Journal of Process Control, Volume 50, February 2017, Pages 40–55
10.1016/j.jprocont.2016.11.007

Identification of Discrete-time Output Error Model for Industrial Processes with Time Delay Subject to Load Disturbance

Shijian Dong [a], Tao Liu [a, *], Wei Wang [a], Jie Bao [b], Yi Cao [c]

a Institute of Advanced Control Technology, Dalian University of Technology, Dalian, 116024, P.R. China
b School of Chemical Sciences and Engineering, The University of New South Wales, UNSW 2052, Australia
c School of Energy, Environment and Agrifood, Cranfield University, Cranfield, MK43 0AL, UK

Abstract: In this paper, a bias-eliminated output error model identification method is proposed for industrial processes with time delay subject to unknown load disturbance with deterministic dynamics. By viewing the output response arising from such load disturbance as a dynamic parameter for estimation, a recursive least-squares identification algorithm is developed in the discrete-time domain to estimate the linear model parameters together with the load disturbance response, while the integer delay parameter is derived by using a one-dimensional searching approach to minimize the output fitting error. An auxiliary model is constructed to realize consistent estimation of the model parameters against stochastic noise. Moreover, dual adaptive forgetting factors are introduced with tuning guidelines to improve the convergence rates of estimating the model parameters and the load disturbance response, respectively. The convergence of model parameter estimation is analyzed with a rigorous proof. Illustrative examples for open- and closed-loop identification are shown to demonstrate the effectiveness and merit of the proposed identification method.

Key words: Output error model, time delay, load disturbance, convergence, forgetting factor

---

Published by Elsevier. This is the Author Accepted Manuscript issued with:Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 4.0).
The final published version (version of record) is available online at DOI:10.1016/j.jprocont.2016.11.007. Please refer to any applicable publisher terms of use.

1. Introduction

Unexpected or time-varying load disturbance is often encountered when performing identification tests for industrial processes [1-3], e.g. an inherent type load disturbance arising from the mold cavity pressure affects the injection velocity to fill up the mold cavity during the filling process of injection molding [4]. Unknown load disturbance may propagate throughout the process and blur the output response when performing an identification test, causing undesired identification error [2-5]. Bias-eliminated model identification against load disturbance has therefore become increasingly appealed in the recent years [6, 7]. Different identification tests have been explored to deal with load disturbance. To enumerate a few, for the use of step response test, a time integral identification method [8] was proposed to eliminate the influence from stochastic or static type load disturbance. A robust step response identification algorithm was developed for unbiased parameter estimation against unexpected deterministic type load disturbance [9], by using the transient response data obtained by adding and subsequently removing a step change to the process input. With a pulse type input excitation, a two-stage identification algorithm was developed to cope with a specific class of load disturbance with a continuous spectrum of amplitude [10]. The approach was further extended to deal with periodic disturbance by using multiple sinusoidal excitations [11]. For the use of a closed-loop relay feedback identification test subject to static type load disturbance, a frequency domain transfer function identification method was proposed based on using the symmetric property of output response to eliminate the influence from such load disturbance [12]. By comparison, a so-called A-locus analysis method was presented to identify integrating and unstable processes subject to static type load disturbance [13]. However, a prior knowledge of the occurrence of load disturbance or its dynamics is needed to apply these identification methods, which may not be available in engineering applications subject to unknown load disturbance.

In discrete-time domain, a few output error (OE) model identification algorithms were presented to deal with stationary stochastic load disturbance for both open-loop and closed-loop identification tasks in the literature [5, 14]. In cases where both the system input and output suffer from colored noises, errors-in-variables methods were developed to procure consistent estimation [15, 16]. In the presence of unknown but bounded disturbance, an error-bounded parameter estimation algorithm was

proposed based on using a membership set [17]. A refined instrumental variable (RIV) approach was recently developed by using a unified operator to estimate the Box–Jenkins model [18]. For the presence of non-stationary disturbance, a bias compensation identification algorithm [19] was given to obtain an extended ARMAX or OE model with good accuracy, by using a variable forgetting factor to estimate the model parameters and disturbance. For time delay systems, only a few papers presented discrete-time domain identification methods for obtaining an OE model with an integer type delay parameter, due to the difficulties for identifying the linear model parameters together with an integer type delay parameter that involves mixed-integer programming, which was recognized to be a non-convex problem for parameter estimation [20-22].

In this paper, to deal with unknown load disturbance having deterministic dynamics, a discrete-time domain OE model identification method is proposed for industrial processes with time delay response, to facilitate computer-aided control design for implementation. The output response to such load disturbance is viewed as a dynamic parameter which is lumped into the model parameters for estimation. To solve the above non-convex problem for parameter estimation, a one-dimensional searching strategy is given by minimizing the output prediction error to determine the optimal integer type delay parameter. An auxiliary model is constructed to realize consistent estimation of the model parameters against stochastic noise, in consideration of that the standard recursive least-squares (RLS) algorithm cannot guarantee consistent estimation of an OE model [23]. Moreover, dual forgetting factors are introduced to expedite the convergence rates of estimating the model parameters and the load disturbance response, respectively, with tuning guidelines to avoid the 'wind-up' of estimation error arising from using a constant forgetting factor as adopted in a traditional RLS algorithm, in particular for using a poor excitation for identification [24-27]. The convergence and unbiased estimation of the proposed algorithm is analyzed with a rigorous proof. The paper is organized as follows. In Section 2, the identification problem is presented. In Section 3, the proposed identification algorithm is detailed, followed by convergence analysis given in Section 4. Two illustrative examples for open- and closed-loop identification together with an application to identify an injection molding process model under time-varying load disturbance are presented in Section 5. Finally, conclusions are drawn in Section 6.

Throughout the paper, the following notations are used. Denote by $\Re$, $\Re^n$, and $\Re^{n\times m}$ the spaces of real number, n-dimensional real vector, and $n\times m$ real matrix, respectively. For any matrix $P\in\Re^{m\times m}$, $P>0$ (or $P\geq 0$) means P is a positive (or semi-positive) definite matrix. For $P\in\Re^{m\times m}$ of full rank, denote by P^{-1} the inverse of P, by P^T the transpose of P, and by $\mathrm{tr}(P)$ the matrix trace. Denote by $\|P\|_2$ the Euclidean norm of $P\in\Re^n$. Denote by $\rho(P)$ the eigenvalue of P, and by $\rho_{\min}(P)$ and $\rho_{\max}(P)$ the minimum and maximum, respectively. The identity/zero vector or matrix with appropriate dimensions is denoted by $\boldsymbol{I}/\boldsymbol{0}$, where $\boldsymbol{I}_m$ indicates $\boldsymbol{I}_m\in\Re^{m\times m}$ and $\boldsymbol{0}_{m\times n}$ for $\boldsymbol{0}_{m\times n}\in\Re^{m\times n}$. Denote by $E[g]$ the mathematical expectation operator with respect to $g\in\Re^n$. Denote by $\hat{\alpha}$ the estimated value of $\alpha\in\Re^m$. Denote by z a discrete-time operator, i.e. $z^{-1}u(t):=u(t-1)$.

2. Problem description

When performing an identification test for a sampled system with time delay subject to load disturbance and measurement noise, the output response can be generally described by the following discrete-time OE model with an integer type delay parameter,

$$\begin{cases} x(t)=\dfrac{B(z^{-1})}{A(z^{-1})}z^{-d}u(t) \\ y(t)=x(t)+\xi(t)+v(t) \end{cases} \tag{1}$$

where d is an integer type delay. The polynomials $A(z^{-1})$ and $B(z^{-1})$ are coprime with the following forms,

$$A(z^{-1})=1+a_1z^{-1}+\cdots+a_{n_a}z^{-n_a}$$
$$B(z^{-1})=b_1z^{-1}+\cdots+b_{n_b}z^{-n_b}$$

All zeros of $A(z^{-1})$ are assumed to be inside the unity circle, i.e. the system is asymptotically stable. Denote by $u(t)$, $x(t)$ and $\xi(t)$ the excitation signal, noise-free response, and load disturbance response, respectively. The output measurement noise is denoted by $v(t)$, which is usually assumed to be a Gaussian white noise with zero mean and unknown variance denoted by σ_v^2.

Generally, it is assumed that the system is causal, i.e. $y(t)$ depends on $u(s)$ for $s\leq t$, but not on future values of $u(t)$ and $v(t)$, while $v(t)$ is uncorrelated with the input sequence, $u(t)$. Assume that $u(t)=0$, $y(t)=0$, and $v(t)=0$ for $t\leq 0$, indicating the initial zero/steady state for identification.

For identifiablity, $\xi(t)$ resulting from load disturbance having deterministic dynamics is considered herein. Stochastic load disturbance response is lumped into $v(t)$ to treat with.

The identification objective is to estimate the above OE model parameters including an integer delay from sampled data, with a prior knowledge or assumption on the orders of n_a and n_b for model fitting. For unknown system dynamics, the optimal model order may be determined by using the Akaike information criterion (AIC), a hypothesis testing condition [14], or a cross-correlation function between the input and the univariate residual sequence [19], so as to check if a higher order model could result in better fitting in terms of the parsimony principle on the number of model parameters.

3. Proposed identification method

3.1 Linear model parameter estimation

Denote the linear model parameter vector and plant information vector, respectively, by

$$\theta_0 = [a_1, \cdots, a_{n_a}, b_1, \cdots, b_{n_b}]^T \in \Re^{n_0} \tag{2}$$

$$\varphi_0(t) = [-x(t-1), \cdots, -x(t-n_a), u(t-1-d), \cdots, u(t-n_b-d)]^T \in \Re^{n_0} \tag{3}$$

where $n_0 = n_a + n_b$.

The noise-free output response can be written by

$$x(t) = \varphi_0^{\ T}(t)\theta_0 \tag{4}$$

For the convenience of identification, the deterministic load disturbance response, $\xi(t)$, is considered as a dynamic parameter to be estimated. We define an augmented parameter vector together with the corresponding information vector, respectively, by

$$\theta(t) = [a_1, \cdots, a_{n_a}, b_1, \cdots, b_{n_b}, \xi(t)]^T \in \Re^{n_{\mathrm{m}}} \tag{5}$$

$$\varphi(t) = [-x(t-1), \cdots, -x(t-n_a), u(t-1-d), \cdots, u(t-n_b-d), 1]^T \in \Re^{n_{\mathrm{m}}} \tag{6}$$

where $n_{\mathrm{m}} = n_a + n_b + 1$.

Hence, the plant description in (1) can be rewritten as a linear regression form,

$$y(t) = \varphi^T(t)\theta(t) + v(t) \tag{7}$$

Note that a prediction of $y(t)$ can be taken as

$$\hat{y}(t) = \hat{\varphi}^T(t)\hat{\theta}(t) \tag{8}$$

where $\hat{\theta}(t)$ and $\hat{\varphi}(t)$ denote the estimated parameter vector and information vector, respectively.

Accordingly, the prediction error is computed by

$$e(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t) \tag{9}$$

Define the following cost function of prediction error with a forgetting factor for minimization,

$$J(t,\hat{\theta}) = \frac{1}{2}\sum_{i=1}^{t} \lambda^{t-i} e(i)^2 \tag{10}$$

where $\lambda \in (0,1]$ is a forgetting factor used for improving the tracking performance.

Taking the first derivative of $J(t,\hat{\theta})$ with respect to $\hat{\theta}$, we have

$$\frac{\partial J(t,\hat{\theta})}{\partial \hat{\theta}} = -\sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i)[y(i) - \hat{\varphi}^T(i)\hat{\theta}(i)] \tag{11}$$

By letting (11) be zero, we obtain

$$\hat{\theta}(t) = \left[\sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i)\right]^{-1} \times \left[\sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i) y(i)\right] \tag{12}$$

Let

$$P(t) = \left[\sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i)\right]^{-1} \tag{13}$$

Given a persistent excitation (PE) condition of $PE(u) \geq n_{\mathrm{m}}$, it can be easily verified that $P(t)$ keeps nonsingular and its inverse can be derived as

$$\begin{aligned} P^{-1}(t) &= \sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) \\ &= \lambda P^{-1}(t-1) + \hat{\varphi}(t)\hat{\varphi}^T(t) \end{aligned} \tag{14}$$

It follows from (12) that

$$\hat{\theta}(t) = P(t)\sum_{i=1}^{t} \lambda^{t-i} \hat{\varphi}(i) y(i) \tag{15}$$

which can be reformulated in a recursive form as below

$$\begin{aligned} \hat{\theta}(t) &= P(t)[\lambda P^{-1}(t-1)\hat{\theta}(t-1) + \hat{\varphi}(t) y(t)] \\ &= P(t)[(P^{-1}(t) - \hat{\varphi}(t)\hat{\varphi}^T(t))\hat{\theta}(t-1) + \hat{\varphi}(t) y(t)] \\ &= \hat{\theta}(t-1) + P(t)\hat{\varphi}(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] \\ &= \hat{\theta}(t-1) + P(t)\hat{\varphi}(t) e_1(t) \end{aligned} \tag{16}$$

where $e_1(t) = y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)$.

To relieve the computation effort on matrix inversion involved with (14) for a recursive estimation of $\hat{\theta}$, we use the following fact of matrix inversion for any matrices, A, B, C, D, with appropriate dimensions [14],

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1} \tag{17}$$

Denoting $A = \lambda P^{-1}(t-1)$, $B = \hat{\varphi}(t)$, $C = \boldsymbol{I}_{n_m \times n_m}$ and $D = \hat{\varphi}^T(t)$ to apply (17) for (14), a recursive computation law for $P(t)$ is obtained as

$$P(t)=\frac{1}{\lambda}\left(P(t-1)-\frac{P(t-1)\hat{\varphi}(t)\hat{\varphi}^T(t)P(t-1)}{\lambda \boldsymbol{I}_{\mathrm{n_m}\times\mathrm{n_m}}+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)}\right) \tag{18}$$

Let

$$K(t)=P(t)\hat{\varphi}(t) \tag{19}$$

It can be derived from (18) that

$$K(t)=\frac{P(t-1)\hat{\varphi}(t)}{\lambda+\hat{\varphi}^T(t)P(t-1)\hat{\varphi}(t)} \tag{20}$$

Substituting (20) into (18) yields

$$P(t)=\frac{1}{\lambda}P(t-1)[\boldsymbol{I}_{\mathrm{n_m}\times\mathrm{n_m}}-K(t)\hat{\varphi}^T(t)] \tag{21}$$

Accordingly, it follows from (16) that

$$\hat{\theta}(t)=\hat{\theta}(t-1)+K(t)e_1(t) \tag{22}$$

Since the information vector $\hat{\varphi}(t)$ contains the unknown inner variables $x(t-i)$ for $i=1,2,\cdots,n_a$, the above parameter estimation cannot be implemented. To solve the problem, an auxiliary model for predicting $x(t)$ is therefore proposed based on the previously estimated result of $\hat{\theta}(t-1)$, i.e.

$$\hat{x}(t)=\hat{\varphi}_0{}^T(t)\hat{\theta}_0(t-1) \tag{23}$$

where $\hat{\varphi}_0(t)$ is an estimation of $\varphi_0(t)$ in (3), and $\hat{\theta}_0(t-1)$ is an estimation of $\theta_0(t-1)$ in (2) at the $(t-1)$-th step for recursion.

3.2 Delay estimation

It is obvious that the prediction error in (9) cannot be computed if the delay parameter is unknown, and therefore, the other model parameters cannot be estimated by using the above recursive algorithm without an estimated delay. In fact, the cost function $J(t,\hat{\theta})$ in (10) is a nonconvex function with respect to the delay parameter. For illustration, consider the following noise-free system,

$$y(t)=\frac{0.5z^{-1}+0.4z^{-2}}{1+\alpha z^{-1}+0.6z^{-2}}z^{-d}u(t) \tag{24}$$

where $\alpha=0.8$ and $d=25$. The input excitation is taken as a Guassian white noise for model identification. By taking $d\in[5,\ 50]$ and $\alpha\in[-1,\ 2]$ to represent identified models by using any identification algorithm, the corresponding J of prediction error is plotted in Fig.1.

It is seen that there exists many local minima for the cost function $J(t,\hat{\theta})$ with respect to an integer type time delay for parameter estimation, relating to a mixed integer programming problem

as studied in [20-22]. Note that if using the existing continuous-time identification algorithms (e.g. [7, 8, 22]) to simultaneously estimate all the linear model parameters together with the delay parameter, fractional delays will be turned out during the recursive procedure for parameter estimation, and thus the convergence of such a recursive algorithm cannot be guaranteed, especially for the presence of stochastic noise or load disturbance [22]. Although cross-correlation analysis [14, 19] on the input and output sequences may be adopted to separately estimate the time delay, the approach may give incorrect estimation because the variance of cross-correlation estimation could be blurred by colored noise or load disturbance.

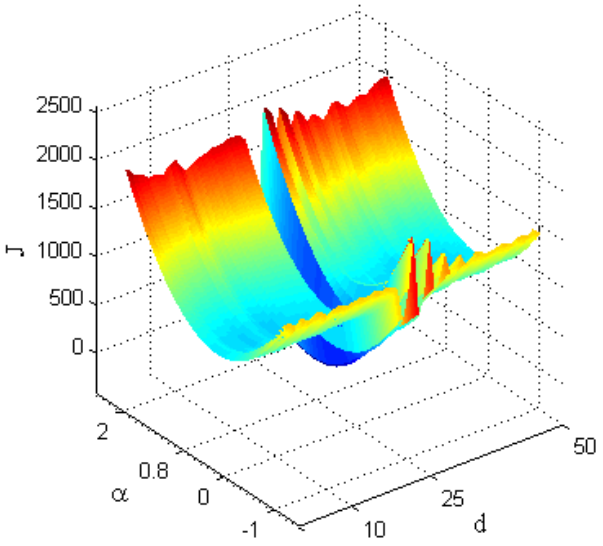


Fig. 1. Plot of the cost function with respect to α and d

Considering that a possible range of the time delay (i.e. $d \in [d_{\min},\ d_{\max}]$) can be determined based on a prior knowledge of system operation or experimental observation such as from a step response test, a one-dimensional searching of integer delay is preferred to implement the above recursive estimation algorithm, i.e.

$$\hat{d}(t) = \min J_i(t, \hat{\theta}_i(t), \hat{d}_i) \tag{25}$$

which means that $\hat{d}(t)$ is determined by minimizing J through a searching range of $\hat{d}_i \in [d_{\min},\ d_{\max}]$, and meanwhile, the remaining linear model parameters are estimated. Initially, it may be taken as $d_{\min}$ or $d_{\max}$. Owing to the limited integer range of $d \in [d_{\min},\ d_{\max}]$, the proposed one-dimensional searching algorithm can be easily executed without heavy computation effort, such that the mixed integer programming problem together with the computation issue of fractional delay can be definitely avoided.

3.3 Recursive identification with dual forgetting factors

It can be seen from (14) and (16) that a constant forgetting factor should be adopted in the standard RLS method, i.e.

$$\lambda P^{-1}(t-1)=\begin{bmatrix} \lambda P_{11}^{-1}(t-1),\cdots,\lambda P_{1n_m}^{-1}(t-1) \\ \vdots \qquad \ddots \qquad \vdots \\ \lambda P_{n_m1}^{-1}(t-1),\cdots,\lambda P_{n_mn_m}^{-1}(t-1) \end{bmatrix} \tag{26}$$

Considering that the dynamic parameter reflecting load disturbance response is time-variant while the true model parameters are time-invariant, it is desirable to forget past values of $\xi(t)$ with a fast rate and adjust the convergency rate for estimating the model parameters θ_0 in terms of fitting accuracy. For this purpose, two adaptive forgetting factors (AFFs) are proposed for estimating θ_0 and $\xi(t)$, respectively, as

$$\lambda_1(t)=1/(1+\left\|\hat{\theta}_0(t)-\hat{\theta}_0(t-1)\right\|_2^2),\quad \lambda_1(t)\geq\lambda_{1\min} \tag{27}$$

$$\lambda_2(t)=1/\exp(\left|y(t)-\hat{\varphi}^T(t)\hat{\theta}(t)\right|),\quad \lambda_2(t)\geq\lambda_{2\min} \tag{28}$$

where $\lambda_{1\min}\in[0.95,1)$ and $\lambda_{2\min}\in[0.8,0.95]$ are the lower limit to avoid over sensitivity in the presence of the output measurement outliers. The key idea behind using these AFFs is to set a smaller forgetting factor to obtain a faster tracking rate when the prediction error is larger, and when the estimated model parameters converge to the true values, these AFFs will become a larger value closed to the unity to maintain fitting stability and accuracy against stochastic noise or disturbance.

Correspondingly the AFF matrix is constructed as

$$\boldsymbol{\lambda}(t)=\begin{bmatrix} \lambda_1(t),\cdots,\lambda_1(t),\lambda_2(t) \\ \vdots \quad \ddots \quad \vdots \qquad \vdots \\ \lambda_1(t),\cdots,\lambda_1(t),\lambda_2(t) \\ \underbrace{\lambda_2(t),\cdots,\lambda_2(t)}_{n_0},\lambda_2(t) \end{bmatrix} \tag{29}$$

Denote the Hadamard product of $\boldsymbol{\lambda}(t)$ and $P^{-1}(t-1)$ by

$$H_\lambda(t)=\boldsymbol{\lambda}(t)\circ P^{-1}(t-1) \tag{30}$$

An updating law for $P^{-1}(t)$ is therefore obtained by substituting (30) into (14) as

$$P^{-1}(t)=H_\lambda(t)+\hat{\varphi}(t)\hat{\varphi}^T(t) \tag{31}$$

Note that the two AFFs should satisfy $\lambda_1(t)\geq\lambda_2(t)$ to ensure the positive definiteness of $P(t)$ and $H_\lambda(t)$ for recursive computation [14, 28].

By using the matrix inverse lemma in (17), it follows that

$$P(t) = [\boldsymbol{I}_{\mathrm{n_m}\times\mathrm{n_m}} - K(t)\hat{\varphi}^T(t)]H_\lambda^{-1}(t) \tag{32}$$

where the gain matrix is computed by

$$K(t) = \frac{H_\lambda^{-1}(t)\hat{\varphi}(t)}{1+\hat{\varphi}^T(t)H_\lambda^{-1}(t)\hat{\varphi}(t)} \tag{33}$$

Hence, θ is estimated by

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] \tag{34}$$

To sum up, the proposed identification algorithm is given below.

(i) Specify the initial values $\hat{\theta}(0)$, $P(0)$, $\lambda_1(0)$, $\lambda_2(0)$, $\lambda_{1\min}$, $\lambda_{2\min}$, $d_{\min}$, $d_{\max}$ for recursion, e.g. $\hat{\theta}(0) = 10^{-3} I_{n_m\times 1}$, $P(0) = 10^6 I_{n_m\times n_m}$, $\lambda_1(0) = 1$, $\lambda_2(0) = 1$, $\lambda_{1\min} = 0.99$, and $\lambda_{2\min} = 0.9$.

(ii) Construct the information vectors $\hat{\varphi}_0(t)$ and $\hat{\varphi}(t)$ in terms of (3) and (6) by using the auxiliary model in (23).

(iii) Choose the two AFFs satisfying $\lambda_1(t) \geq \lambda_2(t)$ to construct the AFF matrix $\boldsymbol{\lambda}(t)$ in (29).

(iv) Compute $\hat{\theta}_i(t)$ by (30), (32)-(34), and the cost function $J_i(t, \hat{\theta}_i(t), \hat{d}_i)$ by (10).

(v) Increase $\hat{d}_i$ by one and return to the above step (ii) for computation, until $\hat{d}_i = d_{\max}$.

(vi) For the smallest $J_i(t, \hat{\theta}_i(t), \hat{d}_i)$ computed from the above step (v), take the correspondingly estimated parameters as $\hat{\theta}(t)$ and $\hat{d}(t)$ at the *t*-th step for recursion.

(vii) Increase t by one and return to the above step (ii), until a specified convergence condition is satisfied, i.e. $\left\|\hat{\theta}(t) - \hat{\theta}(t-1)\right\|_2^2 \leq \varepsilon$, where ε is a user-specified threshold, or the end of the measured data length ($t = N$).

---

4. Convergence analysis

Before the convergence analysis, a preliminary lemma is firstly presented as below.

Lemma 1 *[29]. For any matrices,* $A \in \Re^{m\times m}$ *and* $B \in \Re^{m\times m}$ *while* $B \geq 0$*, and a positive scalar* $\gamma < \rho_{\min}(A)$*, there stand*

$$\rho_{\min}(A)\boldsymbol{I}_m \leq A^T A \leq \rho_{\max}(A)\boldsymbol{I}_m \tag{35}$$

$$(A+\gamma\boldsymbol{I}_m)^T(A+\gamma\boldsymbol{I}_m) \geq [\rho_{\min}(A)-\gamma]^2\boldsymbol{I}_m \tag{36}$$

$$\rho_{\min}(A)\mathrm{tr}(B) \leq \mathrm{tr}(AB) \leq \rho_{\max}(A)\mathrm{tr}(B) \tag{37}$$

Then the following corollary is given for analysis on the persistent excitation (PE) condition for

model identification.

Corollary 1. *For a sampled system described by (7), if there exist* α *and* β *satisfying* $0 < \alpha \le \beta < \infty$, *such that the following PE condition with respect to the memory data length* N_λ *relating to the use of a forgetting factor for recursion holds*

$$\alpha \boldsymbol{I} \le \frac{1}{N_\lambda} \sum_{j=0}^{N_\lambda - 1} \hat{\varphi}(t+j)\hat{\varphi}^T(t+j) \le \beta \boldsymbol{I} \tag{38}$$

where $N_\lambda > n_m$, *and* $t > N_\lambda$, *then the covariance matrix* $P(t)$ *satisfies*

$$\frac{1-\lambda_{\mathrm{M}}}{\beta N_\lambda} \boldsymbol{I} \le \lim_{t\to\infty} P(t) \le \frac{1-\lambda_{\mathrm{m}}}{\alpha \lambda_{\mathrm{m}}^{N_\lambda - 1}} \boldsymbol{I} \tag{39}$$

where λ_{m} *and* λ_{M} *satisfy* $\lambda_{\mathrm{m}} P^{-1}(i) \le \boldsymbol{\lambda}(i) \circ P^{-1}(i) \le \lambda_{\mathrm{M}} P^{-1}(i)$.

Proof : Taking into account on the PE condition in (38), it follows from (30) and (31) that

$$\begin{aligned}
P^{-1}(t) &= \boldsymbol{\lambda}(t) \circ P^{-1}(t-1) + \hat{\varphi}(t)\hat{\varphi}^T(t) \\
&\ge \lambda_{\mathrm{m}} P^{-1}(t-1) + \hat{\varphi}(t)\hat{\varphi}^T(t) \\
&= \frac{1}{N_\lambda}\left(N_\lambda \sum_{i=1}^{t} \lambda_{\mathrm{m}}^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) \right) + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&\ge \frac{1}{N_\lambda}\left(\sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) + \cdots + \sum_{i=N_\lambda}^{t} \lambda_{\mathrm{m}}^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) \right) + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&= \frac{1}{N_\lambda}\left(\sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) + \cdots + \sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i-N_\lambda+1} \hat{\varphi}(i+N_\lambda-1)\hat{\varphi}^T(i+N_\lambda-1) \right) + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&= \frac{1}{N_\lambda} \sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i} \sum_{j=0}^{N_\lambda-1} \lambda_{\mathrm{m}}^{-j} \hat{\varphi}(i+j)\hat{\varphi}^T(i+j) + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&\ge \frac{1}{N_\lambda} \sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i} \sum_{j=0}^{N_\lambda-1} \hat{\varphi}(i+j)\hat{\varphi}^T(i+j) + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&\ge \frac{1}{N_\lambda} \sum_{i=1}^{t-N_\lambda+1} \lambda_{\mathrm{m}}^{t-i} \alpha N_\lambda \boldsymbol{I} + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&= \frac{\lambda_{\mathrm{m}}^{N_\lambda-1} - \lambda_{\mathrm{m}}^t}{1-\lambda} \alpha \boldsymbol{I} + \lambda_{\mathrm{m}}^t P^{-1}(0) \\
&= \frac{\lambda_{\mathrm{m}}^{N_\lambda-1}}{1-\lambda_m} \alpha \boldsymbol{I} + \lambda_{\mathrm{m}}^t N_\lambda \left[P^{-1}(0) - \frac{\alpha}{1-\lambda_m} \boldsymbol{I} \right]
\end{aligned} \tag{40}$$

Note that

$$
\begin{aligned}
P^{-1}(t) &\le \lambda_{\mathrm{M}} P^{-1}(t-1) + \hat{\varphi}(t)\hat{\varphi}^T(t) \\
&= \sum_{i=1}^{t} \lambda_{\mathrm{M}}^{t-i} \hat{\varphi}(i)\hat{\varphi}^T(i) + \lambda_{\mathrm{M}}^{t} P^{-1}(0) \\
&\le \sum_{i=1}^{t} \lambda_{\mathrm{M}}^{t-i} [\sum_{j=0}^{N_\lambda - 1} \hat{\varphi}(i+j)\hat{\varphi}^T(i+j)] + \lambda_{\mathrm{M}}^{t} P^{-1}(0) \\
&\le \sum_{i=1}^{t} \lambda_{\mathrm{M}}^{t-i} \beta N_\lambda \boldsymbol{I} + \lambda_{\mathrm{M}}^{t} P^{-1}(0) \\
&= \frac{1-\lambda_{\mathrm{M}}^{t}}{1-\lambda_{\mathrm{M}}} \beta N_\lambda \boldsymbol{I} + \lambda_{\mathrm{M}}^{t} P^{-1}(0) \\
&= \frac{\beta N_\lambda}{1-\lambda_{\mathrm{M}}} \boldsymbol{I} + \lambda_{\mathrm{M}}^{t} \left[P^{-1}(0) - \frac{N_\lambda \beta}{1-\lambda_{\mathrm{M}}} \boldsymbol{I} \right]
\end{aligned}
\tag{41}
$$

Owing to $\{\lambda_{\mathrm{M}}, \lambda_{\mathrm{m}}\} \in (0,1)$, it follows from (40) and (41) that

$$
\frac{\lambda_{\mathrm{m}}^{N_\lambda - 1}}{1-\lambda_{\mathrm{m}}} \alpha \boldsymbol{I} \le \lim_{t\to\infty} P^{-1}(t) \le \frac{\beta N_\lambda}{1-\lambda_{\mathrm{M}}} \boldsymbol{I} \tag{42}
$$

which is equivalent to (39) in Corollary 1. This completes the proof. □

Consequently, the asymptotic property of the proposed algorithm is addressed in the following proposition.

Proposition 1. *For a sampled system subject to load disturbance as described by (1), the proposed algorithm based on the persistent excitation condition in (38) and the assumption of* $E\left[\left\|\hat{\theta}(0)-\theta(0)\right\|_2^2\right] = \delta_0 < \infty$ *guarantees the parameter estimation error satisfies*

$$
E\left[\left\|\hat{\theta}(t)-\theta(t)\right\|_2^2\right] \le \frac{4\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}} + \frac{12 n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha \lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2+\sigma_\gamma^2+\sigma_\chi^2) + 4(n_m^2 + \frac{\beta^2 {N_\lambda}^2 (1-\lambda_{\mathrm{m}})^2}{\alpha^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}(1-\lambda_{\mathrm{M}})^4})\sigma_w^2 \tag{43}
$$

where $p_0 = P(0)$, σ_χ^2, σ_γ^2 *and* σ_w^2 *denote the variances of* $\chi(t) = \sum_{k=1}^{n_a} a_k[\hat{x}(t-k)-x(t-k)]$, $\gamma(t) = \sum_{j=1}^{n_b} b_j(z^{-d}-z^{-\hat{d}})u(t-j)$, $w(t) = \theta(t)-\theta(t-1)$, *respectively.*

Proof: Define the parameter estimation error vector by

$$
\tilde{\theta}(t) = \hat{\theta}(t) - \theta(t) \tag{44}
$$

It can be seen from (5) that $w(t) = \theta(t) - \theta(t-1)$ arises from the load disturbance response and therefore its variance is denoted by σ_w^2.

Substituting (16) into (44) yields

$$
\begin{aligned}
\tilde{\theta}(t) &= \hat{\theta}(t) - \theta(t) \\
&= \hat{\theta}(t) - [\theta(t-1) + w(t)] \\
&= \hat{\theta}(t-1) + P(t)\hat{\varphi}(t)[y(t) - \hat{\varphi}^T(t)\hat{\theta}(t-1)] - \theta(t-1) - w(t) \\
&= \tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\Bigg(\varphi^T(t)[\theta(t-1) + w(t)] + v(t) + \sum_{k=1}^{n_a} a_k[\hat{x}(t-k) - \hat{x}(t-k)] \\
&\quad + \sum_{j=1}^{n_b} (b_j z^{-\hat{d}} - b_j z^{-\hat{d}})u(t-j) - \hat{\varphi}^T(t)\hat{\theta}(t-1) \Bigg) - w(t) \\
&= \tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\Big[-\hat{\varphi}^T(t)\tilde{\theta}(t-1) + \varphi^T(t)w(t) + v(t) + \gamma(t) + \chi(t) \Big] - w(t) \\
&= \Big(\boldsymbol{I} - P(t)\hat{\varphi}(t)\hat{\varphi}^T(t) \Big)\tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\varphi^T(t)w(t) + P(t)\hat{\varphi}(t)\Gamma(t) - w(t)
\end{aligned}
\tag{45}
$$

where $\Gamma(t) = v(t) + \gamma(t) + \chi(t)$, $\chi(t) = \sum_{k=1}^{n_a} a_k[\hat{x}(t-k) - x(t-k)]$ is the mismatched error of the auxiliary model, $\gamma(t) = \sum_{j=1}^{n_b} b_j(z^{-d} - z^{-\hat{d}})u(t-j)$ arises from the time delay estimation error. Denote by σ_χ^2 and σ_γ^2 the variances of $\chi(t)$ and $\gamma(t)$, respectively.

It can be easily verified from (14) that

$$
\boldsymbol{I} - P(t)\hat{\varphi}(t)\hat{\varphi}^T(t) = \lambda P(t)P^{-1}(t-1) \tag{46}
$$

Substituting (46) into (45) yields

$$
\begin{aligned}
\tilde{\theta}(t) &= \lambda P(t)P^{-1}(t-1)\tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\varphi^T(t)w(t) + P(t)\hat{\varphi}(t)\Gamma(t) - w(t) \\
&= \lambda P(t)P^{-1}(t-1)\Big\{ \lambda P(t-1)P^{-1}(t-2)\tilde{\theta}(t-2) + P(t-1)\hat{\varphi}(t-1)\varphi^T(t-1)w(t-1) \\
&\quad + P(t-1)\varphi(t-1)\Gamma(t-1) - w(t-1) \Big\} + P(t)\hat{\varphi}(t)\varphi^T(t)w(t) + P(t)\hat{\varphi}(t)\Gamma(t) - w(t) \\
&\quad \vdots \\
&= \lambda^t P(t)P^{-1}(0)\tilde{\theta}(0) + P(t)\sum_{i=1}^{t}[\lambda^{t-i}\hat{\varphi}(i)\hat{\varphi}^T(i)w(i)] + P(t)\sum_{i=1}^{t}[\lambda^{t-i}\hat{\varphi}(i)\Gamma(i)] - P(t)\sum_{i=1}^{t}[\lambda^{t-i}P^{-1}(i)w(i)] \\
&= \lambda^t P(t)P^{-1}(0)\tilde{\theta}(0) + P(t)\Phi(t)\Phi^T(t)W(t) + P(t)\Phi(t)\Gamma_\lambda(t) - P(t)\sum_{i=1}^{t}[\lambda^{t-i}P^{-1}(i)w(i)]
\end{aligned}
\tag{47}
$$

where $\Phi(t) = [\mu^{t-1}\hat{\varphi}(1), \cdots, \hat{\varphi}(t)]$, $\Gamma_\lambda(t) = [\mu^{t-1}\Gamma(1), \cdots, \Gamma(t)]^T$, $W(t) = [w(1), \cdots, w(t)]$ and $\mu = \sqrt{\lambda}$.

Considering that $\lambda_{\mathrm{m}}P^{-1}(i) \le \boldsymbol{\lambda}(i) \circ P^{-1}(i) \le \lambda_{\mathrm{M}}P^{-1}(i)$ and using Lemma 1, we have

$$
\begin{aligned}
&E\Big[\big\| \lambda^t P(t)P^{-1}(0)\tilde{\theta}(0) \big\|_2^2 \Big] \\
&= E\Big[\mathrm{tr}[\lambda^t \tilde{\theta}^T(0)P^{-T}(0)P^T(t)P(t)P^{-1}(0)\tilde{\theta}(0)] \Big] \\
&\le \lambda_{\mathrm{M}}^{2t}(t)\rho_{\max}[P^T(t)P(t)]\rho_{\max}[P^{-T}(0)P^{-1}(0)]E\Big[\mathrm{tr}[\tilde{\theta}^T(0)\tilde{\theta}(0)] \Big] \\
&\le \frac{\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2 \lambda_{\mathrm{m}}^{2(N_\lambda - 1)}}
\end{aligned}
\tag{48}
$$

It can be seen from (14) that

$$
P^{-1}(t) = \sum_{i=1}^{t} \lambda^{t-i}\hat{\varphi}(i)\hat{\varphi}^T(i) + \lambda^t P^{-1}(0) \tag{49}
$$

$$
\begin{aligned}
P(t)\Phi(t)\Phi^T(t) &= \boldsymbol{I}_{n_m \times n_m} - \lambda P(t)P^{-1}(0) \\
&\le \boldsymbol{I}_{n_m \times n_m} - \lambda_{\mathrm{m}}^{t} P(t)P^{-1}(0)
\end{aligned} \tag{50}
$$

Correspondingly, it follows that

$$
\begin{aligned}
& E\left[\left\|P(t)\Phi(t)\Phi^T(t)W(t)\right\|_2^2\right] \\
&= E\left[\left\|P(t)\Phi(t)\Phi^T(t)\right\|_2^2\right]\sigma_w^2 \\
&\le n_{\mathrm{m}}^2\sigma_w^2
\end{aligned} \tag{51}
$$

Using (39) and (50), we have

$$
\begin{aligned}
& E\left[\left\|P(t)\Phi(t)\Gamma_\lambda(t)\right\|_2^2\right] \\
&= E\left[P(t)\Phi(t)\Gamma_\lambda(t)\Gamma_\lambda^T(t)\Phi^T(t)P^T(t)\right] \\
&= E\left[\mathrm{tr}[P(t)\Phi(t)\Gamma_\lambda(t)\Gamma_\lambda^T(t)\Phi^T(t)P^T(t)]\right] \\
&\le \frac{(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}\rho_{\max}\left[P(t)\Phi(t)\Phi^T(t)\right]E\left\{\mathrm{tr}\left[\Gamma_\lambda(t)\Gamma_\lambda^T(t)\right]\right\} \\
&\le \frac{n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}\mathrm{tr}\left(E\left[3[v(t)v^T(t)+\gamma(t)\gamma^T(t)+\chi(t)\chi^T(t)]\right]\right) \\
&\le \frac{3n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2+\sigma_\gamma^2+\sigma_\chi^2)
\end{aligned} \tag{52}
$$

$$
\begin{aligned}
& E\left[\left\|P(t)\sum_{i=1}^{t}[\lambda^{t-i}P^{-1}(i)w(i)]\right\|_2^2\right] \\
&\le \frac{(1-\lambda_{\mathrm{m}})^2}{\alpha^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}}E\left[\left\|\sum_{i=1}^{t}[\lambda^{t-i}P^{-1}(i)w(i)]\right\|_2^2\right] \\
&\le \frac{(1-\lambda_{\mathrm{m}})^2}{\alpha^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}}\sigma_w^2E\left[\left\|\sum_{i=1}^{t}[\lambda_{\mathrm{M}}^{t-i}\frac{\beta N_\lambda}{1-\lambda_{\mathrm{M}}}]\right\|_2^2\right] \\
&= \frac{(1-\lambda_{\mathrm{m}})^2}{\alpha^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}}\sigma_w^2\times\left(\frac{1-\lambda_{\mathrm{M}}^t}{1-\lambda_{\mathrm{M}}}\times\frac{\beta N_\lambda}{1-\lambda_{\mathrm{M}}}\right)^2 \\
&\le \frac{\beta^2 N_\lambda{}^2(1-\lambda_{\mathrm{m}})^2\sigma_w^2}{\alpha^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}(1-\lambda_{\mathrm{M}})^4}
\end{aligned} \tag{53}
$$

Taking the expectation on (47), we have

$$
\begin{aligned}
E\left[\left\|\tilde{\theta}(t)\right\|_2^2\right] \le 4\Bigg\{ & E\left[\left\|\lambda^t P(t)P^{-1}(0)\tilde{\theta}(0)\right\|_2^2\right] + E\left[\left\|P(t)\Phi(t)\Phi^T(t)W(t)\right\|_2^2\right] \\
& + E\left[\left\|P(t)\Phi(t)\Gamma_\lambda(t)\right\|_2^2\right] + E\left[\left\|P(t)\sum_{i=1}^{t}[\lambda^{t-i}P^{-1}(i)w(i)]\right\|_2^2\right]\Bigg\}
\end{aligned} \tag{54}
$$

Substituting (51), (52), (53) into (54), it follows the expression in (43) in Proposition 1.

This completes the proof. □

Proposition 1 can be used to evaluate the upper bound error of parameter estimation. It is seen from (43) that a larger α and a smaller β can result in a smaller upper bound of the estimation error, which means the stationarity of the input and output data can improve the parameter estimation accuracy. Besides, it is indicated by (43) that a smaller memory data length N_λ, a lower measurement noise level, or a slow-varying disturbance can facilitate reducing the upper bound of estimation error.

Owing to a one-dimensional searching strategy on the time delay, the optimal delay estimation can surely be obtained, i.e. $\hat{d} = d$, leading to $\gamma(t) = \sum_{j=1}^{n_b} b_j (z^{-d} - z^{-\hat{d}}) u(t-j) = 0$ and $\sigma_\gamma^2 = 0$. Expectation on the upper bound of the parameter estimation error is therefore simplified to

$$E\left[\left\|\hat{\theta}(t) - \theta(t)\right\|_2^2\right] \le \frac{4\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}} + \frac{8n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2 + \sigma_\chi^2) + 4(n_m^2 + \frac{\beta^2 {N_\lambda}^2 (1-\lambda_{\mathrm{m}})^2}{\alpha^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}(1-\lambda_{\mathrm{M}})^4})\sigma_w^2 \quad (55)$$

Note that when the load disturbance response settles down, $w(t)$ goes to zero. Correspondingly, the parameter estimation error in (45) is reduced to

$$\tilde{\theta}(t) = (\boldsymbol{I} - P(t)\hat{\varphi}(t)\hat{\varphi}^T(t))\tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\Gamma(t) \quad (56)$$

Substituting (46) into (56) yields

$$\begin{aligned}\tilde{\theta}(t) &= \lambda P(t)P^{-1}(t-1)\tilde{\theta}(t-1) + P(t)\hat{\varphi}(t)\Gamma(t)\\ &= \lambda P(t)P^{-1}(t-1)[\lambda P(t-1)P^{-1}(t-2)\tilde{\theta}(t-2) + P(t-1)\varphi(t-1)\Gamma(t-1)] + P(t)\hat{\varphi}(t)\Gamma(t)\\ &\ \ \vdots\\ &= \lambda^t P(t)P^{-1}(0)\tilde{\theta}(0) + P(t)\sum_{i=1}^{t}[\lambda^{t-i}\hat{\varphi}(i)\Gamma(i)]\end{aligned} \quad (57)$$

Based on the PE condition in (38), we take the expectation on both sides of (57), obtaining

$$\begin{aligned}E\left[\left\|\tilde{\theta}(t)\right\|_2^2\right] &\le 2\left\{E\left[\left\|\lambda^t P(t)P^{-1}(0)\tilde{\theta}(0)\right\|_2^2 + E\left[\left\|P(t)\Phi(t)\Gamma_\lambda(t)\right\|_2^2\right]\right]\right\}\\ &\le \frac{2\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}} + \frac{6n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2 + \sigma_\gamma^2 + \sigma_\chi^2)\end{aligned} \quad (58)$$

Owing to that the true time delay estimation can surely be obtained by using a one-dimensional searching algorithm, i.e $\hat{d} = d$ together with $\sigma_\gamma^2 = 0$, the above upper error bound can be further reduced to

$$E\left[\left\|\tilde{\theta}(t)\right\|_2^2\right] \le \frac{2\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2 \lambda_{\mathrm{m}}^{2(N_\lambda-1)}} + \frac{4n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2 + \sigma_\chi^2) \quad (59)$$

Note that the auxiliary model prediction $\hat{x}(i)$ could converge to its true value $x(i)$ for linear model identification under stochastic measurement noise and static disturbance, as discussed in the references [5, 23, 30]. The conclusion can be easily extended to time delay system identification

based on the true delay estimation, i.e.

$$\begin{aligned}
&\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}[\hat{x}(i)-x(i)]^2\\
&=\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^{t}[\frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})}u(i-\hat{d})-\frac{B(z^{-1})}{A(z^{-1})}u(i-d)]^2\\
&=\lim_{\substack{t\to\infty\\ \hat{d}=d}}\frac{1}{t}\sum_{i=1}^{t}[(\frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})}-\frac{B(z^{-1})}{A(z^{-1})})u(i-d)]^2\\
&=0
\end{aligned} \tag{60}$$

With the convergence of parameter estimation, the two AFFs will tend to one, i.e. $\lambda_{\mathrm{M}}\to 1$ and $\lambda_{\mathrm{m}}\to 1$. Since $\chi(i)$, $\gamma(i)$, and $\hat{\varphi}(i)$ are all bounded at each step for recursion, there follows for $t\to\infty$ by using (58), (59) and (60) that

$$\begin{aligned}
\lim_{t\to\infty}E\left[\left\|\tilde{\theta}(t)\right\|_2^2\right]&=\lim_{\substack{t\to\infty\\ \lambda_{\mathrm{m}}\to 1}}\frac{2\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}}+\frac{6n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2+\sigma_\gamma^2+\sigma_\chi^2)\\
&=\lim_{\substack{t\to\infty\\ \lambda_{\mathrm{m}}\to 1\\ \hat{d}=d}}\frac{2\lambda_{\mathrm{M}}^{2t}(1-\lambda_{\mathrm{m}})^2\delta_0^2}{\alpha^2 p_0^2\lambda_{\mathrm{m}}^{2(N_\lambda-1)}}+\frac{4n_{\mathrm{m}}(1-\lambda_{\mathrm{m}})}{\alpha\lambda_{\mathrm{m}}^{N_\lambda-1}}(\sigma_v^2+\sigma_\chi^2)\\
&=0
\end{aligned} \tag{61}$$

which indicates that the proposed algorithm can give unbiased estimation when the load disturbance response settles down.

5. Illustrative examples

Example 1. Consider a benchmark example studied in the reference [23],

$$y(t)=\frac{0.6804z^{-1}+0.6303z^{-2}}{1+0.412z^{-1}+0.309z^{-2}}z^{-d}u(t)+\xi(t)+v(t)$$

where $\xi(t)$ is assumed to be a disturbance of multiple sinusoidals as studied in [19], i.e. $\xi(t)=0.5\sin(0.02t)+0.5\sin(0.05t)$, and $d=10$ is assumed for illustration.

To perform an identification test, the input excitation $u(t)$ is taken as a pseudorandom binary sequence (PRBS) with a magnitude switching between ±1. The measurement noise $v(t)$ is assumed to be a white noise sequence with zero mean and variance of $\sigma_v^2=0.18^2$, causing the signal-to-noise ratio, SNR=14 (dB). For applying the proposed algorithm, $\hat{\theta}(0)=10^{-3}I_{5\times 1}$, $P(0)=10^6 I_{5\times 5}$, $\lambda_1(0)=1$, $\lambda_2(0)=1$, $\lambda_{1\min}=0.99$, $\lambda_{2\min}=0.9$, $N=1000$, $d_{\min}=0$ and $d_{\max}=15$ are taken for computation. The identification algorithm given in [19] for identifying an extended OE model is also performed. In addition, the identification algorithm composed of (20)-(23) in section 3.1 using a single forgetting

factor ($\lambda = 0.99$), therefore named as SDLS, is performed for comparison. To procure consistent estimation, a one-dimensional searching approach to estimate the integer time delay is adopted for all these identification methods. The recursive identification results for parameter estimation are shown in Fig.2. It is seen that without using any prior knowledge of the load disturbance, the proposed algorithm gives evidently improved identification accuracy and convergence rate. In contrast, the extended OE model identification algorithm given in [19] results in biased estimation on a few model parameters except for the time delay identified by the one-dimensional searching approach, due to using the predicted output rather than the auxiliary model prediction to construct the regressor, and using a single forgetting factor for recursive estimation. The SDLS algorithm gives unstable estimation arising from the use of a constant forgetting factor that cannot make a good compromise between estimating the time-invariant model parameters and tracking time-varying load disturbance.

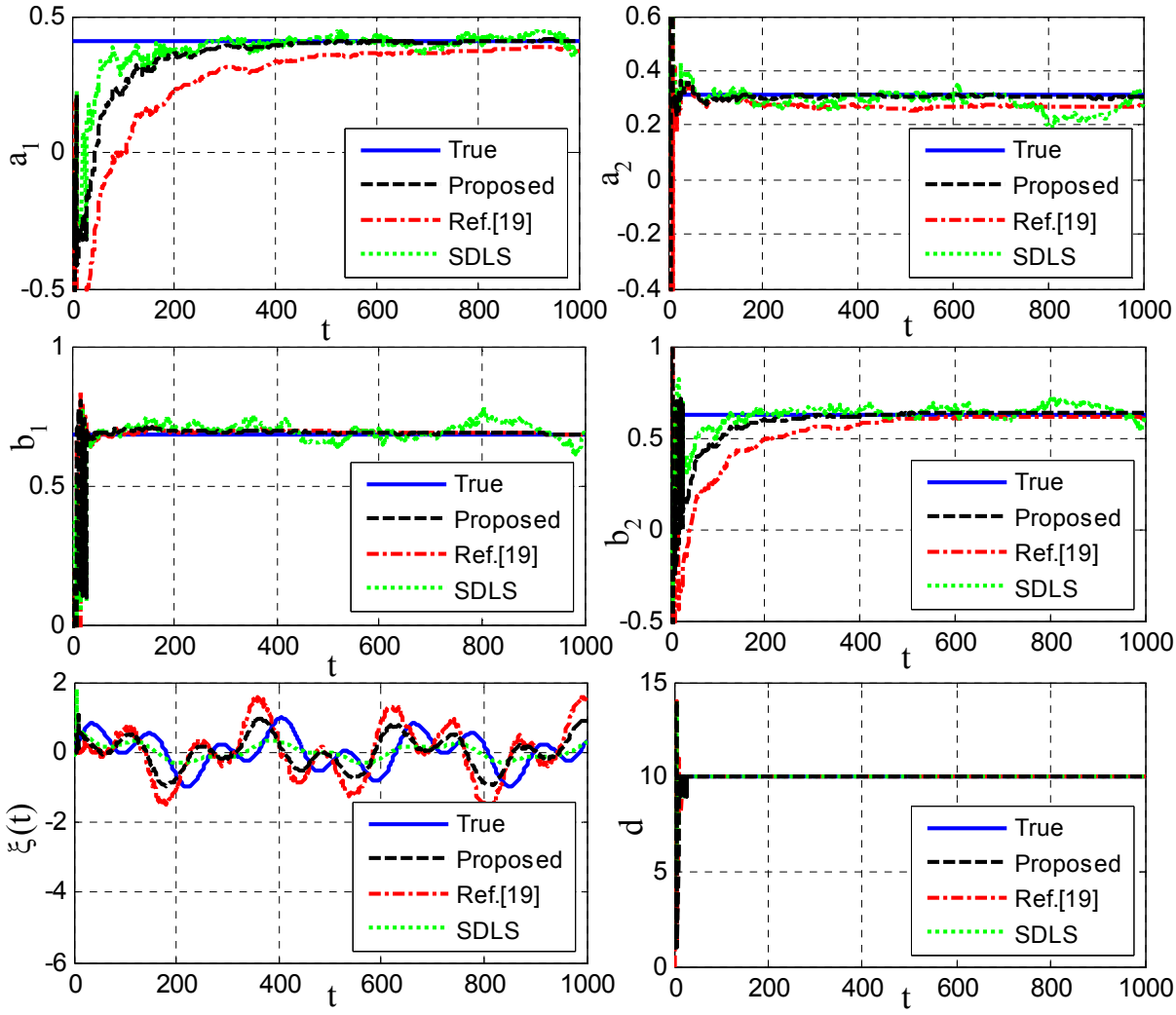


Fig. 2. Identification results by using different algorithms for Example 1

To demonstrate the identification effectiveness in the presence of time-varying load disturbance and stochastic measurement noise, one hundred Monte-Carlo (MC) tests are performed under time-varying load disturbance imitated by

$$\xi(t) = C + A_1 \sin(w_1 t) + A_2 \sin(w_2 t) \tag{62}$$

where $C \in (0, 0.7)$, $A_1 \in (0, 0.6)$, $A_2 \in (0, 0.8)$, $w_1 \in (0, 0.05)$, and $w_2 \in (0, 0.1)$ vary randomly from test to test, while the measurement noise level is the same as above, i.e. SNR≈14 dB.

For model identification, the data length is taken as $N = 4000$ in each test. The identification results are listed in Table 1, where the results are shown by the mean value along with the standard deviation in parentheses, and '*err*' denotes the relative error of the model parameter estimation with respect to the true values (i.e. $err(t) = \left\| \hat{\theta}_0(t) - \theta_0 \right\|_2 / \left\| \theta_0 \right\|_2 \times 100\%$). For assessing the load disturbance estimation, denote by $\xi_{\mathrm{m}}(t)$ the mean value of $\xi(t)$ and by $\sigma_{\xi}(t)$ the standard deviation. The corresponding estimations are denoted by $\hat{\xi}_{\mathrm{m}}(t)$ and $\hat{\sigma}_{\xi}(t)$, and the mean prediction error is denoted by $\Delta\xi_{\mathrm{m}}(t)$. The results are shown in Fig.3 for comparision. It is seen that faster convergence is obtained by the proposed algorithm, e.g. $err = 1.44\%$ at the time step $t = 2000$.

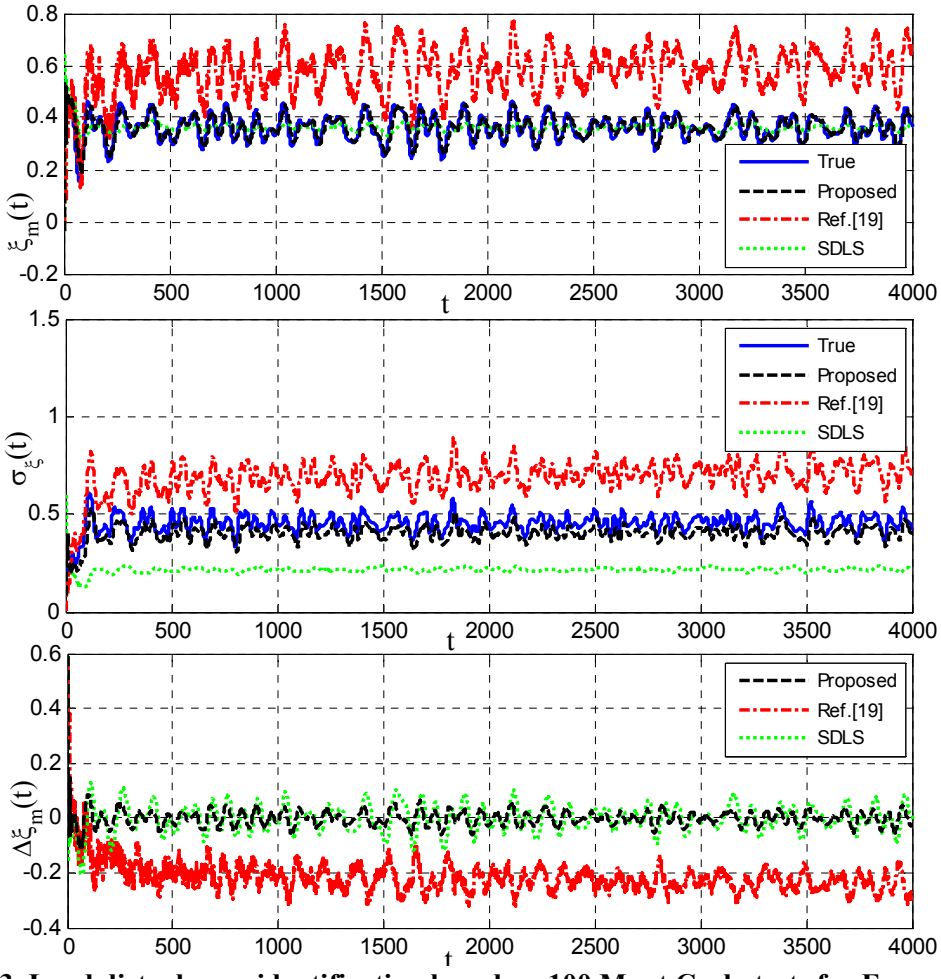


Fig. 3. Load disturbance identification based on 100 Mont Carlo tests for Example 1

Table 1 Identification results based on 100 Mont Carlo tests for Example 1

t	a_1	a_2	b_1	b_2	*d*	*err(t)*
The proposed algorithm						
100	0.3545 (±0.0604)	0.2789 (±0.0313)	0.6818 (±0.0191)	0.5973 (±0.0501)	10.00 (±0.00)	9.17 (±5.41)
500	0.4086 (±0.0164)	0.2953 (±0.0219)	0.6835 (±0.0124)	0.6323 (±0.0187)	10.00 (±0.00)	3.16 (±1.72)
1000	0.4106 (±0.0108)	0.3011 (±0.0145)	0.6823 (±0.0079)	0.6321 (±0.0119)	10.00 (±0.00)	2.05 (±1.06)
2000	0.4115 (±0.0076)	0.3046 (±0.0104)	0.6815 (±0.0061)	0.6320 (±0.0086)	10.00 (±0.00)	1.44 (±0.76)
3000	0.4120 (±0.0059)	0.3058 (±0.0084)	0.6813 (±0.0050)	0.6319 (±0.0074)	10.00 (±0.00)	1.16 (±0.63)
4000	0.4125 (±0.0054)	0.3066 (±0.0070)	0.6808 (±0.0044)	0.6318 (±0.0061)	10.00 (±0.00)	0.99 (±0.53)
True	0.412	0.309	0.6804	0.6303	10	
The OE algorithm in ref.[19]						
100	0.1643 (±0.1639)	0.1924 (±0.0543)	0.6902 (±0.0197)	0.4762 (±0.1204)	9.98 (±0.14)	30.94 (±17.80)
500	0.3382 (±0.0596)	0.2264 (±0.0449)	0.6899 (±0.0126)	0.5842 (±0.0502)	10.00 (±0.00)	12.83 (±6.09)
1000	0.3649 (±0.0445)	0.2488 (±0.0326)	0.6835 (±0.0078)	0.6048 (±0.0345)	10.00 (±0.00)	8.70 (±4.48)
2000	0.3810 (±0.0349)	0.2664 (±0.0243)	0.6830 (±0.0055)	0.6143 (±0.0251)	10.00 (±0.00)	6.06 (±3.48)
3000	0.3874 (±0.0311)	0.2752 (±0.0201)	0.6819 (±0.0044)	0.6178 (±0.0223)	10.00 (±0.00)	4.91 (±3.08)
4000	0.3914 (±0.0289)	0.2805 (±0.0175)	0.6818 (±0.0042)	0.6198 (±0.0206)	10.00 (±0.00)	4.26 (±2.79)
True	0.412	0.309	0.6804	0.6303	10	
The SDLS algorithm						
100	0.3793 (±0.0417)	0.2794 (±0.0399)	0.6911 (±0.0232)	0.6200 (±0.0413)	10.00 (±0.00)	7.56 (± 3.37)
500	0.4008 (±0.0305)	0.2961 (±0.0545)	0.6878 (±0.0298)	0.6289 (±0.0409)	10.00 (±0.00)	6.61 (±4.05)
1000	0.4025 (±0.0279)	0.3065 (±0.0448)	0.6814 (±0.0309)	0.6227 (±0.0435)	10.00 (±0.00)	5.95 (±3.96)
2000	0.4039 (±0.0253)	0.3034 (±0.0374)	0.6811 (±0.0224)	0.6244 (±0.0297)	10.00 (±0.00)	4.89 (±2.74)
3000	0.4091 (±0.0255)	0.3082 (±0.0396)	0.6797 (±0.0270)	0.6293 (±0.0363)	10.00 (±0.00)	5.04 (±3.50)
4000	0.4003 (±0.0313)	0.3192 (±0.0574)	0.6722 (±0.0456)	0.6099 (±0.0629)	10.00 (±0.00)	7.86 (±5.97)
True	0.412	0.309	0.6804	0.6303	10	

Example 2. Consider a closed-loop identification example studied in the reference [31],

$$y(t)=\frac{0.0997z^{-1}-0.0902z^{-2}}{1-1.8858z^{-1}+0.9048z^{-2}}z^{-d}u(t)+\xi(t)+v(t)$$

where $\xi(t)$ is assumed to be a disturbance of multiple sinusoidals and $d=6$ is taken for illustration.

A block diagram of the closed-loop identification test is shown in Fig.4, where $C_c(z^{-1})$ is a proportional-integral (PI) controller, the setpoint $r_2(t)$ is taken as zero, i.e. $r_2(t)=0$, while $r_1(t)$ is an external excitation for test. The plant input $u(t)$ is therefore written as

$$u(t)=r_1(t)-C_c(z^{-1})y(t) \tag{63}$$

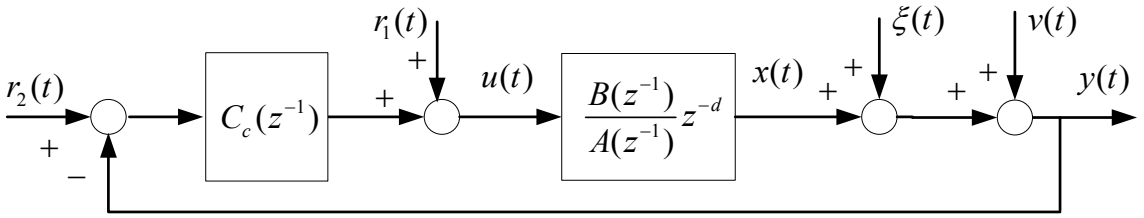


Fig. 4. Schematic of a close-loop system for identification test

To perform an identification test, the external excitation $r_1(t)$ is taken as a PRBS sequence with a magnitude switching between ±1. The noise $v(t)$ is taken as a white noise sequence with zero mean and variance of $\sigma_v^2 = 0.01^2$, causing SNR=29 (dB). One hundred MC tests are performed under different controller settings and time-varying disturbance, i.e.

$$C_c(z^{-1}) = \frac{kz^{-1} - kz^{-2}}{1 - z^{-1}} \tag{64}$$

$$\xi(t) = C + A_1 \sin(w_1 t) + A_2 \sin(w_2 t) \tag{65}$$

where $k \in (0.01, 0.2)$, $C \in (0, 0.3)$, $A_1 \in (0, 0.3)$, $A_2 \in (0, 0.4)$, $w_1 \in (0.02, 0.025)$, and $w_2 \in (0.05, 0.055)$ varies randomly from test to test.

For applying the proposed algorithm, $\hat{\theta}(0) = 10^{-3} I_{5\times1}$, $P(0) = 10^6 I_{5\times5}$, $\lambda_1(0) = 1$, $\lambda_2(0) = 1$, $\lambda_{1\min} = 0.99$, $\lambda_{2\min} = 0.7$, $N = 4000$, $d_{\min} = 0$ and $d_{\max} = 10$ are taken for computation. The extended OE model identification algorithm given in [19] and the above SDLS method with $\lambda = 0.99$ are also performed for comparsion. A one-dimensional searching approach to estimate the integer time delay is adopted for all these identification methods. The averaged parameter estimation results for 100 MC tests are shown in Fig.5. It is seen that convergent and unbiased parameter estimation is obtained by the proposed method, while the OE model identification algorithm given in [19] and the above SDLS method result in biased estimation though the time delay is identified by using the one-dimensional searching approach.

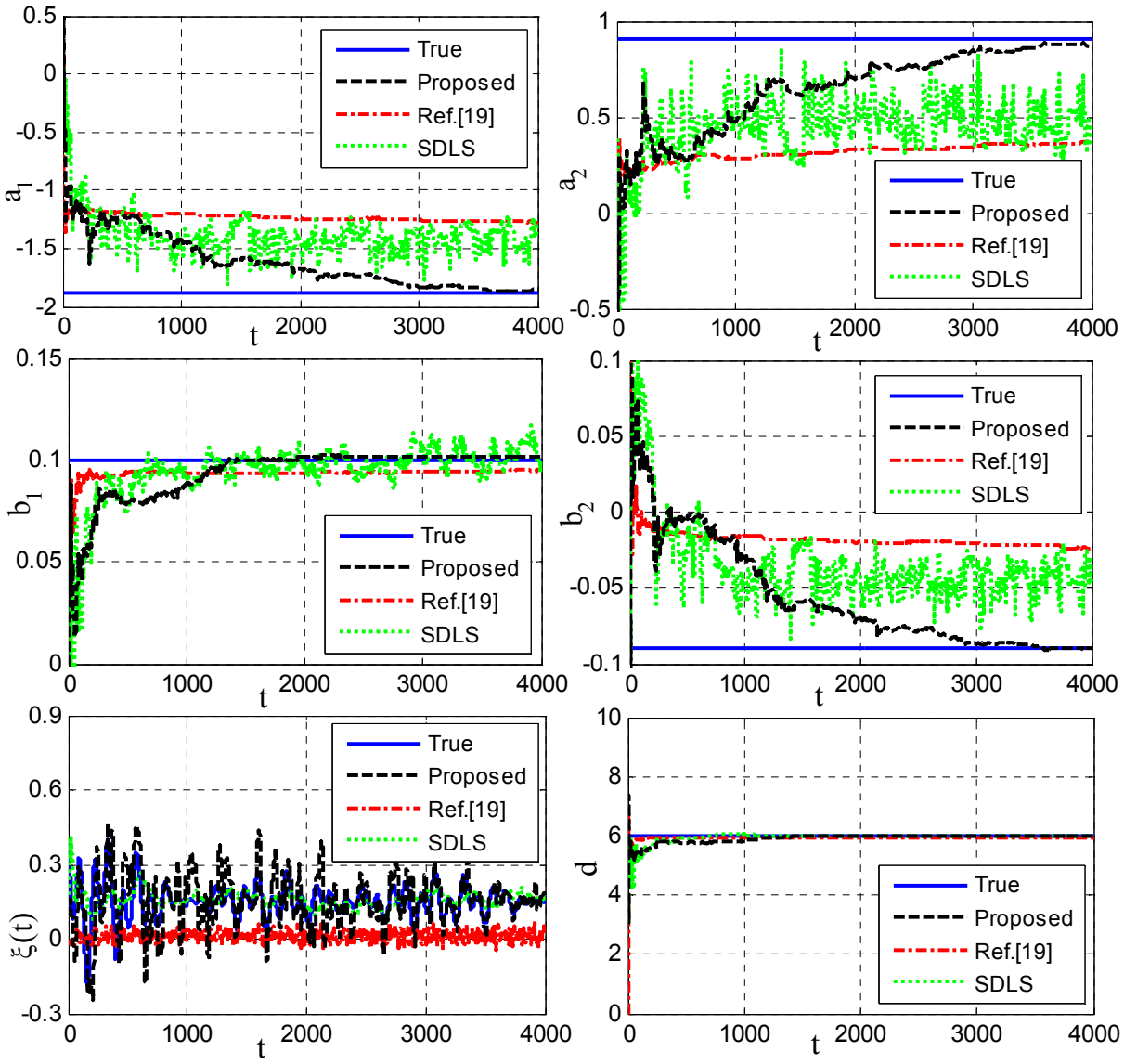


Fig. 5. Mean parameter estimates results based on 100 Mont Carlo tests for Example 2

Example 3. Consider the filling process of injection molding subject to inherent-type load disturbance as studied in the references [4, 32, 33]. The injection velocity response with respect to the proportional valve opening for an industrial injection molding machine was identified [33] as an OE model,

$$y(t)=\frac{1.69z^{-1}+1.419z^{-2}}{1-1.582z^{-1}+0.5916z^{-2}}z^{-5}u(t)+\xi(t)+v(t)$$

where the load disturbance response $\xi(t)$ varies from test to test although the same valve opening is set for operation, due to the existence of time-varying process uncertainties in practice. The fundamental dynamics of the load disturbance response for the production of a rectangular high-density polyethylene mold were estimated in [4] as $\xi(t)=[(-0.15z^{-1}-0.15z^{-2})/(1-0.993z^{-1})]\cdot 1(t)$.

For illustration, the nominal injection velocity response is assumed to be described by the above identified model for injection molding. The measurement noise $v(t)$ is assumed to be a white noise sequence with zero mean and variance of $\sigma_v^2 = 0.2^2$, in view of the fact that the maximal measurement error on the injection velocity is no larger than 1.0 (m/s) [4]. To perform a practical identificaiton test, the excitation sequence of the valve opening is taken a trapezoidal signal between 10% and 70%, which is generated in terms of a PRBS sequence with zero mean and variance of $\sigma_v^2 = 0.07^2$ to determine the duration times for the magnitudes of 10% and 70%, respectively, while the switching time between these two magnitudes is taken as five sampling periods, by monotonously increase or decrease them as shown in Fig.6. A time-varying load disturbance is imitated as

$$\xi(t) = \frac{-\beta_1 z^{-1} - \beta_2 z^{-2}}{1 - \alpha_1 z^{-1}} \delta(t) \tag{66}$$

where $\delta(t) = 1$, and

$$\beta_1 = \begin{cases} 0.05, t < 1000 \\ 0.15, t \geq 1000 \end{cases}, \beta_2 = \begin{cases} 0.08, t < 1000 \\ 0.15, t \geq 1000 \end{cases}, \alpha_1 = \begin{cases} 0.992, t < 1000 \\ 0.993, t \geq 1000 \end{cases}.$$

The injection velocity response with respect to the above trapezoidal input excitation is shown in Fig.7, with a sampling period of 10 (ms). Using the proposed algorithm, we take $\hat{\theta}(0) = 10^{-3} I_{5\times 1}$, $P(0) = 10^6 I_{5\times 5}$, $\lambda_1(0) = 1$, $\lambda_2(0) = 1$, $\lambda_{1\min} = 0.995$, $\lambda_{2\min} = 0.95$, $d_{\min} = 0$ and $d_{\max} = 10$ for computation. The extended OE model identification algorithm given in [19] and the above SDLS with $\lambda = 0.995$ are also used for comparison. The identification results are shown in Fig.8. It is once again seen that the proposed algorithm gives evidently improved estimation. Moreover, the dynamics change of load disturbance is well predicted by the proposed method, while the OE model identification algorithm given in [19] and the above SDLS method could not give proper estimation on the load disturbance and thus result in biased estimation on the model parameters.

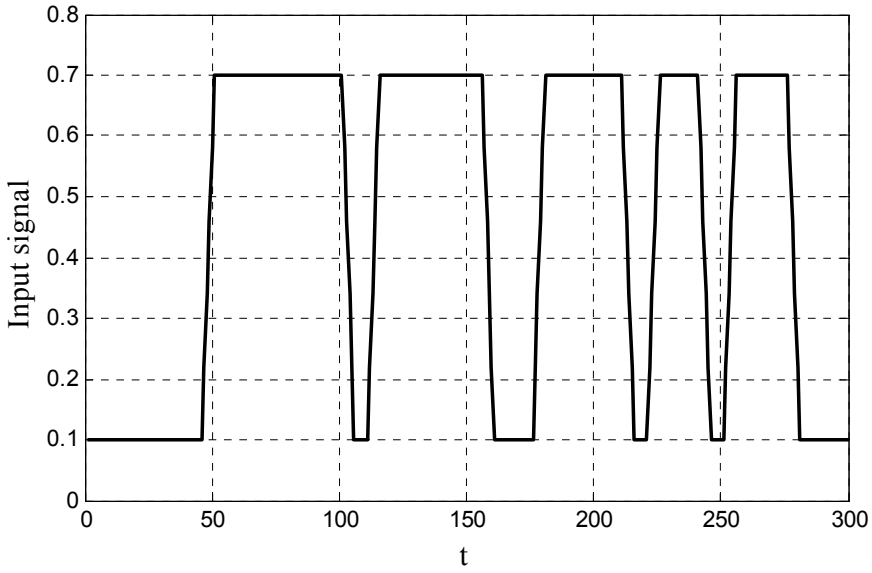


Fig. 6. Trapezoidal input excitation of the valve opening for identifying Example 3

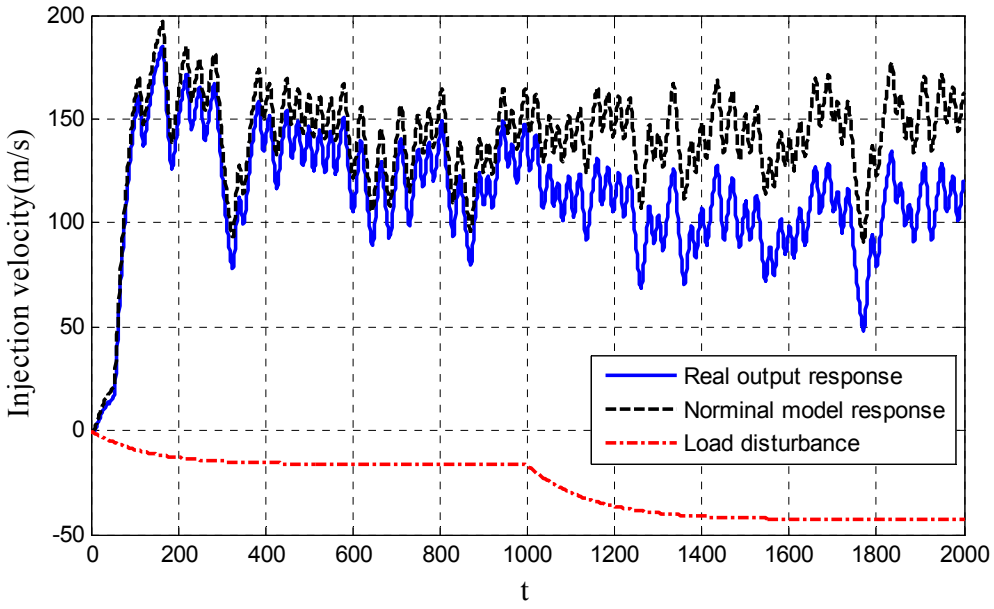


Fig. 7. The injection velocity response under an inherent type load disturbance

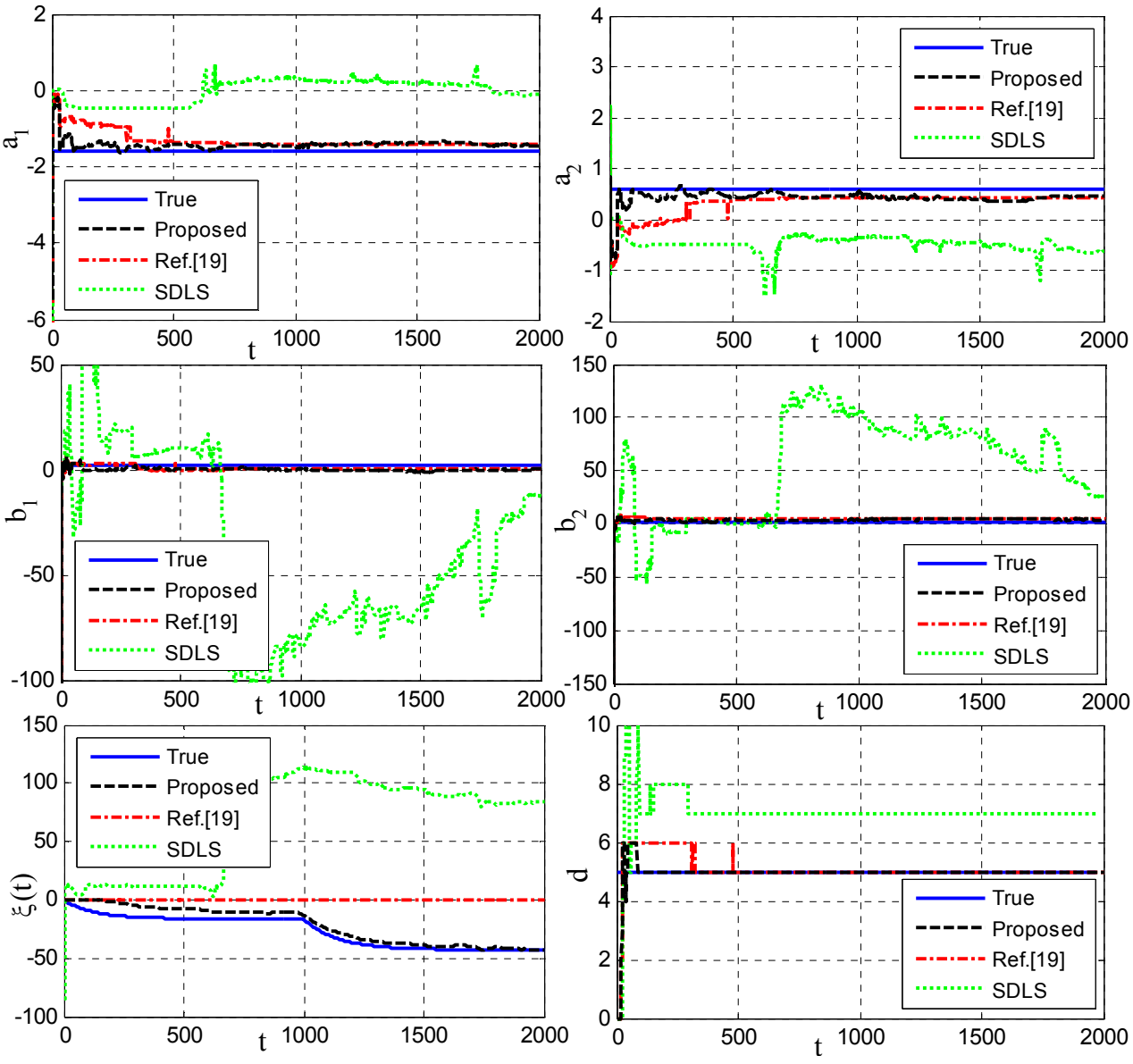


Fig. 8. Identification results using different algorithms for Example 3

6. Conclusions

For identifying industrial sampled systems with time delay subject to unknown load disturbance with deterministic dynamics, a bias-eliminated OE model identification method has been proposed by taking the output response arising from such load disturbance as a dynamic parameter for estimation together with the model parameters. A discrete-time domain recursive identification algorithm is developed to estimate the linear model parameters together with the load disturbance response, while the integer delay parameter is determined by using a one-dimensional searching approach that can guarantee the minimization of the output fitting error. Dual AFFs are introduced to

expedite the convergence rate of parameter estimation and tracking load disturbance, together with tuning guidelines for implementation. A rigorous proof is given to verify the estimation convergence under time-varying load disturbance, and moreover, it is clarified that unbiased estimation can be obtained when the load disturbance response settles down. The applications to two benchmark examples for open- and closed-loop identification, and an injection process model subject to time-varying load disturbance has well demonstrated the effectiveness and merit of the proposed identification method.

Acknowledgement

This work is supported in part by the National Thousand Talents Program of China, NSF China Grants 61473054 and 61633006, and the Fundamental Research Funds for the Central Universities of China.

References

[1] D. E. Seborg, D. A. Mellichamp, T. F. Edgar, F. J. Doyle III, Process Dynamics and Control, 3rd ed, Hoboken, NJ: John Wiley & Sons, 2011.

[2] H. Garnier, L. Wang, (ed). Identification of Continuous-time Models from Sampled Data, Springer-Verlag, London, 2008.

[3] T. Liu, F. Gao, Industrial Process Identification and Control Design: Step-test and Relay-experiment-based Methods, London: Springer, 2012.

[4] T. Liu, F. Zhou, Y. Yang, F. Gao, Step response identification under inherent-type load disturbance with application to injection molding, Ind. Eng. Chem. Res. 49 (2010) 11572-11581.

[5] P. C. Young, Recursive estimation and time-series analysis: an introduction for the student and practitioner. 2nd ed. London: Springer, 2011.

[6] L. Ljung, Perspectives on system identification, Annual Reviews in Control 34 (2010) 1-12.

[7] T. Liu, Q.G. Wang, H.P. Huang, A tutorial review on process identification from step or relay feedback test, J. Process Control 23 (2013) 1597-1623.

[8] Q. G. Wang, Y. Zhang, Robust identification of continuous systems with dead-time from step response, Automatica 37 (2001) 377-390

[9] T. Liu, F. Gao, Robust step-like identification of low-order process model under nonzero initial conditions and disturbance, IEEE Trans. Autom. Control 53 (2008) 2690-2695.

[10] S. H. Hwang, S. T. Lai, Use of two-stage least-squares algorithms for identification of continuous systems with time delay based on pulse responses, Automatica 40 (2004) 1561-1568.

[11] S. H. Hwang, H. C. Ling, S. J. Shiu, Robust identification of continuous parametric models based on multiple sinusoidal testing under slow or periodic disturbances, Ind. Eng. Chem. Res. 43 (19) (2004) 6125-6135.

[12] Q.G. Wang, M. Liu, C.C. Hang, W. Tang, Robust process identification from relay tests in the presence of nonzero initial conditions and disturbance, Ind. Eng. Chem. Res. 45 (2006) 4063-4070.

[13] I. Kaya, Parameter estimation for integrating processes using relay feedback control under static load disturbances, Ind. Eng. Chem. Res. 45(13) (2006) 4726-4731.

[14] T. Söderström, P. Stoica, System Identification, 2nd ed, Prentice Hall, New York, USA, 2001.

[15] T. Söderström, A generalized instrumental variable estimation method for errors-in-variables identification problems, Automatica, 47(8) (2011) 1656-1666.

[16] S. V. Huffel, P. Lemmerling, Total least squares and errors-in-variables modeling: analysis, algorithms and applications, London: Springer, 2013.

[17] E. W. Bai, K. M. Nagpal, R. Tempo, Bounded-error parameter estimation: Noise models and recursive algorithms, Automatica, 32(7) (1996) 985-999.

[18] P. C. Young, Refined instrumental variable estimation: Maximum likelihood optimization of a unified Box–Jenkins model, Automatica, 35 (2015) 35-46.

[19] S. Karra, M. N. Karim, Alternative model structure with simplistic noise model to identify linear time invariant systems subjected to non-stationary disturbances, J. Process Control 19 (2009) 964-977.

[20] G. Jacovitti, G. Scarano, Discrete time techniques for time delay estimation, IEEE Trans. on Signal Processing 41 (1993) 525-533.

[21] G. Ferretti, C. Maffezzoni, R. Scattolini, On the identifiability of the time delay with least-squares methods, Automatica 32 (3) (1996) 449-453.

[22] F. Chen, H. Garnier, M. Gilson, Robust identification of continuous-time models with arbitrary time-delay from irregularly sampled data, J. Process Control 25 (2015) 19-27.

[23] F. Ding, T. Chen, Combined parameter and output estimation of dual-rate systems using an auxiliary model, Automatica 40 (10) (2004) 1739-1748.

[24] M. Evestedt, A. Medvedev, Stationary behavior of an anti-windup scheme for recursive parameter estimation under lack of excitation, Automatica 42 (1) (2006) 151-157.

[25] C. Paleologu, J. Benesty, S. Ciochină, A robust variable forgetting factor recursive least-squares algorithm for system identification, IEEE. Signal Processing Letters 15 (2008) 597-600.

[26] L. Cao, H. Schwartz, A directional forgetting algorithm based on the decomposition of the information matrix, Automatica 36 (2000) 1725-1731.

[27] F. Ding, T. Chen, Performance bounds of forgetting factor least-squares algorithms for time-varying systems with finite measurement data, IEEE Trans. Circuits and Systems 52 (3) (2005) 555-566.

[28] J. E. Parkum, N. K., Poulsen, J. Holst, Recursive forgetting algorithms. International Journal of Control 55(1) (1992) 109-128.

[29] D. S. Bernstein, Matrix Mathematics, 2nd Edition, Princeton: Princeton University Press, 2009.

[30] X. Wang, F. Ding, Convergence of the auxiliary model-based multi-innovation generalized extended stochastic gradient algorithm for Box–Jenkins systems. Nonlinear Dynamics, 82(1-2) (2015) 269-280.

[31] M. Gilson, H. Garnier, P. C. Young, & P. M. Van den Hof, Optimal instrumental variable method for closed-loop identification. IET control theory & applications, 5(10) (2011) 1147-1154.

[32] Y. Yang, F. Gao, Injection velocity control using a self-tuning controller for thermoplastic injection molding, Int. Polym. Process. XIV (1999) 196-208.

[33] J. Shi, F. Gao, T.-J. Wu, Robust design of integrated feedback and iterative learning control of a batch process based on a 2D Roesser system, J. Process Control 15 (2005) 907-924.



https://dspace.org/

Identification of discrete-time output error model for industrial processes with time delay subject to load disturbance

Dong, Shijian; Liu, Tao; Wang, Wei; Bao, Jie; Cao, Yi

2016-12-22
Elsevier
http://dx.doi.org/10.1016/j.jprocont.2016.11.007
Dong S, Liu T, Wang W, et al., Identification of discrete-time output error model for industrial processes with time delay subject to load disturbance. Journal of Process Control, Volume 50, February 2017, Pages 40–55

Downloaded from a DSpace repository. https://dspace.org/