing (unmeasurable) to make the traditional recursive least squares parameter estimation algorithms invalid. In order to overcome this difficulty, this paper develops a bias compensation recursive least squares algorithm for estimating the missing outputs and then the model parameters. The algorithm, based on auxiliary model and particle filter, has four integrated key functions, 1) to establish an auxiliary model to estimate unmeasurable outputs, 2) to compensate bias induced by correlated noise, 3) to add a filter to improve estimation accuracy of the unmeasurable outputs, 4) to obtain unbiased parameter estimation. Three examples are selected for simulation demonstrations to give further guarantees on the usefulness of the proposed algorithms. The comparative studies show that the bias compensation recursive least squares is more effective for such systems with dual-rate input and output data.

Key words: Parameter estimation, Recursive least squares, Auxiliary model, Particle filter, Rational model

1. Introduction

Nonlinear systems widely exist in engineering practice. When design robust controllers for nonlinear systems, researchers often assume that the parameters of the nonlinear systems are known in prior, such assumption may not be true in engineering practice. Thus parameter estimation plays an important role in system control. Recently, a number of identification algorithms have been developed. These algorithms can be roughly divided into two categories: One is the off-line algorithms, such as the least square (LS) algorithm [1, 2], the expectation-maximization algorithm and the variational Bayesian algorithm [3, 4]; The other is the on-line algorithms, such as the stochastic gradient algorithm [5, 6], the recursive least squares algorithm (RLS) and the maximum likelihood algorithm [7, 8]. Compared with the off-line algorithms, the on-line algorithms have less computational efforts and can update the parameters with new data, which makes the on-line algorithms more effective for implementation.

The rational model was first proposed in [9], where the rational model was considered as a special kind of nonlinear systems. Unlike the polynomial nonlinear systems which can be simplified as linear in the parameters and nonlinear in the regression terms, the rational model is defined as the ratio of two polynomial expressions [10]. The parameter identification for rational models is more challenging because a rational model is nonlinear in both the parameters and the regression terms. Since the rational model provides a very concise and parsimonious representation for complex nonlinear systems and has excellent extrapolation properties. The rational model identification is becoming a hot spot of present research [11, 12, 13, 14].

The off-line algorithms are the most widely used identification algorithms for rational models. For example, in [15], an error back propagation parameter estimation algorithm for a class of rational models was studied, and the rational model was depicted in a neural network structure. In [16], an implicit LS algorithm for rational models was proposed, where the algorithm is efficient in dealing with the parameter estimation problems associated with nonlinear in the parameters models. In [17], a globally consistent nonlinear LS estimator for identification of a nonlinear rational system was developed, in which the proposed off-line algorithm is the first globally convergent algorithm for the nonlinear rational systems. However, the off-line algorithms usually have heavy computational loads and cannot be used to update the parameters in real time when new data becomes

[☆]This work is 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), yanjunliu.1983@126.com (Yanjun Liu)

available. In order to overcome these shortcomings of off-line algorithms, an enhanced linear Kalman filter algorithm for parameter estimation of nonlinear rational models was proposed in [18], in which the proposed algorithm is an on-line algorithm. Unfortunately, all the identification algorithms in the above work are limited to rational models whose input and output data are sampled at the same rate. When the rational model has dual-rate input and output data, these algorithms may be invalid. To the best of our knowledge, there is no work reported in the literature on identification of dual-rate rational models. Thus the focus of this paper is to develop a bias compensation on-line algorithm for rational models with dual-rate input and output data.

Dual-rate systems which the inputs and outputs are sampled at two different rates often exist in industrial [19]. For example, in polymer reactors, the manipulated variables can be adjusted at relatively fast rate, whereas the composition, density or molecular weight distribution measurements are typically obtained after several minutes of analysis [20]. The lifting and the polynomial transformation techniques are two main tools which are usually applied to deal with the dual-rate system identification [21, 22]. Both of these two methods first turn the original dual-rate system into a system containing all the input data and only the measured output data and then use all the measured data to estimate the unknown parameters. However, the complex structure of the rational model leads to the dual-rate system be impossible simplified as a systems containing all the measured data, thus these two methods cannot be utilized for rational model identification.

The auxiliary model is another effective method for systems with unmeasurable outputs [23]. Here, those unmeasurable outputs were replaced by the outputs of the auxiliary model, and then the unknown parameters were estimated based on the rich input data, the measured output data and the estimated output data. Compared with the lifting technique and the polynomial transformation technique, the auxiliary model method has no restriction to the polynomial of the output and can keep the number of the unknown parameters unchanged [24]. Thus the auxiliary model method is widely used for the dual-rate system identification [25, 26]. Unfortunately, the auxiliary model method only uses an auxiliary model to predict the missing outputs, while the measured outputs are not applied to improve the predicted outputs during the interval of the slow sampled rate, which makes the errors between the predicted outputs and the true missing outputs large. It is well known that the particle filter is an **optimal** tool for nonlinear state space model [27, 28]. Unlike the auxiliary model method, the particle filter uses the given data so far to improve the predicted states. Thus the particle filter method is more effective than the auxiliary model method [29, 30]. Nevertheless, the standard particle filter is only available for a state space model. In this paper, we will extend the particle filter to a non-state-space model.

This paper takes the above described literature into study, and proposes a bias compensation RLS based auxiliary model (BCRLS-AM) algorithm and a BCRLS based particle filter (BCRLS-PF) algorithm for a dual-rate rational model. In the BCRLS-AM algorithm, the unmeasurable outputs are estimated by the auxiliary model, while in the BCRLS-PF algorithm, the unmeasurable outputs are estimated by the particle filter and are improved by the measured outputs during each interval of the slow sampled rate. Thus the BCRLS-PF algorithm is more effective than the BCRLS-AM algorithm.

Briefly, the rest of this paper is organized as follows. Section 2 introduces the rational model and the RLS algorithm. Section 3 develops the BCRLS-AM algorithm. Section 4 studies the BCRLS-PF algorithm. Section 5 provides three examples. Finally, concluding remarks are given in Section 6.

2. The rational model and the RLS algorithm

Consider the following rational model,

$$y(t) = \frac{a(t)}{b(t)} + e(t), \quad (1)$$

where $y(t)$ is the output, $u(t)$ is the input, $e(t)$ a stochastic white noise with zero mean and variance σ , and $a(t)$ and $b(t)$ are expressed as

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

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

$$\boldsymbol{\theta}_a = [a_1, a_2, \dots, a_n]^T, \quad (2)$$

$$\boldsymbol{\theta}_b = [b_1, b_2, \dots, b_m]^\top. \quad (3)$$

Without loss of generality, assume $b_1 = 1$, then the rational model can be simplified as

$$Y(t) = \boldsymbol{\varphi}^\top(t)\boldsymbol{\theta}_a + y(t)[- \boldsymbol{\psi}^\top(t)\boldsymbol{\theta}_b + b_1\psi_1(t)] + b(t)e(t), \quad (4)$$

where $Y(t) = y(t)\psi_1(t)$ and $\psi_1(t)$ is the first element of the vector $\boldsymbol{\psi}(t)$. Rewriting Equation (4) as

$$Y(t) = \boldsymbol{\phi}^\top(t)\boldsymbol{\theta} + v(t), \quad (5)$$

where

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

For the dual-rate system, we assume that the output sampling period is qh ($q \geq 2$ is an integer), while the input sampling period is h . Clearly, the output sampling period is an integer multiple of the input updating period. Then all the input data $\{u(t), t = 0, 1, 2, \dots\}$ and some output data $\{y(tq), t = 1, 2, \dots\}$ are measurable, while the other outputs are unmeasurable.

Replacing t in (5) with tq gets

$$\begin{aligned} Y(tq) &= \boldsymbol{\phi}^\top(tq)\boldsymbol{\theta} + v(tq), \\ \boldsymbol{\phi}(tq) &= [\varphi_1(tq), \varphi_2(tq), \dots, \varphi_n(tq), -y(tq)\psi_2(tq), \dots, -y(tq)\psi_m(tq)]^\top, \\ v(tq) &= b(tq)e(tq). \end{aligned}$$

Collect tq input, output and noise data and define

$$\begin{aligned} \mathbf{Y}(tq) &:= [Y(tq), Y(tq-q), \dots, Y(q)]^\top, \\ \boldsymbol{\Phi}(tq) &:= [\boldsymbol{\phi}(tq), \boldsymbol{\phi}(tq-q), \dots, \boldsymbol{\phi}(q)]^\top, \\ \mathbf{V}(tq) &:= [v(tq), v(tq-q), \dots, v(q)]^\top. \end{aligned}$$

It follows that

$$\mathbf{Y}(tq) = \boldsymbol{\Phi}(tq)\boldsymbol{\theta} + \mathbf{V}(tq). \quad (6)$$

Then the parameters $\boldsymbol{\theta}$ can be estimated by the following LS algorithm,

$$\hat{\boldsymbol{\theta}} = [\boldsymbol{\Phi}^\top(tq)\boldsymbol{\Phi}(tq)]^{-1}\boldsymbol{\Phi}^\top(tq)\mathbf{Y}(tq). \quad (7)$$

Compared with the LS algorithm, the RLS algorithm has less computational efforts, and can update the parameters in real time when new data becomes available. Next, a RLS algorithm will be derived for the rational model,

$$\begin{aligned} \hat{\boldsymbol{\theta}}(tq) &= [\boldsymbol{\Phi}^\top(tq)\boldsymbol{\Phi}(tq)]^{-1}\boldsymbol{\Phi}^\top(tq)\mathbf{Y}(tq) \\ &= \mathbf{P}(tq) \begin{bmatrix} \boldsymbol{\Phi}(tq-q) \\ \boldsymbol{\phi}^\top(tq) \end{bmatrix}^\top \begin{bmatrix} \mathbf{Y}(tq-q) \\ Y(tq) \end{bmatrix} \\ &= \mathbf{P}(tq)[(\mathbf{P}^{-1}(tq) - \boldsymbol{\phi}^\top(tq)\boldsymbol{\phi}(tq))\hat{\boldsymbol{\theta}}(tq-q) + \boldsymbol{\phi}(tq)Y(tq)] \\ &= \hat{\boldsymbol{\theta}}(tq-q) + \mathbf{P}(tq)\boldsymbol{\phi}(tq)[Y(tq) - \boldsymbol{\phi}^\top(tq)\hat{\boldsymbol{\theta}}(tq-q)]. \end{aligned} \quad (8)$$

Thus the RLS algorithm for estimating the parameter vector $\boldsymbol{\theta}$ can be summarized as follows,

$$\begin{aligned} \hat{\boldsymbol{\theta}}(tq) &= \hat{\boldsymbol{\theta}}(tq-q) + \mathbf{P}(tq)\boldsymbol{\phi}(tq)[Y(tq) - \boldsymbol{\phi}^\top(tq)\hat{\boldsymbol{\theta}}(tq-q)], \\ \mathbf{P}(tq) &= [\boldsymbol{\Phi}^\top(tq)\boldsymbol{\Phi}(tq)]^{-1}, \\ \boldsymbol{\Phi}(tq) &= [\boldsymbol{\phi}(tq), \boldsymbol{\phi}(tq-q), \dots, \boldsymbol{\phi}(q)]^\top. \end{aligned}$$

However, the above RLS algorithm cannot be applied to estimate the unknown parameters for the following two reasons: (1) the LS algorithm is a biased estimation algorithm, which leads to the RLS also be a biased estimation algorithm; (2) the information vector $\boldsymbol{\phi}(iq)$ contains unmeasurable outputs $y(iq-l), l = 1, 2, \dots, q-1, q+1, \dots, iq-1$.

3. The BCRLS-AM algorithm

In order to make the RLS algorithm be an unbiased estimation algorithm, taking the conditional expectation on both sides of (7) gets

$$\begin{aligned} \mathbb{E} \left(\hat{\boldsymbol{\theta}}(tq) \right) &= \mathbb{E} \left\{ [\boldsymbol{\Phi}^T(tq)\boldsymbol{\Phi}(tq)]^{-1}\boldsymbol{\Phi}^T(tq)\mathbf{Y}(tq) \right\} \\ &= \mathbb{E} \left\{ [\boldsymbol{\Phi}^T(tq)\boldsymbol{\Phi}(tq)]^{-1}\boldsymbol{\Phi}^T(tq)(\boldsymbol{\Phi}^T(tq)\boldsymbol{\theta} + \mathbf{V}(tq)) \right\} \\ &= \boldsymbol{\theta} + \mathbf{P}(tq)\mathbb{E} \left\{ \boldsymbol{\Phi}^T(tq)\mathbf{V}(tq) \right\}, \end{aligned} \quad (9)$$

where $\boldsymbol{\Phi}^T(Lq)\mathbf{V}(Lq)$ can be expressed as

$$\begin{aligned} \boldsymbol{\Phi}^T(tq)\mathbf{V}(tq) &= [\boldsymbol{\phi}(tq), \boldsymbol{\phi}(tq-q), \dots, \boldsymbol{\phi}(q)] \begin{bmatrix} v(tq) \\ v(tq-q) \\ \vdots \\ v(q) \end{bmatrix} \\ &= \begin{bmatrix} \varphi_1(tq) & \cdots & \varphi_1(q) \\ \varphi_2(tq) & \cdots & \varphi_2(q) \\ \vdots & \ddots & \vdots \\ \varphi_n(tq) & \cdots & \varphi_n(q) \\ -y(tq)\psi_2(tq) & \cdots & -y(q)\psi_2(q) \\ \vdots & \ddots & \vdots \\ -y(tq)\psi_m(tq) & \cdots & -y(q)\psi_m(q) \end{bmatrix} \begin{bmatrix} b(tq)e(tq) \\ b(tq-q)e(tq-q) \\ \vdots \\ b(q)e(q) \end{bmatrix}. \end{aligned}$$

At time $tq-q$, all the input data and only the scarce output data $\{y(tq-q), \dots, y(q)\}$ are measurable. While the output data $y(tq-1), \dots, y(tq-q+1), \dots, y(1)$ are unknown. In order to overcome this difficulty, an auxiliary model is introduced. The unknown output data are replaced with the outputs of the following auxiliary model,

$$\hat{y}(tq-k) = \frac{\hat{a}(tq-k)}{\hat{b}(tq-k)} \hat{\boldsymbol{\theta}}_m(tq-q), \quad k = q-1, \dots, 1,$$

where $\hat{\boldsymbol{\theta}}_m(tq-q)$ is the unbiased parameter estimate vector at $tq-q$ and the subscript means the estimate of $\frac{\hat{a}(tq-k)}{\hat{b}(tq-k)}$ by using $\hat{\boldsymbol{\theta}}_m(tq-q)$. Replacing y in $\boldsymbol{\Phi}^T(tq)\mathbf{V}(tq)$ with \hat{y} gets

$$\begin{aligned} \boldsymbol{\Phi}^T(tq)\mathbf{V}(tq) &\approx \hat{\boldsymbol{\Phi}}^T(tq)\hat{\mathbf{V}}(tq) \\ &= \begin{bmatrix} \hat{\varphi}_1(tq) & \cdots & \hat{\varphi}_1(q) \\ \hat{\varphi}_2(tq) & \cdots & \hat{\varphi}_2(q) \\ \vdots & \ddots & \vdots \\ \hat{\varphi}_n(tq) & \cdots & \hat{\varphi}_n(q) \\ -y(tq)\hat{\psi}_2(tq) & \cdots & -y(q)\hat{\psi}_2(q) \\ \vdots & \ddots & \vdots \\ -y(tq)\hat{\psi}_m(tq) & \cdots & -y(q)\hat{\psi}_m(q) \end{bmatrix} \begin{bmatrix} \hat{b}(tq)e(tq) \\ \hat{b}(tq-q)e(tq-q) \\ \vdots \\ \hat{b}(q)e(q) \end{bmatrix}. \end{aligned}$$

Unfortunately, $\hat{\varphi}_i(xq)$, $\hat{\psi}_i(xq)$ and $\hat{b}(xq)$ are time-varying, one cannot simplify $\sum_{x=1}^t \hat{\varphi}_i(xq)\hat{b}(xq)e(xq)$ and

$\sum_{x=1}^t \hat{\psi}_i(xq)\hat{b}(xq)e^2(xq)$ as

$$\begin{aligned} \mathbb{E} \left[\sum_{x=1}^t \hat{\varphi}_i(xq)\hat{b}(xq)e(xq) \right] &= 0, \\ \mathbb{E} \left[\sum_{x=1}^t \hat{\psi}_j(xq)\hat{b}(xq)e^2(xq) \right] &= \sigma^2 \sum_{x=1}^t \hat{\psi}_j(xq)\hat{b}(xq). \end{aligned}$$

Thus the parameter estimate vector $\hat{\boldsymbol{\theta}}_m(tq - q)$ at time $tq - q$ is applied to estimate the noise $e(tq)$,

$$\hat{e}(tq) = y(tq) - \frac{\hat{a}(tq - k)}{\hat{b}(tq - k)} \Big|_{\hat{\boldsymbol{\theta}}_m(tq - q)},$$

in which $\hat{e}(tq)$ is the estimate of $e(tq)$.

It follows that

$$\begin{aligned} & \boldsymbol{\Phi}^T(tq) \mathbf{V}(tq) \approx \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq) \\ & = \begin{bmatrix} \hat{\varphi}_1(tq) & \cdots & \hat{\varphi}_1(q) \\ \hat{\varphi}_2(tq) & \cdots & \hat{\varphi}_2(q) \\ \vdots & \ddots & \vdots \\ \hat{\varphi}_n(tq) & \cdots & \hat{\varphi}_n(q) \\ -y(tq)\hat{\psi}_2(tq) & \cdots & -y(q)\hat{\psi}_2(q) \\ \vdots & \ddots & \vdots \\ -y(tq)\hat{\psi}_m(tq) & \cdots & -y(q)\hat{\psi}_m(q) \end{bmatrix} \begin{bmatrix} \hat{b}(tq)\hat{e}(tq) \\ \hat{b}(tq - q)\hat{e}(tq - q) \\ \vdots \\ \hat{b}(q)\hat{e}(q) \end{bmatrix}. \end{aligned} \quad (10)$$

Assume that $\hat{\mathbf{P}}(tq)$ is the estimate of $\mathbf{P}(tq)$ and can be defined as

$$\mathbf{P}(tq) \approx \hat{\mathbf{P}}(tq) = [\hat{\boldsymbol{\Phi}}^T(tq) \hat{\boldsymbol{\Phi}}(tq)]^{-1}, \quad (11)$$

$$\hat{\boldsymbol{\Phi}}(tq) = [\hat{\varphi}(tq), \hat{\varphi}(tq - q), \dots, \hat{\varphi}(q)]^T, \quad (12)$$

$$\hat{\varphi}(iq) = [\hat{\varphi}_1(iq), \hat{\varphi}_2(iq), \dots, \hat{\varphi}_n(iq), -y(iq)\hat{\psi}_2(iq), \dots, -y(iq)\hat{\psi}_m(iq)]^T. \quad (13)$$

Then the following theorem can be obtained.

Theorem 1: For the system in (5), when $\hat{\boldsymbol{\theta}}(tq)$ is expressed by Equation (8), and $\hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq)$ and $\hat{\mathbf{P}}(tq)$ are expressed by Equations (10) and (11). Then the estimate $\hat{\boldsymbol{\theta}}_m(tq)$ is an unbiased estimate of the parameter vector $\boldsymbol{\theta}$ and can be expressed as

$$\hat{\boldsymbol{\theta}}_m(tq) = \hat{\boldsymbol{\theta}}(tq) - \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq). \quad (14)$$

Proof: Taking the conditional expectation on both sides of Equation (14) gets

$$\mathbb{E} [\hat{\boldsymbol{\theta}}_m(tq)] = \mathbb{E} [\hat{\boldsymbol{\theta}}(tq) - \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq)],$$

which means

$$\mathbb{E} [\hat{\boldsymbol{\theta}}_m(tq)] = \mathbb{E} [\hat{\boldsymbol{\theta}}(tq)] - \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq). \quad (15)$$

Substituting Equation (9) into Equation (15) gets

$$\mathbb{E} [\hat{\boldsymbol{\theta}}_m(tq)] = \boldsymbol{\theta} + \mathbf{P}(tq) \mathbb{E} \{ \boldsymbol{\Phi}^T(tq) \mathbf{V}(tq) \} - \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq).$$

According to Equations (10) and (11), one can get

$$\mathbb{E} [\hat{\boldsymbol{\theta}}_m(tq)] = \boldsymbol{\theta}. \quad (16)$$

Equation (16) declares that the estimate $\hat{\boldsymbol{\theta}}_m(tq)$ is an unbiased estimate of parameter vector $\boldsymbol{\theta}$. \square

Then the following BCRLS-AM algorithm for estimating the parameter vector is listed as follows:

$$\hat{\boldsymbol{\theta}}_m(tq) = \hat{\boldsymbol{\theta}}(tq) - \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\Phi}}^T(tq) \hat{\mathbf{V}}(tq), \quad (17)$$

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq - q) + \hat{\mathbf{P}}(tq) \hat{\boldsymbol{\phi}}(tq) [Y(tq) - \hat{\boldsymbol{\phi}}^T(tq) \hat{\boldsymbol{\theta}}(tq - q)], \quad (18)$$

$$\hat{\boldsymbol{\theta}}(tq - k) = \hat{\boldsymbol{\theta}}(tq - q), \quad k = q - 1, \dots, 1, \quad (19)$$

$$\hat{\boldsymbol{\theta}}_m(tq - k) = \hat{\boldsymbol{\theta}}_m(tq - q), \quad (20)$$

$$\hat{\mathbf{P}}(tq) = [\hat{\mathbf{\Phi}}^T(tq)\hat{\mathbf{\Phi}}(tq)]^{-1}, \quad (21)$$

$$\hat{\mathbf{\Phi}}(tq) = [\hat{\phi}(tq), \hat{\phi}(tq-q), \dots, \hat{\phi}(q)]^T, \quad (22)$$

$$\hat{\phi}(iq) = [\hat{\varphi}_1(iq), \hat{\varphi}_2(iq), \dots, \hat{\varphi}_n(iq), -y(iq)\hat{\psi}_2(iq), \dots, -y(iq)\hat{\psi}_m(iq)]^T, \quad (23)$$

$$\hat{e}(tq) = y(tq) - \frac{\hat{a}(tq)}{\hat{b}(tq)} \Big|_{\hat{\boldsymbol{\theta}}_m(tq-q)}, \quad (24)$$

$$\hat{y}(tq-k) = \frac{\hat{a}(tq-k)}{\hat{b}(tq-k)} \Big|_{\hat{\boldsymbol{\theta}}_m(tq-q)}. \quad (25)$$

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

1. Let $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}/p_0$ and $\hat{\mathbf{P}}(0) = p_0\mathbf{I}$ with $\mathbf{1}$ being a column vector whose entries are all unity, \mathbf{I} be an identity matrix of appropriate size and $p_0 = 10^6$.
2. Let $t = 1$, $y(-j) = 0$, $u(-j) = 0$, $e(-j) = 0$, $j = 0, 1, 2, \dots, n-1$, and give a small positive number ε .
3. Let $k = q-1$, and collect the input-output data $\{u(tq-q+1), \dots, u(tq), y(tq)\}$.
4. Form $\hat{\phi}(tq-k)$ by (23).
5. Form $\hat{\mathbf{\Phi}}(tq-k)$ by (22).
6. Compute $\hat{y}(tq-k)$ and $\hat{e}(tq)$ by (25) and (24), respectively. Decrease k by 1, if $k \geq 1$, go to step 4; otherwise, go to next step.
7. Compute $\hat{\mathbf{P}}(tq)$ by (21).
8. Update the parameter estimation vector $\hat{\boldsymbol{\theta}}(tq)$ by (18).
9. Compute $\hat{\boldsymbol{\theta}}_m(tq)$ by (17).
10. Compare $\hat{\boldsymbol{\theta}}_m(tq)$ and $\hat{\boldsymbol{\theta}}_m(tq-q)$: if $\|\hat{\boldsymbol{\theta}}_m(tq) - \hat{\boldsymbol{\theta}}_m(tq-q)\| \leq \varepsilon$, then terminate the procedure and obtain the $\hat{\boldsymbol{\theta}}_m(tq)$; otherwise, increase t by 1 and go to step 3.

Remark 1: According to Equations (17) and (18), we can conclude that the parameter estimates $\hat{\boldsymbol{\theta}}(tq)$ by using the RLS-AM algorithm is biased, while the parameter estimates $\hat{\boldsymbol{\theta}}_m(tq)$ by using the BCRLS-AM algorithm is unbiased.

4. The BCRLS-PF algorithm

The particle filter is an optimal tool for nonlinear state space systems with unknown states [32]. In this section, we will extend the particle filter to dual-rate rational models.

4.1. The particle filter

In the particle filter, the missing outputs can be estimated by

$$\hat{y}(tq-k) = \sum_{j=1}^N \bar{\omega}_j \hat{y}_j(tq-k),$$

in which $\bar{\omega}_j$ is the weight associated with the j th particle and $\sum_{j=1}^N \bar{\omega}_j = 1$, $\hat{y}_j(tq-k)$ is the j th particle drawn from the density function $p(y(tq-k)|y(tq), \dots, y(1), u(tq), \dots, u(1), \boldsymbol{\theta})$. In order to estimate $\hat{y}(tq-k)$, one should get the weights $\bar{\omega}_j$ and the particles $\hat{y}_j(tq-k)$, $j = 1, \dots, N$, respectively.

At time $tq-q$, the density function of the missing outputs can be expressed as

$$p(y(tq-k)|u(tq), \dots, u(1), y(tq), \dots, y(1), \hat{\boldsymbol{\theta}}_m(tq-q)) \approx \sum_{j=1}^N \omega_j \delta(y(tq-k) - \hat{y}_j(tq-k)),$$

where $\delta(\cdot)$ is the Dirac delta function and ω_j is the normalized weight associated with the j th particle. The samples $\{\hat{y}_j(tq-k), j = 1, \dots, N\}$ can be obtained from the following importance density $q(\cdot)$,

$$\begin{aligned} q(y(tq-k)|u(tq-k-1), \dots, u(1), y(tq), \dots, y(1), \hat{\boldsymbol{\theta}}_m(tq-q)) = \\ p(y(tq-k)|u(tq-k-1), \dots, u(1), \hat{y}(tq-k-1), \dots, \hat{y}(1), \hat{\boldsymbol{\theta}}_m(tq-q)). \end{aligned} \quad (26)$$

From Equation (1), Equation (26) can be written as

$$q(y(tq-k)|u(tq-k), \dots, u(1), \hat{y}(tq-k-1), \dots, y(tq-q), \dots, \hat{y}(1), \hat{\theta}_m(tq-q)) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(y(tq-k) - \frac{\hat{a}(tq-k)}{\hat{b}(tq-k)} | \hat{\theta}_m(tq-q))^2}{2\sigma^2} \right]. \quad (27)$$

Thus the new particles can be drew from Equation (27). According to [29], the weight of each new particle can be adjusted as follows,

$$\omega_j(tq) = p(y(tq)|\hat{y}_j(tq))\omega_j(tq-1), \quad (28)$$

$$\omega_j(tq-k) = \omega_j(tq-k-1), \quad k = 1, \dots, q-1. \quad (29)$$

However, the density function $p(y(tq)|\hat{y}_j(tq))$ cannot be obtained directly, which leads to the weight not be updated. Define

$$\|y(tq) - \hat{y}_j(tq)\| := \gamma_j(tq),$$

where $\|\cdot\|$ is the 2 norm. Clearly, we can conclude that the smaller the $\gamma_j(tq)$ is, the more important role of the j th particle at time tq in estimating the true output $y(tq)$ plays.

Let

$$\gamma(tq) = \max\{\gamma_j(tq), j = 1, \dots, N\} + 1,$$

then we can get

$$\frac{\gamma_j(tq)}{\gamma(tq)} < 1, \quad j = 1, \dots, N.$$

Kernel density estimation is a non-parametric method to estimate the probability density function of a random variable. There are many kernel functions, the Epanechnikov kernel function is one of them and is optimal in a mean square error sense. In this paper, the Epanechnikov kernel function is introduced to obtain the density function $p(y(tq)|\hat{y}_j(tq))$ [31],

$$p(y(tq)|\hat{y}_j(tq)) = \frac{3}{4} \left[1 - \left(\frac{\gamma_j(tq)}{\gamma(tq)} \right)^2 \right].$$

By Normalizing the density function, the density function $p(y(tq)|\hat{y}_j(tq))$ can be computed as

$$p(y(tq)|\hat{y}_j(tq)) = \frac{\gamma^2(tq) - \gamma_j^2(tq)}{N\gamma^2(tq) - \sum_{j=1}^N \gamma_j^2(tq)}. \quad (30)$$

Remark 2: If the measured data $y(tq)$ is not applied to update the weight of each particle, then the density function is simplified as $p(y(tq)|\hat{y}_j(tq)) = \frac{1}{N}$ which is the same as that by using the auxiliary model method.

Theorem 2: Assume that the density function $p(y(tq)|\hat{y}_j(tq))$ updated by the particle filter is expressed by Equation (30), and the density function $p(y(tq)|\hat{y}_j(tq))$ updated by the auxiliary model method is $1/N$, then the following inequation holds

$$\sum_{j=1}^N \frac{\gamma^2(tq) - \gamma_j^2(tq)}{N\gamma^2(tq) - \sum_{j=1}^N \gamma_j^2(tq)} (y(tq) - \hat{y}_j(tq))^2 \leq \sum_{j=1}^N \frac{1}{N} (y(tq) - \hat{y}_j(tq))^2.$$

Proof: According to Equation (30), one can get

$$\begin{aligned} & \sum_{j=1}^N \frac{\gamma^2(tq) - \gamma_j^2(tq)}{N\gamma^2(tq) - \sum_{j=1}^N \gamma_j^2(tq)} (y(tq) - \hat{y}_j(tq))^2 - \sum_{j=1}^N \frac{1}{N} (y(tq) - \hat{y}_j(tq))^2 \\ &= \sum_{j=1}^N \frac{-(N-1)\gamma_j^2(tq) + \sum_{k=1, k \neq j}^N \gamma_k^2(tq)}{N(N\gamma^2(tq) - \sum_{k=1}^N \gamma_k^2(tq))} \gamma_j^2(tq). \end{aligned}$$

Since $N(N\gamma^2(tq) - \sum_{k=1}^N \gamma_k^2(tq))$ is greater than zero, one only need to consider the following equation

$$\begin{aligned}
& \sum_{j=1}^N \left[-(N-1)\gamma_j^2(tq) + \sum_{k=1, k \neq j}^N \gamma_k^2(tq) \right] \gamma_j^2(tq) \\
&= \left[(1-N)\gamma_1^4(tq) + \gamma_1^2(tq) \sum_{k=2}^N \gamma_k^2(tq) \right] + \dots \\
& \left[(1-N)\gamma_2^4(tq) + \gamma_2^2(tq) \sum_{k=1, k \neq 2}^N \gamma_k^2(tq) \right] + \dots \\
& \left[(1-N)\gamma_j^4(tq) + \gamma_j^2(tq) \sum_{k=1, k \neq j}^N \gamma_k^2(tq) \right] + \dots \\
& \left[(1-N)\gamma_N^4(tq) + \gamma_N^2(tq) \sum_{k=1}^{N-1} \gamma_k^2(tq) \right] \\
&= -\sum_{k=2}^N (\gamma_1^2(tq) - \gamma_k^2(tq))^2 - \sum_{k=3}^N (\gamma_2^2(tq) - \gamma_k^2(tq))^2 - \dots \\
& - \sum_{k=j+1}^N (\gamma_j^2(tq) - \gamma_k^2(tq))^2 - \dots - (\gamma_{N-1}^2(tq) - \gamma_N^2(tq))^2 \leq 0.
\end{aligned}$$

Clearly, the following inequation holds,

$$\sum_{j=1}^N \frac{\gamma^2(tq) - \gamma_j^2(tq)}{N\gamma^2(tq) - \sum_{j=1}^N \gamma_j^2(tq)} (y(tq) - \hat{y}_j(tq))^2 \leq \sum_{j=1}^N \frac{1}{N} (y(tq) - \hat{y}_j(tq))^2. \quad \square$$

Substituting Equation (30) into Equation (28) gets

$$\omega_j(tq) = \frac{\gamma^2(tq) - \gamma_j^2(tq)}{N\gamma^2(tq) - \sum_{j=1}^N \gamma_j^2(tq)} \omega_j(tq-1). \quad (31)$$

Normalizing $\omega_j(tq-k), k = q-1, \dots, 1, 0$ gets

$$\bar{\omega}_j(tq-k) = \frac{\omega_j(tq-i)}{\sum_{j=1}^N \omega_j(tq-k)}. \quad (32)$$

Then the estimate of $y(tq-k)$ can be computed as

$$\hat{y}(tq-k) = \sum_{j=1}^N \bar{\omega}_j(tq-k) \hat{y}_j(tq-k). \quad (33)$$

The degeneracy phenomenon is an inevitable problem in the particle filter. In order to avoid this phenomenon, an effective sample size N_{eff} is introduced [29],

$$N_{eff} = \frac{1}{\sum_{j=1}^N (\bar{\omega}_j)^2}. \quad (34)$$

Given a threshold N_{hold} in prior, $N_{eff} < N_{hold}$ means severe degeneracy. Thus we should use resampling. The basic idea of resampling is to discard particles with small weights and concentrate on those particles with large weights, then assign the weight of each particles as $\bar{\omega}_j = \frac{1}{N}$.

The steps of the particle filter are summarized as follows.

1. Initialization: At time $tq - q$, draw N initial particles $\{\hat{y}_j(0)\}_{j=1}^N$ from the prior density $p(y(0)|\hat{\boldsymbol{\theta}}_m(tq - q))$ and set each particle's weight to $\frac{1}{N}$. Collect the input data $u(1), \dots, u(tq - q)$ and the output data $y(q), \dots, y(tq - q)$. Give a positive number N_{hold} .
2. Let $k = q - 1$.
3. Collect $\{\hat{y}_j(tq - k)\}_{j=1}^N$ from Equation (27).
4. Compute $\omega_j(tq - k)$ by Equation (28) or Equation (29).
5. Normalize $\bar{\omega}_j(tq - k)$ from Equation (32).
6. Compute the missing output $\hat{y}(tq - k)$ by Equation (33).
7. Compute N_{eff} by Equation (34) and compare N_{eff} with N_{hold} , if $N_{eff} < N_{hold}$, reset the weight of each particle with $\bar{\omega}_j(tq - k) = \frac{1}{N}$ and go to next step; otherwise, go to next step.
8. Decrease k by 1, if $k \geq 0$ go back to step 3; otherwise terminate the procedure.

4.2. The identification algorithm

Since the unmeasurable outputs can be estimated by using the particle filter, then we can get the following BCRLS-PF algorithm:

$$\hat{\boldsymbol{\theta}}_m(tq) = \hat{\boldsymbol{\theta}}(tq) - \hat{\mathbf{P}}(tq)\hat{\boldsymbol{\Phi}}^T(tq)\hat{\mathbf{V}}(tq), \quad (35)$$

$$\hat{\boldsymbol{\theta}}(tq) = \hat{\boldsymbol{\theta}}(tq - q) + \hat{\mathbf{P}}(tq)\hat{\boldsymbol{\phi}}(tq)[Y(tq) - \hat{\boldsymbol{\phi}}^T(tq)\hat{\boldsymbol{\theta}}(tq - q)], \quad (36)$$

$$\hat{\boldsymbol{\theta}}(tq - k) = \hat{\boldsymbol{\theta}}(tq - q), \quad k = q - 1, \dots, 1, \quad (37)$$

$$\hat{\boldsymbol{\theta}}_m(tq - k) = \hat{\boldsymbol{\theta}}_m(tq - q), \quad (38)$$

$$\hat{\mathbf{P}}(tq) = [\hat{\boldsymbol{\Phi}}^T(tq)\hat{\boldsymbol{\Phi}}(tq)]^{-1}, \quad (39)$$

$$\hat{\boldsymbol{\Phi}}(tq) = [\hat{\boldsymbol{\phi}}(tq), \hat{\boldsymbol{\phi}}(tq - q), \dots, \hat{\boldsymbol{\phi}}(q)]^T, \quad (40)$$

$$\hat{\boldsymbol{\phi}}(iq) = [\hat{\varphi}_1(iq), \hat{\varphi}_2(iq), \dots, \hat{\varphi}_n(iq), -y(iq)\hat{\psi}_2(iq), \dots, -y(iq)\hat{\psi}_m(iq)]^T, \quad (41)$$

$$\hat{e}(tq - k) = \hat{y}(tq - k) - \frac{\hat{a}(tq - k)}{\hat{b}(tq - k)} \Big|_{\hat{\boldsymbol{\theta}}_m(tq - q)}. \quad (42)$$

The steps of computing the parameter estimate $\hat{\boldsymbol{\theta}}_m(tq)$ by the BCRLS-PF algorithm are listed below.

1. Let $\hat{\boldsymbol{\theta}}(0) = \mathbf{1}/p_0$ and $\hat{\mathbf{P}}(0) = p_0\mathbf{I}$ with $\mathbf{1}$ being a column vector whose entries are all unity, \mathbf{I} be an identity matrix of appropriate size and $p_0 = 10^6$.
2. Let $t = 1$, $y(-j) = 0$, $u(-j) = 0$, $e(-j) = 0$, $j = 0, 1, 2, \dots, n - 1$, and give a small positive number ε .
3. Let $k = q - 1$, and collect the input-output data $\{u(tq - q + 1), \dots, u(tq), y(tq)\}$.
4. Form $\hat{\boldsymbol{\phi}}^T(tq - k)$ by (41).
5. Form $\hat{\boldsymbol{\Phi}}(tq - k)$ by (40).
6. Compute $\hat{y}(tq - k)$ by the particle filter.
7. Compute $\hat{e}(tq - k)$ by (42). Decrease k by 1, if $k \geq 1$, go to step 4; otherwise, go to next step.
8. Compute $\hat{\mathbf{P}}(tq)$ by (39).
9. Update the parameter estimation vector $\hat{\boldsymbol{\theta}}(tq)$ by (36).
10. Compute $\hat{\boldsymbol{\theta}}_m(tq)$ by (35).
11. Compare $\hat{\boldsymbol{\theta}}_m(tq)$ and $\hat{\boldsymbol{\theta}}_m(tq - q)$: if $\|\hat{\boldsymbol{\theta}}_m(tq) - \hat{\boldsymbol{\theta}}_m(tq - q)\| \leq \varepsilon$, then terminate the procedure and obtain the $\hat{\boldsymbol{\theta}}_m(tq)$; otherwise, increase t by 1 and go to step 3.

Remark 3: It is observed from Theorem 2 that the output estimates by using the particle filter are more accurate than those by using the auxiliary model method, which means that the BCRLS-PF algorithm is more effective than the BCRLS-AM algorithm.

5. Examples

5.1. Example 1

Consider a dual-rate rational model proposed in [9] and assume $q = 2$,

$$y(t) = \frac{0.2y(t-1) + 0.1y(t-1)u(t-1) + u(t-1)}{1 + y^2(t-1) + y^2(t-2)} + \frac{0.8e(t-1) + e(t)}{1 + y^2(t-1) + y^2(t-2)}.$$

The rational model can be turned into the following model,

$$y(t) = 0.2y(t-1) + 0.1y(t-1)u(t-1) + u(t-1) - y(t)y^2(t-1) - y(t)y^2(t-2) + 0.8e(t-1) + e(t),$$

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 $\sigma^2 = 0.10^2$. The parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\phi}(t)$ can be expressed as

$$\begin{aligned} \boldsymbol{\theta} &= [\boldsymbol{\theta}_a^T, \boldsymbol{\theta}_b^T]^T = [a_1, a_2, a_3, b_2, b_3, b_4]^T = [0.2, 0.1, 1, 1, 1, 0.8]^T, \\ \boldsymbol{\phi}(t) &= [\varphi_1(t), \varphi_2(t), \varphi_3(t), -y(t)\psi_2(t), -y(t)\psi_3(t), -y(t)\psi_4(t)]^T \\ &= [y(t-1), y(t-1)u(t-1), u(t-1), -y(t)y^2(t-1), -y(t)y^2(t-2), e(t-1)]^T. \end{aligned}$$

Apply the RLS-AM and the BCRLS-AM algorithms to estimate the parameters of the dual-rate rational model. The estimation errors $\tau := \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t are shown in Figure 1. The parameter estimates and the estimation errors are shown in Tables 1 and 2, respectively. The estimated outputs, the true outputs and their errors are shown in Figure 2 (samples 800-900, BCRLS-AM).

Next, we utilize the BCRLS-PF algorithm to estimate the parameters of the dual-rate rational model. The estimation errors $\tau := \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\|/\|\boldsymbol{\theta}\|$ versus t are shown in Figure 3. The parameter estimates and the estimation errors are shown in Table 3. The estimated outputs, the true outputs and their errors are shown in Figure 4 (samples 800-900).

Table 1: The RLS-AM algorithm estimates and errors

t	100	200	300	500	1000	True Values
a_1	0.14913	0.15563	0.15636	0.13581	0.12754	0.20000
a_2	0.09705	0.09182	0.09369	0.08587	0.08254	0.10000
a_3	0.93340	0.90424	0.89770	0.90463	0.91546	1.00000
b_2	0.91414	0.85764	0.85928	0.87741	0.88076	1.00000
b_3	0.96020	0.90312	0.90746	0.91337	0.92589	1.00000
b_4	0.41447	0.30767	0.37862	0.46169	0.53361	0.80000
τ (%)	21.12177	27.70588	24.32500	20.18904	16.73784	

Table 2: The BCRLS-AM algorithm estimates and errors

t	100	200	300	500	1000	True Values
a_1	0.25349	0.24647	0.23901	0.21536	0.20261	0.20000
a_2	0.10504	0.09381	0.09441	0.08249	0.07914	0.10000
a_3	0.98096	0.94875	0.94225	0.95113	0.96374	1.00000
b_2	1.04686	0.97669	0.97027	0.97959	0.97937	1.00000
b_3	1.05514	0.98822	0.99090	1.00025	1.01154	1.00000
b_4	0.48525	0.38108	0.54288	0.67150	0.79267	0.80000
τ (%)	17.07245	22.14762	13.96512	7.33624	2.53380	

Table 3: The BCRLS-PF algorithm estimates and errors

t	100	200	300	500	1000	True Values
a_1	0.22417	0.23923	0.22300	0.22108	0.21968	0.20000
a_2	0.09526	0.09860	0.09434	0.08776	0.08421	0.10000
a_3	0.98862	0.99921	0.95220	0.97885	0.99325	1.00000
b_2	1.00549	1.03026	0.97506	0.99805	0.98860	1.00000
b_3	1.00977	1.03355	0.91320	0.96229	1.00540	1.00000
b_4	0.24448	0.32093	0.35022	0.66103	0.80504	0.80000
τ (%)	28.95953	25.13350	24.04286	7.68273	1.53259	

From this Example, we can get the following finds.

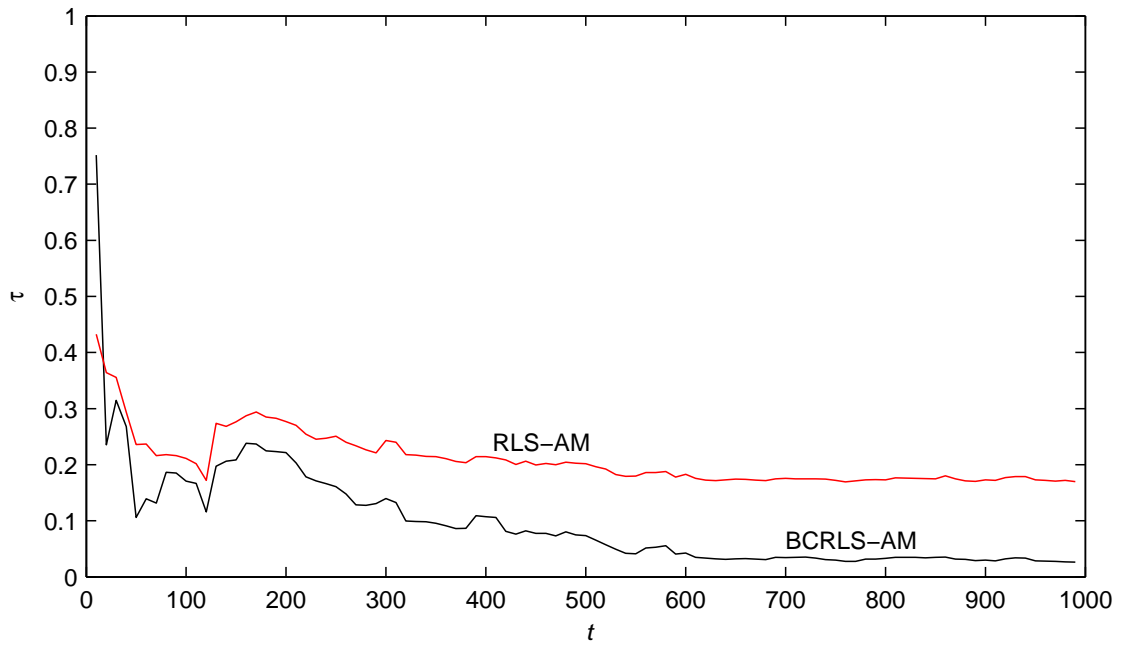


Figure 1: The parameter estimation errors τ versus t

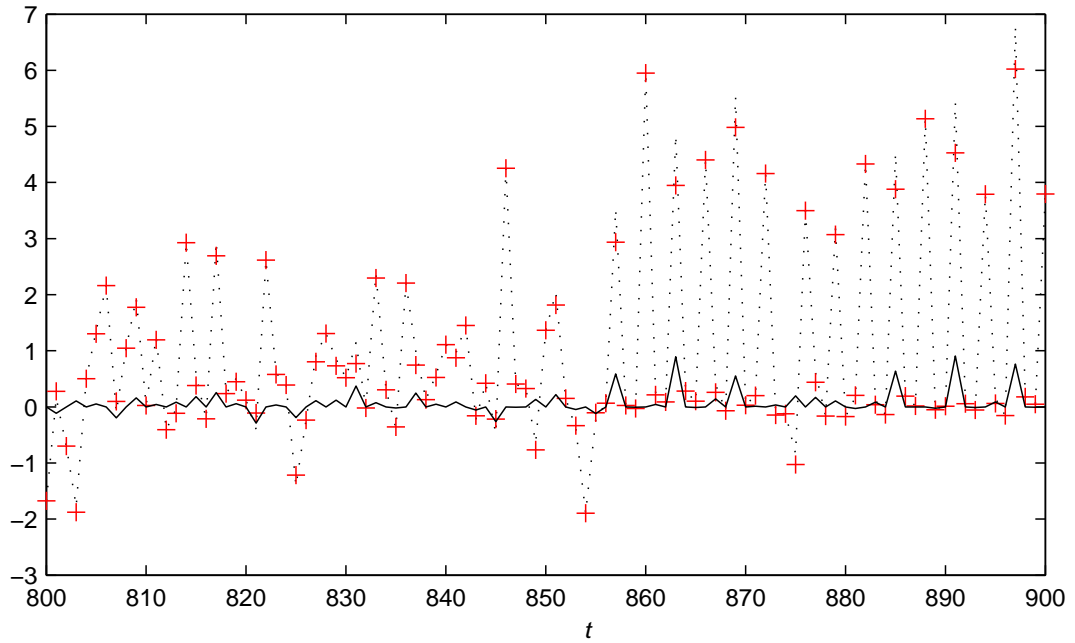


Figure 2: Dotted line : the true outputs; +: the estimated outputs; black line: the output estimation errors (BCRLS-AM)

1. It is observed from Figures 1 and 3 that all the parameter estimation errors of these three algorithms decay when t is increased.
2. Figures 2 and 4 declare that the outputs estimated by using the BCRLS-PF algorithm are more accurate than those by using the BCRLS-AM algorithm.
3. Values in Tables 1-3 witness that the BCRLS-PF algorithm has the smallest parameter estimation errors

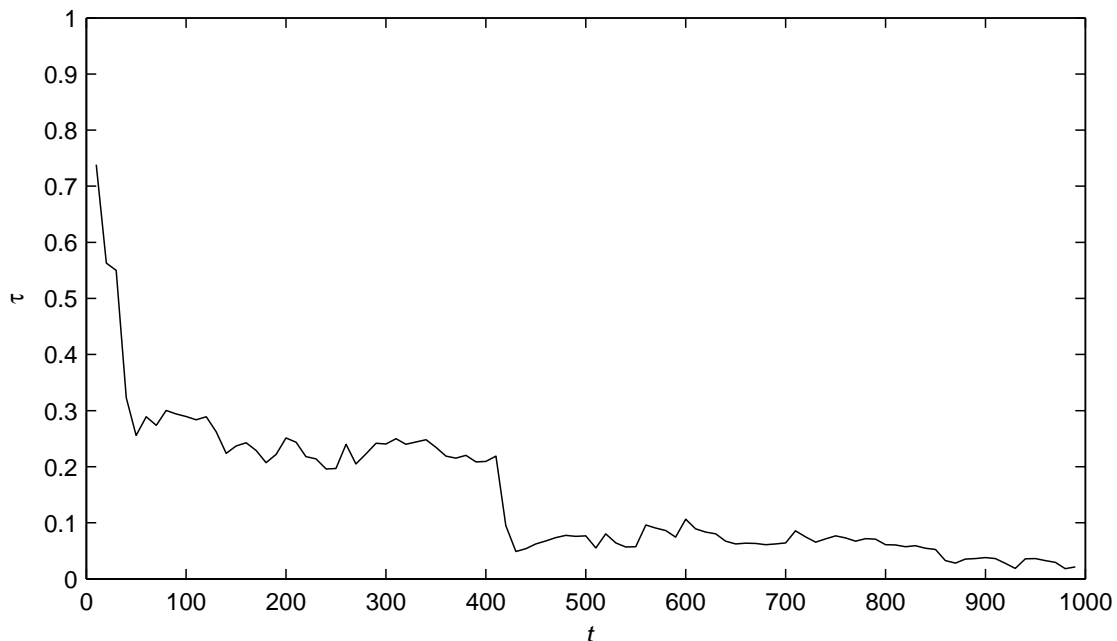


Figure 3: The parameter estimation errors τ versus t (BCRLS-PF)

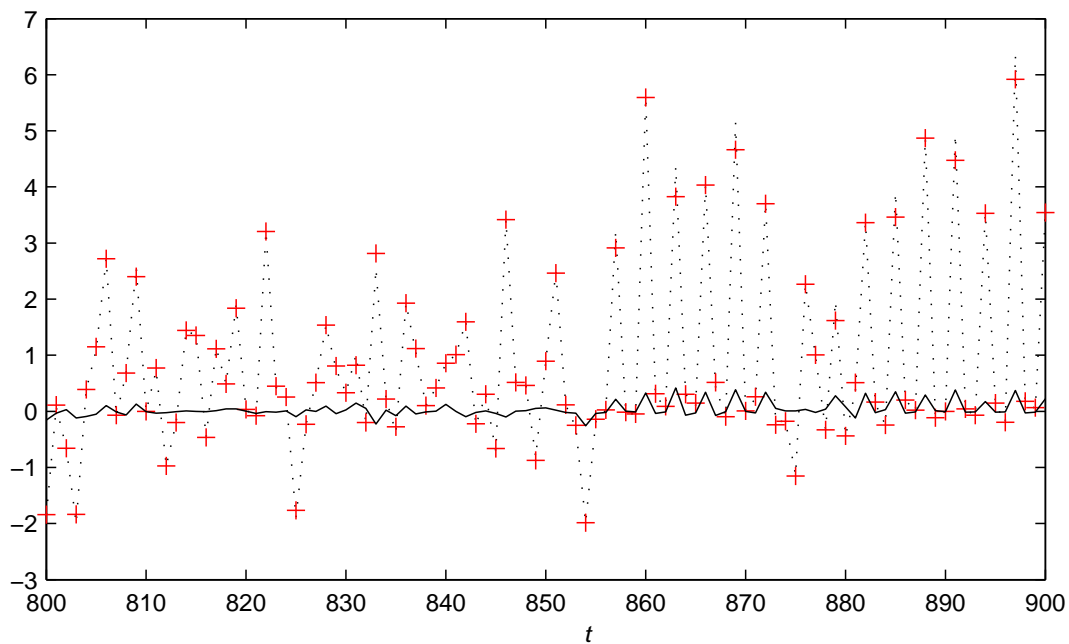


Figure 4: Dotted line : the true outputs; +: the estimated outputs; black line: the output estimation errors (BCRLS-PF)

5.2. Example 2

In this example, a chemical model which is used to describe propylene catalytic oxidation is written as follows [16],

$$y(t) = \frac{k_1 C_p(t)}{1 + k_2 \frac{C_p(t)}{C_o^{0.5}(t)}} + \frac{v(t)}{1 + k_2 \frac{C_p(t)}{C_o^{0.5}(t)}}$$

where $C_o(t)$ and $C_p(t)$ are the oxygen and propylene concentrations at time instant t respectively, and can be seen as the inputs. $y(t)$ is the rate of disappearance of propylene, and can be seen as the output. $k_1 = 7.28 \times 10^{-4}$,

and $k_2 = 0.231$.

In simulation, we manually impose $q = 2$. The inputs $C_o(t)$ and $C_p(t)$ are taken as persistent excitation signal sequences with zero mean and unit variance, $\{v(t)\}$ is taken as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$. Apply the RLS-AM, the BCRLS-AM and the BCRLS-PF algorithms to estimate the parameters of the dual-rate rational model. The estimation errors $\tau := \|\hat{\theta} - \theta\|/\|\theta\|$ versus t by using these three algorithms are shown in Figures 5 and 6. The parameter estimates and the estimation errors are shown in Table 4. From Figure 5, we can conclude that the BCRLS-AM algorithm is better than the RLS-AM algorithm, and from Table 4, we can get that the BCRLS-PF algorithm is the most effective algorithm among these three algorithms.

Table 4: The parameter estimates and errors

Algorithms	t	100	200	300	500	1000	True Values
RLS-AM	k_1	0.00090	-0.00175	-0.00221	-0.00096	-0.00008	0.00073
	k_2	0.22563	0.23248	0.22909	0.22673	0.22332	0.23100
	τ (%)	2.32408	1.25193	1.51540	1.98599	3.34285	
BCRLS-AM	k_1	0.00091	-0.00178	-0.00230	-0.00099	-0.00008	0.00073
	k_2	0.23895	0.22752	0.23456	0.23309	0.23249	0.23100
	τ (%)	3.44428	1.85894	2.02276	1.17098	0.73332	
BCRLS-PF	k_1	0.00069	0.00046	0.00043	0.00056	0.00064	0.00073
	k_2	0.22923	0.22716	0.23077	0.22838	0.23018	0.23100
	τ (%)	0.76744	1.66781	0.16397	1.13639	0.35489	

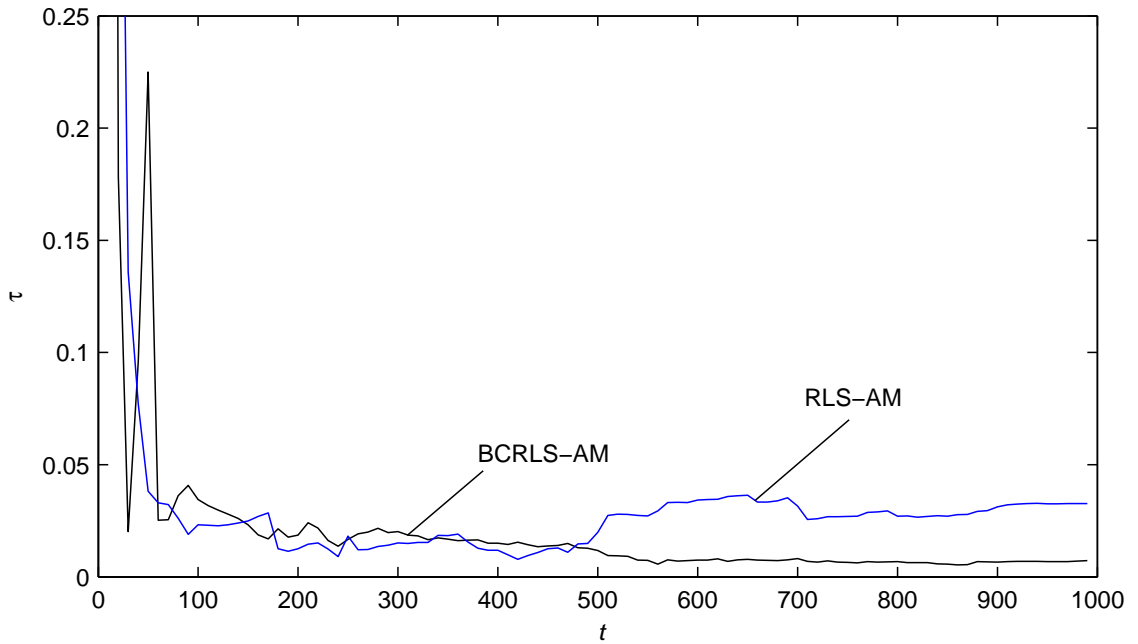


Figure 5: The parameter estimation errors τ versus t

5.3. Example 3

The Michaelis-Menten model which is expressed for enzyme kinetics is considered as follows [12],

$$y(t) = \frac{\beta}{1 + \frac{\alpha}{u(t)}} + \frac{v(t)}{1 + \frac{\alpha}{u(t)}},$$

where $y(t)$ is the initial velocity of an enzymatic reaction, $u(t)$ is the substrate concentration. There are only 12 input-output data $\{u(t), y(t)\}, t = 1, \dots, 12$ in [12], thus the on-line algorithms proposed in this paper are

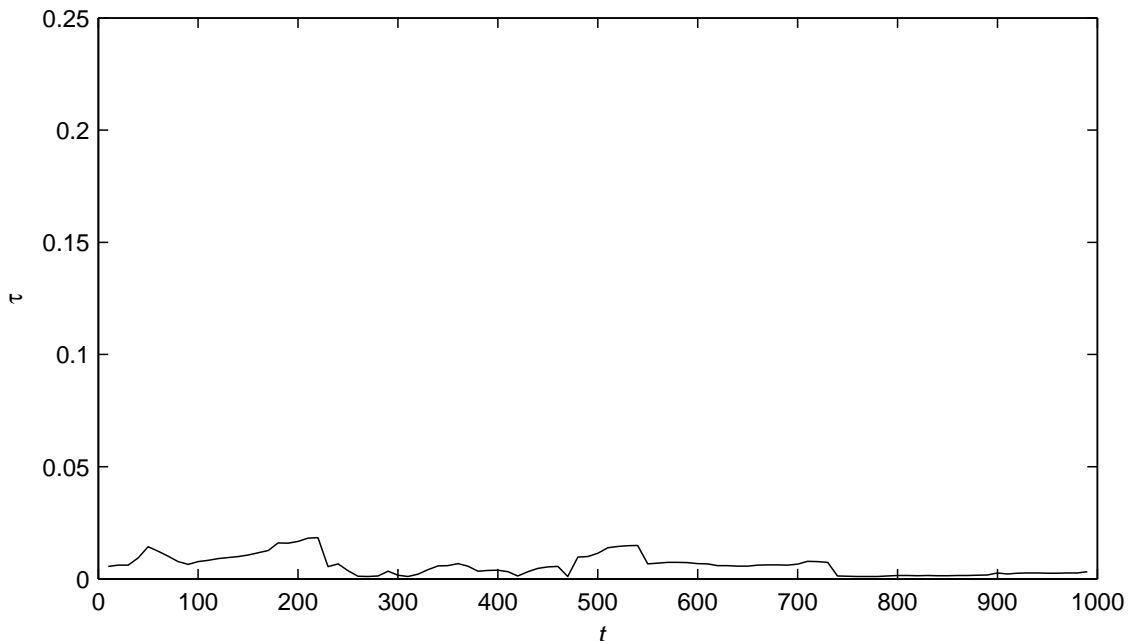


Figure 6: The parameter estimation errors τ versus t (BCRLS-PF)

invalid. In this example, we assign the parameter estimates (the CLS-GN method) proposed in [17] to the unknown parameters: $\alpha = 0.0641$ and $\beta = 212.6837$, and assume $q = 2$.

In simulation, the input $\{u(t)\}$ is taken as a persistent excitation signal sequence with zero mean and unit variance. For a large β , in order to limit the noise to signal ratio to a reasonable level, we manually impose a white noise sequence $\{v(t)\}$ with zero mean and variance $\sigma^2 = 5^2$. Apply the RLS-AM, the BCRLS-AM and the BCRLS-PF algorithms to estimate the parameters of the dual-rate rational model. The comparisons of these three algorithms are shown in Table 5 ($t=1000$). The output estimates by using these three algorithms and the true output data are displayed in Figure 7 (samples 980-1000). From Table 5 and Figure 7, we can conclude that the BCRLS-PF algorithm is the most effective algorithm.

Table 5: The parameter estimates and errors

Algorithms	The parameter estimates	τ (%)
RLS-AM	$\alpha = 0.06386, \beta = 208.61513$	1.91297
BCRLS-AM	$\alpha = 0.06383, \beta = 209.01998$	1.72261
BCRLS-PF	$\alpha = 0.06381, \beta = 210.23803$	1.14991

6. Conclusions

In research methodology, this study presents an example to integrate several across-boundary technical techniques to accommodate some challenging academic issues linked with wide range of applications. To the best of our knowledge, the BCRLS algorithm is the first on-line algorithm proposed for dual-rate rational models, and it is believed that this study will bring forward a new direction in using on-line algorithm to identify rational models with time-delay, random missing observations and colored noise.

In technique development/novelty, a BCRLS-AM algorithm is proposed for dual-rate rational models in this paper, in which the unmeasurable outputs are estimated by an auxiliary model. Based on the rich input data, the measurable output data and the estimated output data, the unknown parameters can be estimated by the BCRLS-AM algorithm. In order to increase the estimation accuracy of the BCRLS-AM algorithm, a BCRLS-PF algorithm is developed. Furthermore, these two algorithms can also be extended to rational models with random missing outputs.

In applications, the step by step procedure plus simulation demonstrations — a format of computational experiments provides potential users a user-friendly framework for their ad hoc demands.

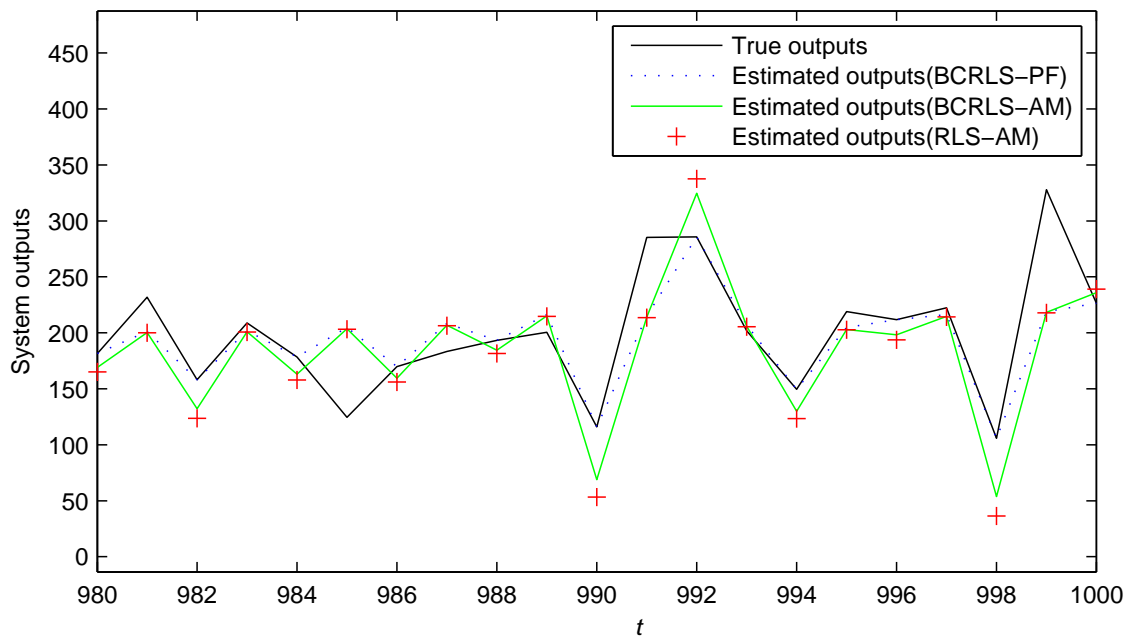


Figure 7: The true outputs and the estimated outputs

In further research, there are still some interesting topics not discussed in this paper. For example, if the denominator and numerator polynomial terms contain noise terms, how to eliminate the bias in the RLS algorithm. Another topic is how to prove the convergence properties of the BCRLS-PF algorithm. These topics will remain as the challenging issues in the future.

References

- [1] H. Sarmadi, A. Karamodin, A. Entezami, A new iterative model updating technique based on least squares minimal residual method using measured modal data, *Applied Mathematical Modelling* 40 (23-24) (2016) 10323-10341.
- [2] F. Ding, F.F. Wang, L. Xu, et al., Decomposition based least squares iterative identification algorithm for multivariate pseudo-linear ARMA systems using the data filtering, *Journal of the Franklin Institute* 354 (3) (2017) 1321-1339.
- [3] F.Y. Chen, F. Ding, The filtering based maximum likelihood recursive least squares estimation for multiple-input single-output systems, *Applied Mathematical Modelling* 40 (3) (2016) 2106-2118.
- [4] Y.J. Lu, B. Huang, S. Khatibisepehr, A variational Bayesian approach to robust identification of switched ARX models, *IEEE Transactions on Cybernetics* 46 (12) (2016) 3195-3208.
- [5] D.Q. Wang, L. Mao, F. Ding, Recasted models based hierarchical extended stochastic gradient method for MIMO nonlinear systems, *IET Control Theory and Applications* 11 (4) (2017) 476-485.
- [6] L. Xu, F. Ding, The parameter estimation algorithms for dynamical response signals based on the multi-innovation theory and the hierarchical principle, *IET Signal Processing* 11 (2) (2017) 228-237.
- [7] D.Q. Wang, Y.P. Gao, Recursive maximum likelihood identification method for a multivariable controlled autoregressive moving average system, *IMA Journal of Mathematical Control and Information* 33 (4) (2016) 1015-1031.
- [8] D.Q. Wang, Z. Zhang, J.Y. Yuan, Maximum likelihood estimation method for dual-rate Hammerstein systems, *International Journal of Control, Automation and Systems* 15 (2) (2017) 698-705.
- [9] S.A. Billings, Q.M. Zhu, Rational model identification using extended least squares algorithm, *International Journal of Control* 54 (3) (1991) 529-546.

- [10] Q.M. Zhu, Z. Ma, K. Warwick, Neural network enhanced generalised minimum variance self-tuning controller for nonlinear discrete time systems, *IEE Proceedings Control Theory and Applications* 146 (4) (1999) 319-326.
- [11] E. Klipp, R. Herwig, A. Kowald, *Systems biology in practice: concepts, implementation and application*, Weinheim, Germany: Wiley-VCH, (2005).
- [12] D.M. Bates, D.G. Watts, *Nonlinear regression analysis and its applications*, Hoboken, NJ: John Wiley & Sons, (2007).
- [13] S.A. Billings, K.Z. Mao, Structure detection for nonlinear rational models using genetic algorithms, *International Journal of Systems Science* 29 (3) (1998) 223-231.
- [14] Q.M. Zhu, Y.J. Wang, D.Y. Zhao, et al., Review of rational (total) nonlinear dynamic system modelling, identification and control, *International Journal of Systems Science* 46 (12) (2015) 2122-2133.
- [15] Q.M. Zhu, A back propagation algorithm to estimate the parameters of nonlinear dynamic rational models, *Applied Mathematical Modelling* 27 (3) (2003) 169-187.
- [16] Q.M. Zhu, An implicit least squares algorithm for nonlinear rational model parameter estimation, *Applied Mathematical Modelling* 29 (7) (2005) 673-689.
- [17] 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.
- [18] 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.
- [19] D.Q. Wang, H.B. Liu, F. Ding, Highly efficient identification methods for dual-rate Hammerstein systems, *IEEE Transactions on Control Systems Technology* 23 (5) (2015) 1952-1960.
- [20] Y.J. Liu, F. Ding, Y. Shi, An efficient hierarchical identification method for general dual-rate sampled-data systems, *Automatica* 50 (3) (2014) 962-970.
- [21] D. Li, S. Shah, T. Chen, Analysis of dual-rate inferential control systems, *Automatica* 38 (6) (2002) 1053-1059.
- [22] J. Ding, F. Ding, X.P. Liu, et al., Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data, *IEEE Transactions on Automatic Control* 56 (11) (2011) 2677-2683.
- [23] F. Ding, F.F. Wang, L. Xu, et al., Parameter estimation for pseudo-linear systems using the auxiliary model and the decomposition technique, *IET Control Theory and Applications* 11 (3) (2017) 390-400.
- [24] J. Chen, Several gradient parameter estimation algorithms for dual-rate sampled systems, *Journal of the Franklin Institute* 351 (1) (2014) 543-554.
- [25] J. Chen, R. Ding, An auxiliary-model-based stochastic gradient algorithm for dual-rate sampled-data Box-Jenkins systems, *Circuits, Systems and Signal Processing* 32 (5) (2013) 2475-2485.
- [26] F. Ding, X.M. Liu, Y. Gu, An auxiliary model based least squares algorithm for a dual-rate state space system with time-delay using the data filtering, *Journal of the Franklin Institute* 353 (2) (2016) 398-408.
- [27] T. Yang, R.S. Laugesen, P.G. Mehta, et al., Multivariable feedback particle filter, *Automatica* 71 (2016) 10-23.
- [28] E. Ozkan, V. Smidl, S. Saha, et al., Marginalized adaptive particle filtering for nonlinear models with unknown time-varying noise parameters, *Automatica* 49 (6) (2013) 1566-1575.
- [29] J. Deng, B. Huang, Identification of nonlinear parameter varying systems with missing output data, *AIChE Journal* 58 (2012) 3454-3467.
- [30] D. Simon, *Optimal State Estimation: Kalman, H_∞ , and Nonlinear Approaches*, John Wiley & Sons, Inc., New Jersey, (2006).
- [31] L. Qi, R.J. Racine, *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, (2007).
- [32] L.E. Calvet, V. Czellar, Accurate Methods for Approximate Bayesian Computation Filtering, *Journal of Financial Econometrics*, 13 (4) (2015) 798-838.