Bias compensation-based parameter and state estimation for a class of time-delay non-linear state-space models

Ya Gu[1] ✉, *Quanmin Zhu*[2], *Hassan Nouri*[2]
[1]*School of Electrical Engineering and Automation, Changshu Institute of Technology, Changshu 215500, Jiangsu, People's Republic of China*
[2]*Department of Engineering Design and Mathematics, University of the West of England, Bristol BS16 1QY, UK*
✉ *E-mail: guya927819@163.com*

Abstract: This study presents, based on bias compensation, an integrated parameter and state estimation algorithm for a class of time-delay non-linear systems which are described by canonical observable state-space model. In technical development, the state-space system model is transformed into an input–output representation/realisation by eliminating the state variables, which is accordingly used as a feasible identification model. With such an input–output structure, directly data measurable to accommodate the estimation bias, an augmented least-squares algorithm (by adding the bias correction terms into the estimates) is proposed for estimating the parameters and states interactively. Regarding the estimator properties, the proposed algorithm is proved unbiased. The simulation results show that the proposed algorithm has good performance in estimating the parameters of state-space systems.

1 Introduction

Systems can be divided into linear systems, bilinear systems and non-linear systems. Linear system identification methods have been well developed [1–3]. Recently, Zhang *et al.* studied the parameter and state estimation for bilinear stochastic systems [4–6]. Non-linearity and time delay have been two critical and challenging factors in modelling and identification of dynamic systems. In process industry, almost all processes have certain degree of non-linearity in conjunction with dynamics, which bring challenge into the modelling and design of the corresponding control systems [7–9]. For example, as the amplifier transistor magnifying glass, once the components are out of linear working range, the non-linear saturation is unavoidable, which induces windup problem in the following control system design [10, 11]. To cope with such non-linear effects in designing control systems, all parameters of the non-linear system models should be available in advance. The importance of the non-linear models is not only for control system design, also for wide range of applications in prediction, decision making, signal processing and so on. Data-driven model identification has been almost always the first step in the system analysis and design. Therefore, the research for non-linear system identification has been a hot topic over the past decades [12–15].

Time-delay systems often exist in engineering practice [16, 17]. For example, in the communication network, due to the network congestion, the signals that are transmitted over a communication channel often have time delays. In chemical processes, some variables such as chemical component concentrations are often measured through laboratory analysis, which will introduce time delays. Recently, there are various analysis algorithms for time-delay systems [18–20]. It is obvious that models of the systems are the foundation of system analysis and control [21–23].

Identification of the state-space systems has received considerable attention in the past few decades [24–26]. There are various identification methods, representatively such as the subspace identification algorithm, the recursive least-squares algorithm, and the Kalman filter algorithm [27]. Most existing works assume that the system is linear and that the system has no time delay. However, it should be noted that the most of the practical systems are subject to non-linear and may have time delays from hardware limitations, relationships between variables, to transfer delay. Therefore, how to identify a fidelity model for a time-delay system has attracted widespread attention [28, 29].

Parameter estimation, state estimation and filtering are important in the area of control. Recently, Zou *et al.* presented the moving horizon estimation algorithms for networked time-delay systems under round-robin protocol [30] and for networked non-linear systems with random access protocol [31], considering unknown inputs under dynamic quantisation effects [32]. In practice, there is a need for estimation of a class of non-linear time-delay state-space model with coloured noise. Such studies are meaningful and significant for applications and academic research. On the basis of the work in [33], this study proposes a solution by using model conversion and bias compensation least-squares principle. The main contributions of the paper are as follows.

- The state-space model is the most popular description of dynamic systems, particularly for control system design. However, its identification and estimation are difficult because of the state variables cannot always be measurable. This study converts the state-space model into its input–output expression including coloured noise and time delay, which is feasible for indentification/estimation as input and output data are always measurable. It should be noted such a model structure conversion has nothing in change of original model's properties such as stability and observability.
- This study presents a bias compensation-based identification algorithm for jointly estimating the system parameters and states based on the bias compensation, which effectively accommodates non-linearity, time-delay dynamics, and coloured noise corruption. This work generalises the linear model-based results to the time-delay case.
- In assurance of using the proposed identification procedure, this study analyses the convergence of the algorithm using stochastic process theory.

For the rest of the study, Section 2 demonstrates the problem formulation about the state-space system with time delay and derives the identification model. Sections 3 and 4 develop a bias compensation-based parameter and state estimation algorithm. Section 5 provides two illustrative examples to show that the proposed algorithm is effective. Finally, Section 6 offers some concluding remarks.

2 System descriptive model and identification model

Descriptive model: This work expands the linear model in [33] to the following descriptive model for a class of non-linear state-space systems with time delay:

$$x(t+1) = Ax(t) + Bx(t-1) + g\bar{u}(t), \tag{1}$$

$$\bar{u}(t) = f(u(t)), \tag{2}$$

$$y(t) = cx(t) + w(t), \tag{3}$$

$$w(t) = \sum_{s=1}^{n_d} d_s v(t-s) + v(t).$$

$$A := \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_1 \end{bmatrix} \in \mathbb{R}^{n\times n}, \tag{4}$$

$$B := [b_1^{\mathrm{T}}, b_2^{\mathrm{T}}, \ldots, b_n^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{n\times n},\ b_i \in \mathbb{R}^{1\times n},$$
$$f := [f_1, f_2, \ldots, f_n]^{\mathrm{T}} \in \mathbb{R}^n,$$
$$c := [1, 0, \ldots, 0] \in \mathbb{R}^{1\times n},$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}$ is the system input, $y(t) \in \mathbb{R}$ is the system output, and $v(t) \in \mathbb{R}$ is a random noise with zero mean, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times n}$, $g \in \mathbb{R}^n$ and $c \in \mathbb{R}^{1\times n}$ are the system parameter matrices/vectors. Assume that (c, A) is observable, and $u(t) = 0$, $y(t) = 0$ and $v(t) = 0$ for $t \leqslant 0$.

The external disturbance $w(t)$ can be fitted by a moving average process, an autoregressive process or an autoregressive moving average process. Without loss of generality, we consider $w(t)$ as a moving average noise process:

$$w(t) := (1 + w_1 z^{-1} + w_2 z^{-2} + \cdots + w_{n_d} z^{-n_d}) v(t). \tag{5}$$

Since $w(t)$ is a coloured noise, the parameter estimation using conventional least-squares algorithm is biased. Here, the principle of deviation compensation is introduced to study the identification problem of such systems. The basic idea is to add a correction term on the basis of the least-squares estimate to obtain the unbiased estimate. The non-linear block is a non-linear function

$$\bar{u}(t) = f(u(t)) = \sum_{l=1}^{m} p_l(t) f_l(u(t)) = f(u(t))p, \tag{6}$$

where basis $f := (f_1, f_2, \ldots, f_m)$ are the known functions of the system input $u(t)$, vector $p := (p_1, p_2, \ldots, p_m)$ are the unknown parameters associated with the non-linear functions and m is the number of the parameters of the non-linear block.

The system in (1)–(4) is an observability canonical form, and its transformation matrix T is a non-singular matrix, i.e.

$$T := [c^{\mathrm{T}}, (cA)^{\mathrm{T}}, \ldots, (cA^{n-1})^{\mathrm{T}}]^{\mathrm{T}} = I_n. \tag{7}$$

Identification model: Due to the unknown state variable in (1), it needs to derive an alternative expression as feasible identification structure that only involves the available input–output data $\{u(t), y(t)\}$. The following transforms the state-space model with time delay into an input–output representation as its identification model.

From (1) to (4), it has

$$\begin{aligned} y(t+i) &= cA^i x(t) + cA^{i-1}Bx(t-1) + cA^{i-2}Bx(t) \\ &+ \cdots + cBx(t-1+i-1) + cA^{i-1}gf(u(t))p_1 \\ &+ cA^{i-2}gf(u(t+1))p_2 + \cdots \\ &+ cgf(u(t+i-1))p_m + w(t+i), \\ &i = 0, 1, \ldots, n-1, \end{aligned} \tag{8}$$

$$\begin{aligned} y(t+n) &= cA^n x(t) + cA^{n-1}Bx(t-1) + cA^{n-2}Bx(t) \\ &+ \cdots + cBx(t-2+n) + cA^{n-1}gf(u(t))p_1 \\ &+ cA^{n-2}gf(u(t+1))p_2 + \cdots \\ &+ cgf(u(t+n-1))p_m + w(t+n). \end{aligned} \tag{9}$$

Define the following vectors/matrices:

$$\begin{aligned} \phi_y(t) &:= [y(t), y(t+1), \ldots, y(t+n-1)]^{\mathrm{T}} \in \mathbb{R}^n, \\ \phi_u(t) &:= [f(u(t)), f(u(t+1)), \ldots, \\ &\quad f(u(t+n-1))]^{\mathrm{T}} \in \mathbb{R}^n, \\ X(t-1+n) &:= [x^{\mathrm{T}}(t-1), x^{\mathrm{T}}(t), \ldots, \\ &\quad x^{\mathrm{T}}(t-2+n)]^{\mathrm{T}} \in \mathbb{R}^{n^2}, \\ \phi_w(t) &:= [w(t), w(t+1), \ldots, w(t+n-1)]^{\mathrm{T}} \in \mathbb{R}^n, \end{aligned}$$

$$M := \begin{bmatrix} 0 & \cdots & 0 & 0 \\ cB & \cdots & 0 & 0 \\ cAB & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & 0 \\ cA^{n-2}B & \cdots & cB & 0 \end{bmatrix} \in \mathbb{R}^{n\times n^2},$$

$$Q := \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ cgp_m & \cdots & \mathbf{0} & \mathbf{0} \\ cAgp_{m-1} & \ddots & \vdots & \vdots \\ \vdots & \ddots & \mathbf{0} & \mathbf{0} \\ cA^{n-2}gp_2 & \cdots & cgp_m & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

From (8) and (9) and the above definitions, it gives

$$\begin{aligned} \phi_y(t+n) &= Tx(t) + MX(t-1+n) + Q\phi_u(t+n) \\ &+ \phi_w(t+n) \\ &= x(t) + MX(t-1+n) + Q\phi_u(t+n) \\ &+ \phi_w(t+n), \end{aligned}$$

or

$$\begin{aligned} x(t) &= \phi_y(t+n) - MX(t-1+n) - Q\phi_u(t+n) \\ &- \phi_w(t+n). \end{aligned} \tag{10}$$

Define the information vectors $\varphi_s(t)$, $\varphi_n(t)$ and the parameter vector θ as

$$\varphi_s(t) := \begin{bmatrix} \phi_y(t) \\ X(t-1) \\ \phi_u(t) \end{bmatrix} \in \mathbb{R}^{2n+n^2},$$

$$\varphi_n(t) := \begin{bmatrix} -\phi_w(t) \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{2n+n^2},$$

$$\theta := \begin{bmatrix} \theta_a \\ \theta_b \\ \theta_c \end{bmatrix} \in \mathbb{R}^{2n+n^2},$$

$$\theta_a := [cA^n]^{\mathrm{T}} \in \mathbb{R}^n,$$

$$\theta_b := \left[-cA^n M + [cA^{n-1}B, \ldots, cB]\right]^{\mathrm{T}} \in \mathbb{R}^{n^2},$$

$$\theta_c := \left[-cA^n Q + [cA^{n-1}gp_1, \ldots, cgp_m]\right]^{\mathrm{T}} \in \mathbb{R}^n.$$

Inserting (10) into (9) yields

$$
\begin{aligned}
y(t+n) &= cA^{n}[\phi_{y}(t+n) - MX(t-1+n) \\
&\quad -Q\phi_{u}(t+n) - \phi_{w}(t+n)] \\
&\quad +cA^{n-1}Bx(t-1) + cA^{n-2}Bx(t) + \cdots \\
&\quad +cBx(t-2+n) + cA^{n-1}gp_{1}f(u(t)) \\
&\quad +cA^{n-2}gp_{2}f(u(t+1)) + \cdots \\
&\quad +cgf(u(t+n-1))p_{m} + w(t+n) \\
&= cA^{n}[\phi_{y}(t+n) - MX(t-1+n) \\
&\quad -Q\phi_{u}(t+n) - \phi_{w}(t+n)] \\
&\quad +[cA^{n-1}B, \ldots, cB]\begin{bmatrix} x(t-1) \\ x(t) \\ \vdots \\ x(t-2+n) \end{bmatrix} \\
&\quad +[cA^{n-1}gp_{1}, cA^{n-2}gp_{2}, \ldots, cgp_{m}] \\
&\quad \times \begin{bmatrix} f(u(t)) \\ f(u(t+1)) \\ \vdots \\ f(u(t+n-1)) \end{bmatrix} + w(t+n) \\
&= cA^{n}\phi_{y}(t+n) + [-cA^{n}M + [cA^{n-1}B, \\
&\quad cA^{n-2}B, \ldots, cB]]X(t-1+n) \\
&\quad +[-cA^{n}Q + [cA^{n-1}gp_{1}, \ldots, cgp_{m}]] \\
&\quad \times \phi_{u}(t+n) - cA^{n}\phi_{w}(t+n) + w(t+n) \\
&= \phi_{y}^{\mathrm{T}}(t+n)\theta_{a} + X^{\mathrm{T}}(t-1+n)\theta_{b} \\
&\quad +\phi_{u}^{\mathrm{T}}(t+n)\theta_{c} - \phi_{w}^{\mathrm{T}}(t+n)\theta_{a} + w(t+n) \\
&= \varphi_{s}^{\mathrm{T}}(t+n)\theta + \varphi_{n}^{\mathrm{T}}(t+n)\theta + w(t+n).
\end{aligned}
\tag{11}
$$

Replacing $t+n$ with t in (11) gives

$$y(t) = \varphi_{s}^{\mathrm{T}}(t)\theta + \varphi_{n}^{\mathrm{T}}(t)\theta + w(t).$$

Assume that $\varphi_{n}^{\mathrm{T}}(t)\theta + w(t) =: e(t)$, we have

$$y(t) = \varphi_{s}^{\mathrm{T}}(t)\theta + e(t). \tag{12}$$

The above expression is the system identification model in regression form of the state-space descriptive model of (1)–(4), where the information vector $\phi_{y}(t)$ and $\phi_{u}(t)$ can be formed with known data. Define the parameter vector θ_{v} and the information vector $\varphi_{v}(t)$ as

$$
\begin{aligned}
\theta_{v} &:= [d_{1}, d_{2}, \ldots, d_{n_d}] \in \mathbb{R}^{n_d}, \\
\varphi_{v}(t) &:= [v(t-1), v(t-2), \ldots, v(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^{n_d}.
\end{aligned}
$$

From (3), it has the noise model

$$w(t) = \varphi_{v}^{\mathrm{T}}(t)\theta_{v} + v(t). \tag{13}$$

Remark 1: As the state $x(t)$ and the noise $w(t)$ are unknown in descriptive model of (1)–(4), it is necessary to derive its alternative feasible identification model for only using the available input–output data $\{u(t), y(t)\}$.

Remark 2: According to the authors' knowledge, the parameter estimation of non-linear state-space systems has been studied for many years. Although a least-squares algorithm has been developed for state-space system with colored noise, the state variable is eliminated, so the calculation is large, but the state estimation is not considered. In addition, less work is focused on the theoretical analysis of the performance of the proposed algorithm. This inspired the research to explore new and effective methods for state and parameter combinational estimation, and to provide the theoretical analysis for the convergence of the proposed algorithm.

3 Parameter estimation algorithm

Base on the system model of (12) and the noise model of (13), this section derives the estimation algorithm with the bias compensation.

Using the least-squares search and minimising the cost function give the least-squares estimate of the parameter vector θ

$$\hat{\theta}_{\mathrm{LS}}(t) = \left[\sum_{i=1}^{t}\varphi_{s}(i)\varphi_{s}^{\mathrm{T}}(i)\right]^{-1}\sum_{i=1}^{t}\varphi_{s}(i)y(i). \tag{14}$$

Using (12) and (14) give

$$
\begin{aligned}
&\left[\sum_{i=1}^{t}\varphi_{s}(i)\varphi_{s}^{\mathrm{T}}(i)\right][\hat{\theta}_{\mathrm{LS}}(t) - \theta] \\
&\quad = \sum_{i=1}^{t}\varphi_{s}(i)y(i) - \sum_{i=1}^{t}\varphi_{s}(i)\varphi_{s}^{\mathrm{T}}(i)\theta \\
&\quad = \sum_{i=1}^{t}\varphi_{s}(i)[y(i) - \varphi_{s}^{\mathrm{T}}(i)\theta] \\
&\quad = \sum_{i=1}^{t}\varphi_{s}(i)[\varphi_{n}^{\mathrm{T}}(i)\theta + w(i)].
\end{aligned}
$$

Dividing by t and taking limits on both sides give

$$
\begin{aligned}
&\lim_{t\to\infty}\frac{1}{t}\left[\left[\sum_{i=1}^{t}\varphi_{s}(i)\varphi_{s}^{\mathrm{T}}(i)\right][\hat{\theta}_{\mathrm{LS}}(t) - \theta]\right] \\
&\quad = \lim_{t\to\infty}\left[\frac{1}{t}\sum_{i=1}^{t}\varphi_{s}(i)\varphi_{n}^{\mathrm{T}}(i)\right]\theta \\
&\qquad + \lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}\varphi_{s}(i)w(i).
\end{aligned}
\tag{15}
$$

Note that $w(t)$ in (15) is the moving average noise and $v(t)$ is the white noise with zero mean and variance δ and is independent of the inputs. From (5), we have

$$
\begin{aligned}
H(0) &:= E[w^{2}(t)] \\
&= [1 + \theta_{v}^{\mathrm{T}}\theta_{v}]\delta, \\
H(i) &:= E[w(t-i)w(t)] \\
&:= \theta_{v}^{\mathrm{T}}(i:n_d)\begin{bmatrix} 1 \\ \theta_{v}(1:n_d - i) \end{bmatrix}\delta, \ i = 1, 2, \ldots, n_d,
\end{aligned}
$$

where $H(i)$ is the autocorrelation function of the noise $w(t)$ and $H(i) = 0$ when $i > n_d$. Define the autocorrelation function vectors h and the autocorrelation function matrices H, Λ as

$$h := [H(n^{2}), H(n^{2}-1), \ldots, H(1), 0, 0, \ldots, 0] \in \mathbb{R}^{2n+n^{2}}, \tag{16}$$

$$H := \mathrm{diag}[\Lambda, 0] \in \mathbb{R}^{(2n+n^{2})\times(2n+n^{2})}, \tag{17}$$

$$\Lambda := \begin{bmatrix} H(0) & \cdots & H(n^{2}-1) \\ H(1) & \cdots & H(n^{2}-2) \\ \vdots & \ddots & \vdots \\ H(n^{2}-1) & \cdots & H(0) \end{bmatrix} \in \mathbb{R}^{n^{2}\times n^{2}}. \tag{18}$$

Equation (15) can be rewritten as

$$\lim_{t\to\infty}\hat{\theta}_{\mathrm{LS}}(t) = \theta - tP(t)(H\vartheta - h).$$

The above equation shows that the least-squares estimate of the model parameters is biased. If compensation $tP(t)(H\vartheta - h)$ is introduced into the least-squares estimate $\hat{\theta}_{LS}(t)$, $\hat{\theta}_C(t) = \hat{\theta}_{LS}(t) + tP(t)[\hat{H}(t)\hat{\theta}_C(t) - \hat{h}(t)]$ is the unbiased estimation of θ. This is the basic idea of the deviation compensated least-squares method. It is written recursively as

$$\hat{\theta}_C(t) = \hat{\theta}_{LS}(t) + tP(t)[\hat{H}(t)\hat{\theta}_C(t) - \hat{h}(t)], \tag{19}$$

where $\hat{\theta}_C(t)$ and $\hat{\theta}_{LS}(t)$ are the bias compensation least-squares estimation and the least-squares estimation at time t of parameter θ, respectively. $\hat{\theta}_C(t)$ is related to the estimates of H and h (i.e. δ and θ_v). The following gives the estimates based on the interactive identification. Define the covariance matrix

$$P(t) = \left[\sum_{i=1}^{t} \varphi_s(i)\varphi_s^T(i)\right]^{-1}.$$

The question turns into how to estimate $\hat{\delta}(t)$. Define

$$\varepsilon_{LS}(i) = y(i) - \varphi_s^T(i)\hat{\theta}_{LS}(t), \quad i = 1, 2, \ldots, t.$$

Using (12) and the following

$$\sum_{i=1}^{t} \varepsilon_{LS}(i)\varphi_s^T(i) = \mathbf{0},$$

we have

$$\begin{aligned}\sum_{i=1}^{t} \varepsilon_{LS}^2(i) &= \sum_{i=1}^{t} \varepsilon_{LS}(i)[y(i) - \varphi_s^T(i)\hat{\theta}_{LS}(t)] \\ &= \sum_{i=1}^{t} \varphi_s^T(i)[\theta - \hat{\theta}_{LS}(t)][\varphi_n^T(i)\theta + w(i)] \\ &+ \sum_{i=1}^{t} [\varphi_n^T(i)\theta + w(i)]^2.\end{aligned} \tag{20}$$

Divide both sides of the above equation by t to take the limit, and use the white noise characteristics of $v(t)$ to obtain

$$\begin{aligned}\lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} \varepsilon_{LS}^2(i) &= \lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} \varphi_s^T(i)[\theta - \hat{\theta}_{LS}(t)] \\ &\times [\varphi_n^T(i)\theta + w(i)] \\ &+ \lim_{t\to\infty} \frac{1}{t}\sum_{i=1}^{t} [\varphi_n^T(i)\theta + w(i)]^2 \\ &= \theta^T h\hat{\theta}_{LS}(t) - h^T[\theta + \hat{\theta}_{LS}(t)] \\ &+ \delta[1 + \theta_v^T\theta_v].\end{aligned} \tag{21}$$

In conclusion, the bias compensation-based recursive least-squares algorithm is as follows:

$$\hat{\theta}_C(t) = \hat{\theta}_{LS}(t) + tP(t)[\hat{H}(t)\hat{\theta}_C(t) - \hat{h}(t)], \tag{22}$$

$$\hat{\theta}_{LS}(t) = \hat{\theta}_{LS}(t-1) + L(t)[y(t) - \varphi_s^T(t)\hat{\theta}_{LS}(t-1)], \tag{23}$$

$$L(t) = \frac{P(t-1)\varphi_s(t)}{1 + \varphi_s^T(t)P(t-1)\varphi_s(t)}, \tag{24}$$

$$P(t) = [I - L(t)\varphi_s^T(t)]P(t-1),\ P(0) = p_0 I, \tag{25}$$

$$\hat{\delta}(t) = \frac{1}{\chi(t)t}J(t), \tag{26}$$

$$\begin{aligned}\chi(t) &= \hat{\theta}_C^T(t)\hat{R}(t)\hat{\theta}_{LS}(t) - \hat{h}^T(t)[\hat{\theta}_C(t) + \hat{\theta}_{LS}(t)]/\delta \\ &+ 1 + \hat{\theta}_v^T(t)\hat{\theta}_v(t),\end{aligned} \tag{27}$$

$$J(t) = J(t-1) + \frac{[y(t) - \varphi_s^T(t)\hat{\theta}_{LS}(t-1)]^2}{1 + \varphi_s^T(t)P(t-1)\varphi_s(t)}, \tag{28}$$

$$\hat{h}(t) = [\hat{H}(n), \hat{H}(n-1), \ldots, \hat{H}(1), 0, 0, \ldots, 0], \tag{29}$$

$$\hat{R}(t) = \begin{bmatrix} 1 + \hat{\theta}_v^T(t)\hat{\theta}_v(t) & \cdots & \frac{\hat{H}(n-1)}{\hat{\delta}(t-1)} & \mathbf{0} \\ \frac{\hat{H}(1)}{\hat{\delta}(t-1)} & \cdots & \frac{\hat{H}(n-2)}{\hat{\delta}(t-1)} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\hat{H}(n-1)}{\hat{\delta}(t-1)} & \cdots & 1 + \hat{\theta}_v^T(t)\hat{\theta}_v(t) & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{30}$$

$$\frac{\hat{H}(i)}{\hat{\delta}(t-1)} = \hat{\theta}_v^T(i{:}n_d)\begin{bmatrix} 1 \\ \hat{\theta}_v(1{:}n_d - i)\end{bmatrix}, \quad i = 1, 2, \ldots, n_d.$$

The above equation involves the estimate $\hat{\theta}_v(t)$ of the noise parameter vector θ_v, which can be computed by the noise model. Using the least-squares principle, the estimate $\hat{\theta}_v(t)$ of $\theta_v(t)$ is as follows:

$$\hat{\theta}_v(t) = \hat{\theta}_v(t-1) + P_v(t)\hat{\varphi}_v(t)[\hat{w}(t) - \varphi_v^T(t)\hat{\theta}_v(t-1)], \tag{31}$$

$$P_v^{-1}(t) = P_v^{-1}(t-1) + \hat{\varphi}_v(t)\hat{\varphi}_v^T(t), \tag{32}$$

$$\hat{\varphi}_v(t) = [\hat{v}(t-1), \hat{v}(t-2), \ldots, \hat{v}(t-n_d)]^T, \tag{33}$$

$$\hat{w}(t) = y(t) - \varphi_s^T(t)\hat{\theta}_C(t) - \varphi_n^T(t)\hat{\theta}_C(t), \tag{34}$$

$$\hat{v}(t) = \hat{w}(t) - \varphi_v^T(t)\hat{\theta}_v(t). \tag{35}$$

Remark 3: The difficulty of such identification is that the information vectors $\varphi_s(t)$ and $\varphi_n(t)$ contain the unknown inner state variable $X(t)$ and coloured noise $\phi_w(t)$, respectively. The new solution here is to use the measured data $u(t)$ and $y(t)$, and to replace the unknown $x(t-i)$ and $w(t)$ in the identification algorithm with the estimated values $\hat{x}(t-i)$ and $\hat{w}(t)$.

Remark 4: For the purpose of reducing the influence of the coloured noise on parameter estimation, the interactive estimation process is introduced by using a compensation-based parameter estimation algorithm: compute the estimates $\hat{\theta}_v(t)$ and $\hat{\delta}(t)$ by the output $y(t)$ and the information vectors $\varphi_s(t)$, $\varphi_v(t)$; then update the unbiased estimate $\hat{\theta}_C(t)$ by these obtained variable estimates.

The computational burden of the bias compensation-based recursive least-squares algorithm is shown in Table 1 and we assume that $n_l = 2n + n^2$.

4 State estimation algorithm

In the principle of the hierarchical identification, the first step is to recursively calculate the parameter estimation based on the input–output data of the system, and the second step is to recursively calculate the state estimation of the system based on the input–output data and the obtained parameter estimation.

Since the relationship between the parameter vector θ and the matrices/vectors A, B, g, p and d has been established, here look for simple algebra to establish the relationship between the parameter vector θ and the parameters A, B, g, p and d.

Table 1 Computational burden of the bias compensation-based recursive least- squares algorithm

Computational sequences	Number of multiplications	Number of additions
$\hat{\boldsymbol{\theta}}_{\mathrm{C}}(t) = \hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t) + t\boldsymbol{P}(t)[\hat{\boldsymbol{H}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{C}}(t) - \hat{\boldsymbol{h}}(t)] \in \mathbb{R}^{n_l}$	$2n_l^2 + n_l$	$2n_l^2$
$\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t) = \hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t-1) + \boldsymbol{L}(t)[y(t) - \boldsymbol{\varphi}_s^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t-1)] \in \mathbb{R}^{n_l}$	$2n_l$	$2n_l$
$\boldsymbol{L}(t) = \frac{\boldsymbol{P}(t-1)\boldsymbol{\varphi}_s(t)}{1+\boldsymbol{\varphi}_s^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}_s(t)} \in \mathbb{R}^{n_l}$	$n_l^2 + 2n_l$	n_l^2
$\boldsymbol{P}(t) = [\boldsymbol{I} - \boldsymbol{L}(t)\boldsymbol{\varphi}_s^{\mathrm{T}}(t)]\boldsymbol{P}(t-1) \in \mathbb{R}^{n_l \times n_l}$	$2n_l^2$	$2n_l^2 - n_l$
$\hat{\delta}(t) = \frac{1}{\chi(t)t}J(t)$	2	
$\chi(t) = \hat{\boldsymbol{\theta}}_{\mathrm{C}}^{\mathrm{T}}(t)\hat{\boldsymbol{R}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t) - \hat{\boldsymbol{h}}^{\mathrm{T}}(t)[\hat{\boldsymbol{\theta}}_{\mathrm{C}}(t) + \hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t)]/\delta + 1 + \hat{\boldsymbol{\theta}}_v^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_v(t)$	$n_l^2 + 2n_l + n_d + 1$	$n_l^2 + 2n_l + n_d$
$J(t) = J(t-1) + \frac{[y(t) - \boldsymbol{\varphi}_s^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(t-1)]^2}{1+\boldsymbol{\varphi}_s^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\boldsymbol{\varphi}_s(t)}$	$n_l^2 + 2n_l + 2$	$n_l^2 + n_l + 1$
sum	$7n_l^2 + 9n_l + n_d + 5$	$7n_l^2 + 4n_l + n_d + 1$
total flops	$14n^2 + 13n + 2n_d + 6$	

Post-multiplying (7) by $\boldsymbol{B}$ gives

$$\boldsymbol{c}\boldsymbol{A}^{i-1}\boldsymbol{B} = \boldsymbol{b}_i, \quad i = 1, 2, \ldots, n.$$

Then, the parameter matrix $\boldsymbol{M}$ can be expressed as

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{b}_1 & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{b}_2 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{b}_{n-1} & \cdots & \boldsymbol{b}_1 & \boldsymbol{0} \end{bmatrix}.$$

Post-multiplying (7) by $\boldsymbol{A}$ and write down the nth row

$$\boldsymbol{c}\boldsymbol{A}^n = [-a_n, -a_{n-1}, \ldots, -a_1].$$

From the definitions of $\boldsymbol{\theta}_a$, $\boldsymbol{\theta}_b$ and $\boldsymbol{\theta}_c$, it gives

$$\begin{aligned}
\boldsymbol{\theta}_a &= [\boldsymbol{c}\boldsymbol{A}^n]^{\mathrm{T}} = [-a_n, -a_{n-1}, \ldots, -a_1]^{\mathrm{T}}, \\
\boldsymbol{\theta}_b &= [-\boldsymbol{c}\boldsymbol{A}^n\boldsymbol{M} + [\boldsymbol{c}\boldsymbol{A}^{n-1}\boldsymbol{B}, \boldsymbol{c}\boldsymbol{A}^{n-2}\boldsymbol{B}, \ldots, \boldsymbol{c}\boldsymbol{B}]]^{\mathrm{T}} \\
&= \begin{bmatrix} \boldsymbol{b}_1^{\mathrm{T}}a_{n-1} + \boldsymbol{b}_2^{\mathrm{T}}a_{n-2} + \cdots + \boldsymbol{b}_{n-1}^{\mathrm{T}}a_1 + \boldsymbol{b}_n^{\mathrm{T}} \\ \boldsymbol{b}_1^{\mathrm{T}}a_{n-2} + \boldsymbol{b}_2^{\mathrm{T}}a_{n-3} + \cdots + \boldsymbol{b}_{n-2}^{\mathrm{T}}a_1 + \boldsymbol{b}_{n-1}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{b}_1^{\mathrm{T}}a_1 + \boldsymbol{b}_2^{\mathrm{T}} \\ \boldsymbol{b}_1^{\mathrm{T}} \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{b}_1^{\mathrm{T}} & \cdots & \boldsymbol{b}_{n-1}^{\mathrm{T}} & \boldsymbol{b}_n^{\mathrm{T}} \\ \boldsymbol{0} & \cdots & \boldsymbol{b}_{n-2}^{\mathrm{T}} & \boldsymbol{b}_{n-1}^{\mathrm{T}} \\ \vdots & \vdots & \vdots & \\ \boldsymbol{0} & \cdots & \boldsymbol{b}_1^{\mathrm{T}} & \boldsymbol{b}_2^{\mathrm{T}} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{b}_1^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \\ 1 \end{bmatrix}, \\
\boldsymbol{\theta}_c &= [-\boldsymbol{c}\boldsymbol{A}^n\boldsymbol{Q} + [\boldsymbol{c}\boldsymbol{A}^{n-1}\boldsymbol{g}p_1, \boldsymbol{c}\boldsymbol{A}^{n-2}\boldsymbol{g}p_2, \cdots, \boldsymbol{c}\boldsymbol{g}p_m]]^{\mathrm{T}} \\
&= \begin{bmatrix} g_1p_ma_{n-1} + \cdots + g_{n-1}p_2a_1 + g_np_1 \\ g_1p_ma_{n-2} + \cdots + g_{n-2}p_3a_1 + g_{n-1}p_2 \\ \vdots \\ g_1p_ma_1 + g_2p_{m-1} \\ g_1p_m \end{bmatrix} \\
&= \begin{bmatrix} a_{n-1} & \cdots & a_1 & 1 \\ a_{n-2} & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} g_1p_m \\ g_2p_{m-1} \\ \vdots \\ g_{n-1}p_2 \\ g_np_1 \end{bmatrix}.
\end{aligned}$$

Then the state estimation algorithm is formulated as below:

$$\begin{aligned} \hat{\boldsymbol{x}}(t-n) = {} & \boldsymbol{\phi}_y(t) - \hat{\boldsymbol{M}}(t)\hat{\boldsymbol{X}}(t-1) \\ & -\hat{\boldsymbol{Q}}(t)\boldsymbol{\phi}_u(t)\hat{\boldsymbol{p}}(t) - \hat{\boldsymbol{\phi}}_w(t), \end{aligned} \tag{36}$$

$$\boldsymbol{\phi}_y(t) = [y(t), y(t+1), \ldots, y(t+n-1)]^{\mathrm{T}}, \tag{37}$$

$$\boldsymbol{\phi}_u(t) = [\boldsymbol{f}(u(t)), \boldsymbol{f}(u(t+1)), \ldots, \boldsymbol{f}(u(t+n-1))]^{\mathrm{T}}, \tag{38}$$

$$\hat{\boldsymbol{X}}(t-1) = \hat{\boldsymbol{x}}^{\mathrm{T}}(t-1-n), \hat{\boldsymbol{x}}^{\mathrm{T}}(t-n), \ldots, \hat{\boldsymbol{x}}^{\mathrm{T}}(t-2)]^{\mathrm{T}}, \tag{39}$$

$$\hat{\boldsymbol{\phi}}_w(t) = [\hat{w}(t), \hat{w}(t+1), \ldots, \hat{w}(t+n-1)]^{\mathrm{T}}, \tag{40}$$

$$\hat{\boldsymbol{M}}(t) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \hat{\boldsymbol{b}}_1(t) & 0 & \cdots & 0 & 0 \\ \hat{\boldsymbol{b}}_2(t) & \hat{\boldsymbol{b}}_1(t) & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ \hat{\boldsymbol{b}}_{n-1}(t) & \hat{\boldsymbol{b}}_{n-2}(t) & \cdots & \hat{\boldsymbol{b}}_1(t) & 0 \end{bmatrix}, \tag{41}$$

$$\hat{\boldsymbol{Q}}(t) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ \hat{g}_1(t)\hat{p}_m(t) & \cdots & 0 & 0 \\ \hat{g}_2(t)\hat{p}_{m-1}(t) & \ddots & \vdots & \vdots \\ \vdots & \ddots & 0 & 0 \\ \hat{g}_{n-1}(t)\hat{p}_2(t) & \cdots & \hat{g}_1(t)\hat{p}_m(t) & 0 \end{bmatrix}. \tag{42}$$

The proposed parameter and state estimation algorithm can be extended to other scalar and multivariable stochastic systems [34–38] and non-linear stochastic systems [39–45] and can be applied to other fields [46–50] such as signal processing.

Theorem 1: For the system in (1)–(4), suppose that $\{v(t)\}$ is a white noise sequence with zero mean and variance σ^2, i.e. $E[v(t)] = 0$, $E[v^2(t)] = \sigma^2$, and that there exist constants $0 < \tau_i < \infty$ such that for large t, the following persistent excitation conditions hold:

$$\tau_1\boldsymbol{I} \leqslant \frac{1}{t}\sum_{j=1}^{t}\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(j)\hat{\boldsymbol{\phi}}_y(j) \leqslant \tau_2\boldsymbol{I}, \text{ a.s.}$$

$$\tau_3\boldsymbol{I} \leqslant \frac{1}{t}\sum_{j=1}^{t}\hat{\boldsymbol{X}}^{\mathrm{T}}(j)\hat{\boldsymbol{X}}^{\mathrm{T}}(j) \leqslant \tau_4\boldsymbol{I}, \text{ a.s.}$$

$$\tau_5\boldsymbol{I} \leqslant \frac{1}{t}\sum_{j=1}^{t}\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(j)\hat{\boldsymbol{\phi}}_u(j) \leqslant \tau_6\boldsymbol{I}, \text{a.s.},$$

$$\tau_7\boldsymbol{I} \leqslant \frac{1}{t}\sum_{j=1}^{t}\hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(j)\hat{\boldsymbol{\phi}}_w(j) \leqslant \tau_8\boldsymbol{I}, \text{ a.s.}$$

The parameter estimation error given by the bias compensation least-squares algorithm converges to zero under the mean square.

Proof: The output equation in identification model is as follows:

$$y(t) = \boldsymbol{\phi}_y^{\mathrm{T}}(t)\boldsymbol{\theta}_a + \boldsymbol{X}^{\mathrm{T}}(t-1)\boldsymbol{\theta}_b + \boldsymbol{\phi}_u^{\mathrm{T}}(t)\boldsymbol{\theta}_c - \boldsymbol{\phi}_w^{\mathrm{T}}(t)\boldsymbol{\theta}_a + w(t)$$
$$= [\boldsymbol{\phi}_y^{\mathrm{T}}(t) - \boldsymbol{\phi}_w^{\mathrm{T}}(t)]\boldsymbol{\theta}_a + \boldsymbol{X}^{\mathrm{T}}(t-1)\boldsymbol{\theta}_b + \boldsymbol{\phi}_u^{\mathrm{T}}(t)\boldsymbol{\theta}_c + w(t).$$

Define the parameter estimation error vectors

$$\tilde{\boldsymbol{\theta}}_a(t) := \hat{\boldsymbol{\theta}}_a(t) - \boldsymbol{\theta}_a,$$
$$\tilde{\boldsymbol{\theta}}_b(t) := \hat{\boldsymbol{\theta}}_b(t) - \boldsymbol{\theta}_b,$$
$$\tilde{\boldsymbol{\theta}}_c(t) := \hat{\boldsymbol{\theta}}_c(t) - \boldsymbol{\theta}_c,$$
$$\tilde{\boldsymbol{\theta}}_v(t) := \hat{\boldsymbol{\theta}}_v(t) - \boldsymbol{\theta}_v.$$

Define some variables

$$\tilde{y}(t) := [\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\tilde{\boldsymbol{\theta}}_a(t-1) + \hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1) + \hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1),$$
$$\tilde{w}(t) := \hat{\boldsymbol{\varphi}}_v^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_v(t-1),$$
$$\Upsilon_1(t) := \hat{y}(t) - y(t) + [\boldsymbol{\phi}_y^{\mathrm{T}}(t) - \boldsymbol{\phi}_w^{\mathrm{T}}(t) - [\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]]\boldsymbol{\theta}_a + [\boldsymbol{X}^{\mathrm{T}}(t-1) - \hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)]\boldsymbol{\theta}_b + [\boldsymbol{\phi}_u^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)]\boldsymbol{\theta}_c,$$
$$\Upsilon_2(t) := \hat{w}(t) - w(t) + [\boldsymbol{\varphi}_v^{\mathrm{T}}(t) - \hat{\boldsymbol{\varphi}}_v^{\mathrm{T}}(t)]\boldsymbol{\theta}_v,$$
$$e_v(t) := y(t) - [\boldsymbol{\phi}_y^{\mathrm{T}}(t) - \boldsymbol{\phi}_w^{\mathrm{T}}(t)]\hat{\boldsymbol{\theta}}_a(t-1) - \hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\hat{\boldsymbol{\theta}}_b(t-1) - \hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_c(t-1) - \hat{\boldsymbol{\varphi}}_v^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_v(t-1)$$
$$= \hat{w}(t) - \hat{\boldsymbol{\varphi}}_v^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_v(t-1).$$

From (12) and (13), it gives

$$e(t) := -\tilde{y}(t) + \Upsilon_1(t) + v(t),$$
$$e_v(t) := -\tilde{w}(t) + \Upsilon_2(t) + v(t).$$

The parameter estimation error vectors are in the least-squares form

$$\tilde{\boldsymbol{\theta}}_a(t) = \tilde{\boldsymbol{\theta}}_a(t-1) + \boldsymbol{P}_a(t)[\boldsymbol{\phi}_y^{\mathrm{T}}(t) - \boldsymbol{\phi}_w^{\mathrm{T}}(t)]e(t),$$
$$\tilde{\boldsymbol{\theta}}_b(t) = \tilde{\boldsymbol{\theta}}_b(t-1) + \boldsymbol{P}_b(t)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)e(t),$$
$$\tilde{\boldsymbol{\theta}}_c(t) = \tilde{\boldsymbol{\theta}}_c(t-1) + \boldsymbol{P}_c(t)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)e(t),$$
$$\tilde{\boldsymbol{\theta}}_v(t) = \tilde{\boldsymbol{\theta}}_v(t-1) + \boldsymbol{P}_v(t)\hat{\boldsymbol{\phi}}_v^{\mathrm{T}}(t)e(t).$$

Define the non-negative functions

$$\Omega_1(t) := \tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t)\boldsymbol{P}_a^{-1}(t)\tilde{\boldsymbol{\theta}}_a(t),$$
$$\Omega_2(t) := \tilde{\boldsymbol{\theta}}_b^{\mathrm{T}}(t)\boldsymbol{P}_b^{-1}(t)\tilde{\boldsymbol{\theta}}_b(t),$$
$$\Omega_3(t) := \tilde{\boldsymbol{\theta}}_c^{\mathrm{T}}(t)\boldsymbol{P}_c^{-1}(t)\tilde{\boldsymbol{\theta}}_c(t),$$
$$\Omega_4(t) := \tilde{\boldsymbol{\theta}}_v^{\mathrm{T}}(t)\boldsymbol{P}_v^{-1}(t)\tilde{\boldsymbol{\theta}}_v(t),$$
$$\Omega_{123}(t) := \Omega_1(t) + \Omega_2(t) + \Omega_3(t),$$
$$\Omega(t) := \Omega_{123}(t) + \Omega_4(t),$$
$$\alpha_1(t) := [\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}} + \hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\boldsymbol{P}_b(t)\hat{\boldsymbol{X}}(t-1) + \hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\boldsymbol{P}_c(t)\hat{\boldsymbol{\phi}}_u(t),$$
$$\alpha_2(t) := \hat{\boldsymbol{\phi}}_v^{\mathrm{T}}(t)\boldsymbol{P}_v(t)\hat{\boldsymbol{\phi}}_v(t).$$

Then, we have

$$\Omega_1(t) = \tilde{\boldsymbol{\theta}}_a(t-1) + \boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)^{\mathrm{T}}\boldsymbol{P}_a^{-1}(t)$$
$$\times \{\tilde{\boldsymbol{\theta}}_a(t-1) + \boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)\}$$
$$= \tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t-1)\boldsymbol{P}_a^{-1}(t)\tilde{\boldsymbol{\theta}}_a(t-1)$$
$$+2\tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t-1)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)$$
$$+e^{\mathrm{T}}(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)$$
$$= \Omega_1(t-1) + \tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t-1)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}$$
$$\times [\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]$$
$$+2\tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t-1)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)$$
$$+e^{\mathrm{T}}(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)$$
$$\leqslant \Omega_1(t-1) + [[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\tilde{\boldsymbol{\theta}}_a(t-1)]^2$$
$$+2\tilde{\boldsymbol{\theta}}_a^{\mathrm{T}}(t-1)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e(t)$$
$$+[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\boldsymbol{P}_a(t)[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]^{\mathrm{T}}e^2(t).$$

Similarly, we have

$$\Omega_2(t) = \Omega_2(t-1) + [\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)]^2$$
$$+2e(t)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)$$
$$+\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\boldsymbol{P}_b(t)\hat{\boldsymbol{X}}(t-1)e^2(t),$$
$$\Omega_3(t) = \Omega_3(t-1) + [\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1)]^2$$
$$+2e(t)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1)$$
$$+\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\boldsymbol{P}_c(t)\hat{\boldsymbol{\phi}}_u(t)e^2(t),$$
$$\Omega_4(t) \leqslant \Omega_4(t-1) + [\hat{\boldsymbol{\phi}}_v^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_v(t-1)]^2$$
$$+2\tilde{\boldsymbol{\theta}}_v^{\mathrm{T}}(t-1)\hat{\boldsymbol{\phi}}_v(t)e_v(t)$$
$$+\hat{\boldsymbol{\phi}}_v^{\mathrm{T}}(t)\boldsymbol{P}_v(t)\hat{\boldsymbol{\phi}}_v(t)e_v^2(t)$$
$$= \Omega_4(t-1) + \tilde{w}^2(t) + 2\tilde{w}(t)e_v(t) + \alpha_2(t)e_v^2(t)$$
$$= \Omega_4(t-1) - [1-\alpha_2(t)]\tilde{w}^2(t)$$
$$+2[1-\alpha_2(t)]\tilde{w}(t)[\Upsilon_2(t) + v(t)]$$
$$+\alpha_2(t)[v^2(t) + \Upsilon_2^2(t) + 2\Upsilon_2(t)v(t)],$$
$$\Omega_{123}(t) \leqslant \Omega_{123}(t-1) + \tilde{y}^2(t) - 2[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\tilde{\boldsymbol{\theta}}_a(t-1)$$
$$\times \hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1)$$
$$+2\tilde{y}(t)e(t) + \alpha_1(t)e^2(t)$$
$$= \Omega_{123}(t-1) - [1-\alpha_1(t)]\tilde{y}^2(t) - 2[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]$$
$$\times \tilde{\boldsymbol{\theta}}_a(t-1)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1)$$
$$+2[1-\alpha_1(t)]\tilde{y}(t)[\Upsilon_1(t) + v(t)]$$
$$+\alpha_1(t)[v^2(t) + \Upsilon_1^2(t) + 2\Upsilon_1(t)v(t)],$$

$$
\begin{aligned}
\Omega(t) \leqslant\ & \Omega(t-1) - [1-\alpha_1(t)]\tilde{y}^2(t) - 2[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)] \\
& \times \tilde{\boldsymbol{\theta}}_a(t-1)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1) \\
& + 2[1-\alpha_1(t)]\tilde{y}(t)[\Upsilon_1(t) + v(t)] \\
& - [1-\alpha_2(t)]\tilde{w}^2(t) + 2[1-\alpha_2(t)] \\
& \times \tilde{w}(t)[\Upsilon_2(t) + v(t)] \\
& + [\alpha_1(t) + \alpha_2(t)][2v^2(t) + \Upsilon_1^2(t) + 2\Upsilon_1(t)v(t) \\
& + \Upsilon_2^2(t) + 2\Upsilon_2(t)v(t)] \\
=\ & \Omega(t-1) - [1-\alpha_1(t)]\tilde{y}^2(t) - 2[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)] \\
& \times \tilde{\boldsymbol{\theta}}_a(t-1)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1) \\
& + 2[1-\alpha_1(t)]\tilde{y}(t)\Upsilon_1(t) - [1-\alpha_2(t)]\tilde{w}^2(t) \\
& + 2[1-\alpha_2(t)]\tilde{w}(t)\Upsilon_2(t) \\
& + [\alpha_1(t) + \alpha_2(t)][2v^2(t) + \Upsilon_1^2(t) + \Upsilon_2^2(t)] \\
& + 2[1-\alpha_1(t)]\tilde{y}(t)v(t) + 2[1-\alpha_2(t)]\tilde{w}(t)v(t) \\
& + [\alpha_1(t) + \alpha_2(t)][2\Upsilon_1(t)v(t) + 2\Upsilon_2(t)v(t)].
\end{aligned}
$$

Let

$$
\begin{aligned}
\alpha_3(t) :=\ & [1-\alpha_1(t)]\tilde{y}^2(t) + 2[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)] \\
& \times \tilde{\boldsymbol{\theta}}_a(t-1)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1) \\
& - 2[1-\alpha_1(t)]\tilde{y}(t)\Upsilon_1(t) + [1-\alpha_2(t)]\tilde{w}^2(t) \\
& - 2[1-\alpha_2(t)]\tilde{w}(t)\Upsilon_2(t).
\end{aligned}
$$

Since $\alpha_1(t)$, $\alpha_2(t)$, $\tilde{y}(t)$, $[\hat{\boldsymbol{\phi}}_y^{\mathrm{T}}(t) - \hat{\boldsymbol{\phi}}_w^{\mathrm{T}}(t)]\tilde{\boldsymbol{\theta}}_a(t-1)\hat{\boldsymbol{X}}^{\mathrm{T}}(t-1)\tilde{\boldsymbol{\theta}}_b(t-1)\hat{\boldsymbol{\phi}}_u^{\mathrm{T}}(t)\tilde{\boldsymbol{\theta}}_c(t-1)$, $\Upsilon_1(t)$, $\Upsilon_2(t)$ and $\tilde{w}(t)$ are uncorrelated with $v(t)$, taking the mathematical expectation yields

$$
\begin{aligned}
E[\Omega(t)] \le\ & E[\Omega(t-1)] - E[\alpha_3(t)] + E\{[\alpha_1(t) + \alpha_2(t)][\Upsilon_1^2(t) \\
& + \Upsilon_2^2(t) + 2v^2(t)]\}.
\end{aligned}
$$

If $\alpha_3(t) \ge 0$ and $\Upsilon_i^2(t) \le \varepsilon$, $\varepsilon < \infty$, we have

$$
\begin{aligned}
E[\Omega(t)] \le\ & E[\Omega(t-1)] + E\{[\alpha_1(t) + \alpha_2(t)][\Upsilon_1^2(t) \\
& + \Upsilon_2^2(t) + 2v^2(t)]\} \\
\le\ & E[\Omega(t-1)] + E\{[\alpha_1(t) + \alpha_2(t)](2\sigma^2 + 2\varepsilon)\}.
\end{aligned}
$$

Otherwise, let $\hat{\boldsymbol{\theta}}_a(t) = \hat{\boldsymbol{\theta}}_a(t-1)$, $\hat{\boldsymbol{\theta}}_b(t) = \hat{\boldsymbol{\theta}}_b(t-1)$, $\hat{\boldsymbol{\theta}}_c(t) = \hat{\boldsymbol{\theta}}_c(t-1)$, and $\hat{\boldsymbol{\theta}}_v(t) = \hat{\boldsymbol{\theta}}_v(t-1)$. Then we always have

$$
E[\Omega(t)] \le E[\Omega(t-1)] + E\{[\alpha_1(t) + \alpha_2(t)](2\sigma^2 + 2\varepsilon)\}.
$$

Then we can obtain the result in a similar way in [51]. □

5 Examples

Example 1: Consider an input non-linear state time-delay system, modelled by

$$
\begin{aligned}
\boldsymbol{x}(t+1) =\ & \begin{bmatrix} 0 & 1 \\ -0.05 & 0.10 \end{bmatrix}\boldsymbol{x}(t) \\
& + \begin{bmatrix} 0.24 & 0.11 \\ 0.19 & 0.10 \end{bmatrix}\boldsymbol{x}(t-1) + \begin{bmatrix} 0.50 \\ 1.00 \end{bmatrix}\bar{u}(t), \\
\bar{u}(t) =\ & f(u(t)) = p_1 f_1(u(t)) + p_2 f_2(u(t)) + p_3 f_3(u(t)) \\
=\ & u(t) + 5.00u^2(t) + 0.10u^3(t), \\
y(t) =\ & [1, 0]\boldsymbol{x}(t) + 0.10v(t-1) + v(t).
\end{aligned}
$$

The parameter vector to be identified is

$$
\begin{aligned}
\boldsymbol{\theta} =\ & [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}]^{\mathrm{T}} \\
=\ & [0.19, 0.21, 0.19, 0.10, 0.50, 1.00, 2.50, 5.00, 0.05, \\
& 0.10, 0.10]^{\mathrm{T}}.
\end{aligned}
$$

In simulation, the input $\{u(t)\}$ was assigned as an uncorrelated stochastic signal sequence with zero mean and unit variance, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$ and $\sigma^2 = 1.00^2$. Applying the compensation-based parameter estimation algorithm in (22)–(30) to estimate the parameters and the algorithm in (36)–(42) to estimate the states of this example system. The comparison of the bias compensation-based recursive least-squares algorithm (BC-RLS) and the least-squares algorithm (RLS) is shown in Fig. 1. The parameter estimates and their errors are shown in Table 2, the parameter estimation error versus t is shown in Fig. 2 and the state estimates $\hat{x}_1(t)$ and $\hat{x}_2(t)$ versus t are shown in Figs. 3 and 4.

From Table 2 and Figs. 1–4, we can draw the following conclusions:

• Compared with the RLS algorithm, the BC-RLS algorithm can enhance the parameter estimation accuracy.
• The proposed algorithm is effective for estimating the parameters of the non-linear state-space system with time delay. With the data length increasing, the parameter estimation errors become smaller and converge to zero.
• It is clear that the proposed state observer can generate accurate state estimates because the state estimates are close to their true values as t increases.

Example 2: An experiment was performed on a three tank system which is shown in Fig. 5. In this experiment, a non-linear function was designed for the input signal $f(u(t)) = u(t) + 5.00u^2(t) + 0.10u^3(t)$ to drive the pump to change the flow of the dollar inflow water. The water level in the middle tank was used as the system output signal. Therefore, the input function was considered as a static non-linear block. With the flow rate as input and the intermediate tank liquid level as the target output, a linear dynamic subsystem was constructed. The entire system was structured as a Hammerstein model. Applying the BC-RLS algorithm and the RLS algorithm to identify this three tank system, the simulation result is displayed in Fig. 6. The validation result of the identification model is given in Fig. 7, where the solid line is the measured output and the dashed line is the predicted output. It can be seen that the model output fits well with the measured output, indicating that the identification model can reflect the dynamic characteristics of the process and prove the effectiveness of the presented method.

6 Conclusions

The necessity of this study has been well justified with relevant research issues. Consequently, it has developed an interactive compensation-based procedure for the estimation of the parameters and states of the system descriptive model from input and noise

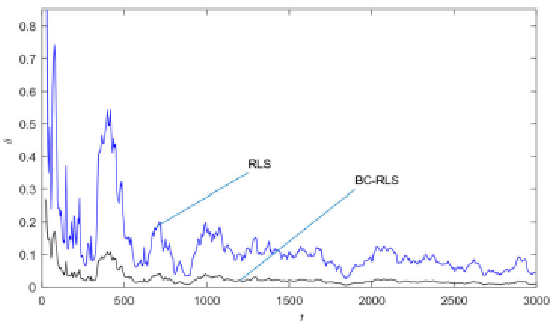

Fig. 1 *Parameter estimation errors δ versus t with $\sigma^2 = 1.00^2$*

Table 2 Parameter estimates and errors

σ^2	t	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	θ_8	θ_9	θ_{10}	θ_{11}	δ%
0.10^2	100	0.16820	0.23286	0.20131	0.10873	0.63414	0.84759	2.51580	4.85989	-0.01254	0.13945	0.12257	4.88064
	200	0.15739	0.23406	0.20003	0.12220	0.56086	0.90025	2.50998	4.93119	0.02649	0.12156	0.11783	3.11804
	500	0.18939	0.21148	0.19012	0.10204	0.53059	0.93883	2.49241	4.97458	0.03688	0.11644	0.11346	2.20195
	1000	0.19120	0.21016	0.19007	0.09857	0.51011	0.97640	2.49812	4.98643	0.04669	0.10557	0.10793	1.82653
	2000	0.19308	0.20849	0.18968	0.09720	0.50217	0.98568	2.50119	4.99251	0.05051	0.10425	0.10381	1.77662
	3000	0.19262	0.20863	0.18979	0.09755	0.50079	0.99011	2.50039	4.99541	0.05089	0.10265	0.10386	1.76249
1.00^2	100	0.09970	0.29621	0.21272	0.16476	0.60702	1.12974	2.51363	4.70900	0.04721	-0.07817	0.11751	7.32544
	200	0.01317	0.34109	0.22453	0.23365	0.61154	0.84719	2.52238	4.86943	0.04799	0.09310	0.11231	6.33194
	500	0.16739	0.22915	0.18822	0.13397	0.63417	0.75695	2.40858	4.93747	0.00207	0.14538	0.11025	5.69091
	1000	0.19109	0.21569	0.18949	0.09499	0.52430	0.94984	2.47511	4.95838	0.05112	0.09468	0.10234	2.18209
	2000	0.21377	0.19814	0.18657	0.07773	0.48632	0.95326	2.50924	4.97018	0.07091	0.10961	0.10301	2.14980
	3000	0.21082	0.19866	0.18795	0.07985	0.48241	0.96714	2.50170	4.98449	0.07025	0.10362	0.10396	1.99755
true values		0.19000	0.21000	0.19000	0.10000	0.50000	1.00000	2.50000	5.00000	0.05000	0.10000	0.10000	

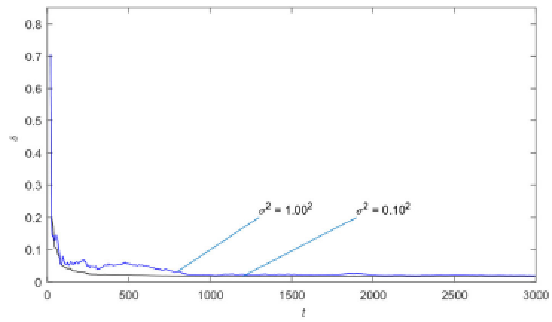

Fig. 2 *Parameter estimation errors δ versus t with $\sigma^2 = 0.10^2$ and $\sigma^2 = 1.00^2$*

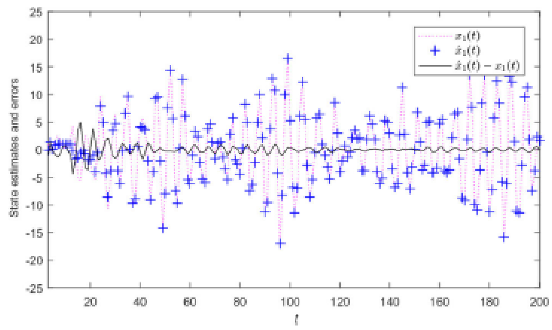

Fig. 3 *State and state estimate $\hat{x}_1(t)$ versus t*

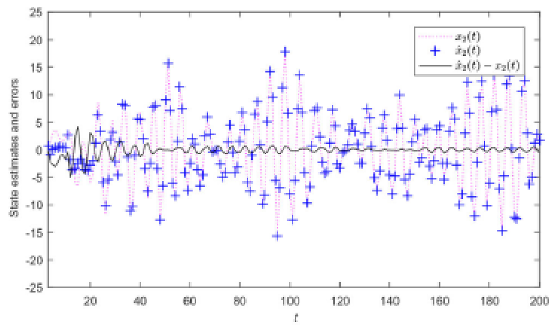

Fig. 4 *State and state estimate $\hat{x}_2(t)$ versus t*

contaminated output data sequences. The proof of the estimation convergence provides analytical assurance for applications. Two

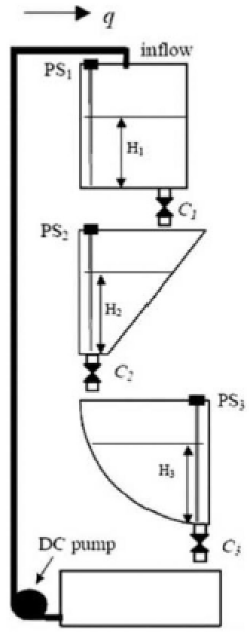

Fig. 5 *Experimental apparatus*

case studies show the effectiveness of the proposed procedure/algorithms. This study has potential impact to various real-world applications and academic research, which surely need more thorough bench tests and analytical development. The proposed methods in this paper can combine different tools and techniques and the identification principle [52–57] to study new parameter estimation algorithms of different stochastic systems with coloored noises and can be applied to other literatures [58–65] such as signal modelling and communication networked systems.

7 Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 61903050) and the Natural Science Foundation of Jiangsu Province (No. BK20181033). The authors are grateful to Professor Feng Ding for his helpful and constructive suggestions.

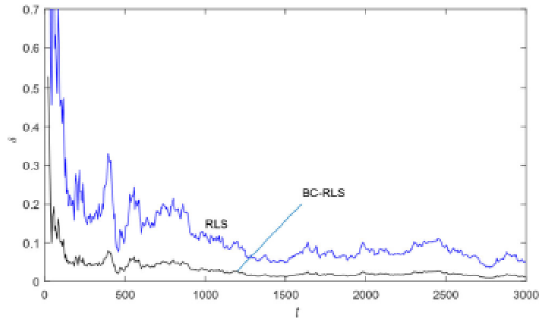

Fig. 6 *Parameter estimation errors δ versus t with* $\sigma^2 = 1.00^2$

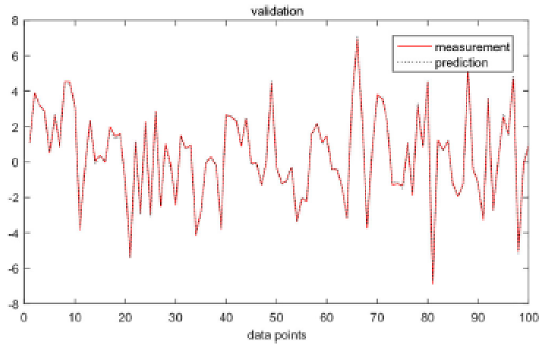

Fig. 7 *Validation result*

8 References

[1] Ding, F., Lv, L., Pan, J., *et al.*: 'Two-stage gradient-based iterative estimation methods for controlled autoregressive systems using the measurement data', *Int. J. Control Autom.*, 2020, **18**, (4), pp. 886–896

[2] Pan, J., Jiang, X., Wan, X.K., *et al.*: 'A filtering based multi-innovation extended stochastic gradient algorithm for multivariable control systems', *Int. J. Control Autom. Syst.*, 2017, **15**, (3), pp. 1189–1197

[3] Pan, J., Ma, H., Zhang, X., *et al.*: 'Recursive coupled projection algorithms for multivariable output-error-like systems with coloured noises', *IET Signal Process.*, 2020, **14**, (7), pp. 455–466

[4] Zhang, X., Ding, F., Yang, E.F.: 'State estimation for bilinear systems through minimizing the covariance matrix of the state estimation errors', *Int. J. Adapt. Control*, 2019, **33**, (7), pp. 1157–1173

[5] Zhang, X., Ding, F.: 'Hierarchical parameter and state estimation for bilinear systems', *Int. J. Syst. Sci.*, 2020, **51**, (2), pp. 275–290

[6] Ding, F., Xu, L., Meng, D.D., *et al.*: 'Gradient estimation algorithms for the parameter identification of bilinear systems using the auxiliary model', *J. Comput. Appl. Math.*, 2020, **369**, p. 112575

[7] Ibrir, S.: 'Joint state and parameter estimation of non-linearly parameterized discrete-time nonlinear systems', *Automatica*, 2018, **97**, pp. 226–233

[8] Yang, X.Q., Xiong, W.L., Wang, Z.Y., *et al.*: 'Parameter identification of nonlinear multirate time-delay system with uncertain output delays', *Trans. Inst. Meas. Control*, 2018, **40**, (12), pp. 3498–3506

[9] Ding, F.: 'Hierarchical multi-innovation stochastic gradient algorithm for Hammerstein nonlinear system modeling', *Appl. Math. Model.*, 2013, **37**, (4), pp. 1694–1704

[10] Beikzadeh, H., Liu, G.J., Marquez, H.J.: 'Robust sensitive fault detection and estimation for single-rate and multirate nonlinear sampled-data systems', *Syst. Control Lett.*, 2018, **119**, pp. 71–80

[11] Zhang, Y., Chai, T., Wang, D., *et al.*: 'Virtual unmodeled dynamics modeling for nonlinear multivariable adaptive control with decoupling design', *IEEE Trans. Syst. Man Cybern. Syst.*, 2018, **48**, (3), pp. 342–353

[12] Ding, F., Liu, X.G., Chu, J.: 'Gradient-based and least-squares-based iterative algorithms for Hammerstein systems using the hierarchical identification principle', *IET Control Theory Applic.*, 2013, **7**, (2), pp. 176–184

[13] Wang, D.Q., Zhang, S., Gan, M., *et al.*: 'A novel EM identification method for Hammerstein systems with missing output data', *IEEE Trans. Ind. Inf.*, 2019, **16**, (4), pp. 2500–2508, to be published, doi: 10.1109/TII.2019.2931792

[14] Chen, G.Y., Gan, M., Chen, C.L.P., *et al.*: 'A regularized variable projection algorithm for separable nonlinear least squares problems', *IEEE Trans. Autom. Control*, 2019, **64**, (2), pp. 526–537

[15] Chen, J., Liu, Y., Ding, F., *et al.*: 'Gradient-based particle filter algorithm for an ARX model with nonlinear communication output', *IEEE Trans. Syst. Man Cybern. Syst.*, 2020, **50**, (6), pp. 2198–2207

[16] Zeng, H.B., Liu, X.G., Wang, W.: 'A generalized free-matrix-based integral inequality for stability analysis of time-varying delay systems', *Appl. Math. Comput.*, 2019, **354**, pp. 1–8

[17] Dong, C.Y., Yang, S.F., Jia, H.J., *et al.*: 'Pade-based stability analysis for a modular multilevel converter considering the time delay in the digital control system', *IEEE Trans. Ind. Electron.*, 2019, **66**, (7), pp. 5242–5253

[18] Sanz, R., Garcia, P., Krstic, M.: 'Observation and stabilization of LTV systems with time-varying measurement delay', *Automatica*, 2019, **103**, pp. 573–579

[19] Javier, B.F., Zheng, G.: 'Unknown input functional observability of descriptor systems with neutral and distributed delay effects', *Automatica*, 2017, **85**, pp. 186–192

[20] Chen, J., Huang, B., Ding, F., *et al.*: 'Variational Bayesian approach for ARX systems with missing observations and varying time-delays', *Automatica*, 2018, **94**, pp. 194–204

[21] Xu, L., Chen, L., Xiong, W.L.: 'Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration', *Nonlinear Dyn.*, 2015, **79**, (3), pp. 2155–2163

[22] Xu, L.: 'The damping iterative parameter identification method for dynamical systems based on the sine signal measurement', *Signal Process.*, 2016, **120**, pp. 660–667

[23] Xu, L, Xiong, W.L., Alsaedi, A., *et al.*: 'Hierarchical parameter estimation for the frequency response based on the dynamical window data', *Int. J. Control Autom. Syst.*, 2018, **16**, (4), pp. 1756–1764

[24] Gu, Y., Chou, Y.X., Liu, J.C., *et al.*: 'Moving horizon estimation for multirate systems with time-varying time-delays', *J. Franklin. Inst.*, 2019, **356**, (4), pp. 2325–2345

[25] Ansari, A., Bernstein, D.S.: 'Deadbeat unknown-input state estimation and input reconstruction for linear discrete-time systems', *Automatica*, 2019, **103**, pp. 11–19

[26] Yan, L.P., Jiang, L., Liu, J., *et al.*: 'Optimal distributed Kalman filtering fusion for multirate multisensor dynamic systems with correlated noise and unreliable measurements', *IET Signal Process.*, 2018, **12**, (4), pp. 522–531

[27] Miguez, J., Marino, I.P., Vazquez, M.A.: 'Analysis of a nonlinear importance sampling scheme for Bayesian parameter estimation in state-space models', *Signal Process.*, 2019, **142**, pp. 281–291

[28] Xie, L., Yang, H.Z., Huang, B.: 'FIR model identification of multirate processes with random delays using EM algorithm', *AIChE J.*, 2013, **59**, (11), pp. 4124–4132

[29] Gupta, N., Maqbul, M.: 'Solutions to Rayleigh-love equation with constant coefficients and delay forcing term', *Appl. Math. Comput.*, 2019, **355**, pp. 123–134

[30] Zou, L., Wang, Z.D., Han, Q.L., *et al.*: 'Moving horizon estimation for networked time-delay systems under round-robin protocol', *IEEE Trans. Autom. Control*, 2019, **64**, (12), pp. 5191–5198

[31] Zou, L., Wang, Z.D., Han, Q.L., *et al.*: 'Moving horizon estimation of networked nonlinear systems with random access protocol', *IEEE Trans. Syst. Man Cybern. Syst.*, 2020, doi: 10.1109/TSMC.2019.2918002

[32] Zou, L., Wang, Z.D., Hu, J., *et al.*: 'Moving horizon estimation with unknown inputs under dynamic quantization effects', *IEEE Trans. Autom. Control*, 2020, doi: 10.1109/TAC.2020.2968975

[33] Gu, Y., Ding, F., Li, J.H.: 'State filtering and parameter estimation for linear systems with d-step state-delay', *IET Signal Process.*, 2014, **8**, (6), pp. 639–646

[34] Ding, F.: 'State filtering and parameter estimation for state space systems with scarce measurements', *Signal Process.*, 2014, **104**, pp. 369–380

[35] Ji, Y., Jiang, X.K., Wan, L.J.: 'Hierarchical least squares parameter estimation algorithm for two-input Hammerstein finite impulse response systems', *J. Franklin. Inst.*, 2020, **357**, (8), pp. 5019–5032

[36] Li, M.H., Liu, X.M.: 'Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique', *Int. J. Control Autom. Syst.*, 2020, **18**, (6), pp. 1581–1592

[37] Ding, F.: 'Combined state and least squares parameter estimation algorithms for dynamic systems', *Appl. Math. Model*, 2014, **38**, (1), pp. 403–412

[38] Wan, L.J., Ding, F.: 'Decomposition- and gradient-based iterative identification algorithms for multivariable systems using the multi-innovation theory', *Circuits, Syst. Signal Process.*, 2019, **38**, (7), pp. 2971–2991

[39] Ji, Y.M., Zhang, C., Kang, Z., *et al.*: 'Parameter estimation for block-oriented nonlinear systems using the key term separation', *Int. J. Robust Nonlinear*, 2020, **30**, pp. 3727–3752

[40] Ren, H.P., Fan, J.T., Kaynak, O.: 'Optimal design of a fractional-order proportional-integer-differential controller for a pneumatic position servo system', *IEEE Trans. Ind. Electron.*, 2019, **66**, (8), pp. 6220–6229

[41] Gan, M., Li, H.X., Peng, H.: 'A variable projection approach for efficient estimation of RBF-ARX model', *IEEE Trans. Cybern.*, 2015, **45**, (3), pp. 462–471

[42] Ma, F.Y., Yin, Y.K., Li, M.: 'Start-up process modelling of sediment microbial fuel cells based on data driven', *Math. Probl. Eng.*, 2019, **2019**, doi: 10.1155/2019/7403732

[43] Wang, Y.J.: 'Filtering-based iterative identification for multivariable systems', *IET Control Theory Appl.*, 2016, **10**, (8), pp. 894–902

[44] Ma, F.Y., Yin, Y.K., Pang, S.P., *et al.*: 'A data-driven based framework of model optimization and neural network modeling for microbial fuel cells', *IEEE Access*, 2019, **7**, pp. 162036–162049

[45] Ma, F.Y., Fu, C.C., Yang, J., *et al.*: 'Control strategy for adaptive active energy harvesting in sediment microbial fuel cells', *J. Energy Eng.*, 2020, **146**, (1), Article Number: 04019034

[46] Chang, Y.F., Zhai, G.S., Fu, B., *et al.*: 'Quadratic stabilization of switched uncertain linear systems: a convex combination approach', *IEEE-CAA J. Autom. Sin.*, 2019, **6**, (5), pp. 1116–1126

[47] Chang, Y.F., Sun, C.J., Qiu, Y.: 'Effective notch stress method for fatigue assessment of sheet alloy material and bi-material welded joints', *Thin-Walled Struct.*, 2020, **151**, p. 106745

[48] Geng, L., Xiao, R.B.: 'Control and backbone identification for the resilient recovery of a supply network utilizing outer synchronization', *Appl. Sci.*, 2020, **10**, (1), p. 213, Article Number: 313

[49] Tang, L., Liu, G.J., Yang, M., *et al.*: 'Joint design and torque feedback experiment of rehabilitation robot', *Adv. Mech. Eng.*, 2020, **12**, pp. 1–11, doi: 10.1177/1687814020924498

[50] Wang, L.J., Feng, B.Y., Wang, Y., *et al.*: 'Bidirectional short-circuit current blocker for DC microgrid based on solid-state circuit breaker', *Electronics*, 2020, **9**, (2), p. 306

[51] Ding, F.: '*System identification - performances analysis for identification methods*' (Science Press, Beijing, 2014)

[52] Wu, T.Z., Ye, F.C., Su, Y.H., *et al.*: 'Coordinated control strategy of DC microgrid with hybrid energy storage system to smooth power output fluctuation', *Int. J. Low-Carbon Technol.*, 2020, **15**, (1), pp. 46–54

[53] Zhang, Y., Huang, M.M., Wu, T.Z., *et al.*: 'Reconfigurable equilibrium circuit with additional power supply', *Int. J. Low-Carbon Technol.*, 2020, **15**, (1), pp. 106–111

[54] Liu, Y.J., Ding, F., Shi, Y.: 'An efficient hierarchical identification method for general dual-rate sampled-data systems', *Automatica*, 2014, **50**, (3), pp. 962–970

[55] Wang, Y.J., Ding, F.: 'Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model', *Automatica*, 2016, **71**, pp. 308–313

[56] Gu, Y., Liu, J., Li, X., *et al.*: 'State space model identification of multirate processes with time-delay using the expectation maximization', *J. Franklin Inst.*, 2019, **356**, (3), pp. 1623–1639

[57] Ding, F., Zhang, X., Xu, L.: 'The innovation algorithms for multivariable state-space models', *Int. J. Adapt. Control.*, 2019, **33**, (11), pp. 1601–1608

[58] Wu, M.H., Chen, R., Tong, Y.: 'Shadow elimination algorithm using color and texture features', *Comput. Intel. Neurosci.*, 2020, **2020**, p. 2075781

[59] Zhao, N., Fan, P., Cheng, Y.: 'Dynamic contract incentives mechanism for traffic offloading in multi-UAV networks', *Wirel. Commun. Mob. Comput.*, 2020, **2020**, Article ID: 2361029, doi: 10.1155/2020/2361029

[60] Xu, G.H., Qi, F., Lai, Q., *et al.*: 'Fixed time synchronization control for bilateral teleoperation mobile manipulator with nonholonomic constraint and time delay', *IEEE Trans. Circuits-II*, 2020, doi: 10.1109/TCSII.2020.2990698

[61] Wang, L., Liu, H., Dai, L.V., *et al.*: 'Novel method for identifying fault location of mixed lines', *Energies*, 2018, **11**, (6), p. 1529. Article Number: 1529. https://doi.org/10.3390/en11061529

[62] Liu, H., Zou, Q.X., Zhang, Z.P.: 'Energy disaggregation of appliances consumptions using ham approach', *IEEE Access*, 2019, **7**, pp. 185977–185990

[63] Liu, J.C., Liu, G.Y., Chou, Y.X.: 'Seismic data reconstruction via complex shearlet transform and block coordinate relaxation', *J. Seism. Explor.*, 2019, **28**, (4), pp. 307–332

[64] Zhao, X.L., Lin, Z.Y., Fu, B., *et al.*: 'Research on automatic generation control with wind power participation based on predictive optimal 2-degree-of-freedom PID strategy for multi-area interconnected power system', *Energies*, 2018, **11**, (12), p. 3325, Article Number: 3325. http://dx.doi.org/10.3390/en11123325

[65] Ding, F., Wang, X.H., Chen, Q.J., *et al.*: 'Recursive least squares parameter estimation for a class of output nonlinear systems based on the model decomposition', *Circuits Syst. Signal Process.*, 2016, **35**, (9), pp. 3323–3338