Preprints of the
9th International Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
June 7-10, 2015, Whistler, British Columbia, Canada

Iterative Identification of Output Error Model with Time Delay

Shijian Dong, Tao Liu *, Fengwei Chen

Institute of Advanced Control Technology, Dalian University of Technology, Dalian 116024, China
**Corresponding Author, Tel: 0411-84796465; E-mail: liurouter@ieee.org*

Abstract: In this paper an iterative identification method is proposed to simultaneously estimate the parameters of an output error model together with a delay parameter for industrial processes with time delay, to facilitate model based control design. An extended observation vector is constructed to establish a negative-gradient based iterative identification algorithm. An auxiliary model is established to estimate the unknown noise-free output of the system for consistent estimation. Moreover, a variable forgetting factor (VFF) is introduced to enhance the convergence rate and identification accuracy against measurement noise. The convergence of the proposed algorithm is analyzed with a strict proof. An illustrative example is shown to demonstrate the effectiveness of the proposed identification method.

Keywords: Time delay, output error model, auxiliary model, variable forgetting factor, convergence

1. INTRODUCTION

Control-oriented model identification methods have been increasingly explored in the recent years, owing to the fact that model based control strategies have been widely used to obtain superior system performance for various industrial and chemical processes, as surveyed by Liu, Wang, and Huang (2013) and Ljung (2010). Since time delay is usually associated with industrial applications, linear transfer function models with time delay have been extensively used for control system design and controller tuning (Liu and Gao, 2012; Richard, 2013). In fact, it is quite challenging to identify a transfer function model with time delay due to the nonlinear relationship between the delay parameter and the other model parameters, especially in the presence of measurement noise. Based on step or relay tests, identification methods for obtaining frequency domain transfer function models with time delay have been developed in the literature (Hang, Åström, Wang, 2002; Atherton, 2006; Ahmed, Huang, and Shah, 2007; Liu and Gao, 2009 and 2010). It should be noted that most of these methods are implemented offline.

For the convenience of online control tuning, online model identification has become appealing in modern process industry. In continuous-time domain, a number of online identification methods have been developed by Gomez, Orlov, and Kolmanovsky (2007), Gawthrop, Nihtila, and Rad. (1980), Rad, Lo, and Tsang (2003), Ren, Rad, Chan, and Lo (2005) and Na, Ren, and Xia (2014). In discrete-time domain, a recursive least-squares (RLS) identification method for sampled time delay systems was proposed by Ferretti, Maffezzoni, and Scattolini (1991), where a recursive update of the time delay was suggested by inspecting the phase change of the identified model. Elnaggar, Dumont, and Elshafei (1990) proposed an online identification algorithm for estimating all the model parameters including time delay in terms of a two-step procedure, the first step assumes a known time delay to estimate the other model parameters, and then the second step determines the optimal model parameters by minimizing the squared output error index with respect to the assumed time delay. By taking partial derivatives with respect to the time-shift operator to construct a generalized regression vector, Bedoui, Abderrahim, and Ltaief (2013, 2012) proposed a few algorithms for simultaneously estimating all the model parameters including time delay, but these algorithms for output error model would cause bias.

Because the identification of time delay systems is involved with nonlinear formulation of model parameters, existing linear recursive algorithms cannot be directly applied to estimate all the model parameters. To solve this problem, this paper proposed an extended observation vector for iterative identification of an output error model with time delay. In view of that the standard RLS algorithm cannot guarantee consistent estimation for output error model (Söderström and Stoica, 1989; Ljung, 2002), an auxiliary model as developed by Peter C. Young (2011) is introduced to construct instrumental variables for improving consistent estimation. To relieve computation load, it is suggested to use a variable forgetting factor (VFF) in combination with a stochastic gradient (SG) algorithm for parameter estimation, which can apparently improve the convergence rate and estimation accuracy. For clarity, the paper is organized as follows. Section 2 presents the problem formulation. The proposed identification algorithm is detailed in section 3. The convergence analysis of the algorithm is given in section 4. An illustrative example is shown in section 5. Finally, some conclusions are given in Section 6.

2. PROBLEM FORMULATION

Consider an industrial process with time delay described by the following output error model with time delay,

$$y(t)=\frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)}z^{-d}u(t)+v(t) \tag{1}$$

where $u(t)$ and $y(t)$ are the input and output signals of the plant, respectively, $v(t)$ denotes measurement noise or disturbance, d indicates the process time delay, and

$$A\left(z^{-1}\right)=1+a_1z^{-1}+a_2z^{-2}+\cdots+a_{n_a}z^{-n_a} \tag{2}$$

$$B\left(z^{-1}\right)=b_1z^{-1}+b_2z^{-2}+\cdots+b_{n_b}z^{-n_b} \tag{3}$$

In this study, the following assumptions are made for model identification.

A1. Measurement noise $v(t)$ satisfies $\mathrm{E}\left[v(t)\right]=0$, $\mathrm{E}\left[\left\|v(t)\right\|^2\right]\le\sigma_v{}^2<\infty$ and $\mathrm{E}\left[v(t)v(i)\right]=0$ for $i\ne t$.

A2. The process input, $u(t)$, is uncorrelated with $v(t)$.

A3. The process to be identified is observable and controllable, while $A\left(\mathrm{z}^{-1}\right)$ and $B\left(\mathrm{z}^{-1}\right)$ are coprime. The polynomial degrees, n_a and n_b, are known.

A4. The system is causal, i.e. $y(t)$ depends on $u(s)$ and $v(s)$ for $s\le t$, but not on future values of $u(t)$ and $v(t)$.

A5. In the parameter estimation, the time delay is rounded to be an integer for practical implementation.

3. PROPOSED ALGORITHM

Define the unknown noise-free output (i.e. true output)

$$x(t)=\frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)}z^{-d}u(t) \tag{4}$$

Define the parameter vector, generalized parameter vector including the time delay and information vector, respectively, by

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

$$\theta_G=\left[\theta^T,d\right]^T\in\Re^{n}$$

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

where $n_0=n_a+n_b$ and $n=n_a+n_b+1$.

Correspondingly, Eq. (4) can be rewritten as

$$x(t)=\varphi^T(t)\theta(t) \tag{5}$$

The model in (1) can be rewritten as,

$$y(t)=x(t)+v(t)=\varphi^T(t)\theta+v(t) \tag{6}$$

Note that the process output can be estimated by

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

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

Hence, the prediction error can be computed by

$$e\left(t,\hat{d}\right)=y(t)-\hat{\varphi}^T(t)\hat{\theta}(t) \tag{8}$$

It is obvious that the unknown time delay impede computing the prediction error in (8). To resolve the problem, the following cost function is adopted herein,

$$J\left(t,\hat{\theta}_G(t)\right)=\frac{1}{2}e\left(t,\hat{d}\right)^2 \tag{9}$$

Using the negative gradient search approach (Goodwin and Sin, 1984), the SG algorithm for estimating $\hat{\theta}_G(t)$ is presented as follows.

$$\begin{aligned}\hat{\theta}_G(t)&=\hat{\theta}_G(t-1)-\mu(t)\,grad\left[J\left(t,\hat{\theta}_G(t-1)\right)\right]\\&=\hat{\theta}_G(t-1)-\mu(t)\left[-\hat{\varphi}^T(t),\frac{\partial e\left(t,\hat{d}\right)}{\partial\hat{d}}\right]^T e\left(t,\hat{d}\right)\end{aligned} \tag{10}$$

where $\mu(t)$ is the step size or convergence factor.

Taking the first derivative of $e\left(t,\hat{d}\right)$ with respect to $\hat{d}$, we have

$$\begin{aligned}\frac{\partial e\left(t,\hat{d}\right)}{\partial\hat{d}}&=-\sum_{j=1}^{n_b}\frac{\hat{b}_j\partial u\left(t-j-\hat{d}\right)}{\partial\hat{d}}\\&=-\sum_{j=1}^{n_b}\hat{b}_j\left(\frac{\partial u\left(t-j-\hat{d}\right)}{\partial\left(t-j-\hat{d}\right)}\times\frac{\partial\left(t-j-\hat{d}\right)}{\partial\hat{d}}\right)\\&=\sum_{j=1}^{n_b}\hat{b}_j\frac{\partial u\left(t-j-\hat{d}\right)}{\partial\left(t-j-\hat{d}\right)}\\&=\sum_{j=1}^{n_b}\hat{b}_j\lim_{\gamma\to0}\frac{u\left(t-j-\hat{d}\right)-u\left(t-j-\hat{d}-\delta\right)}{\left(t-j-\hat{d}\right)-\left(t-j-\hat{d}-\delta\right)}\end{aligned} \tag{11}$$

Given a sampled system, there exists $\gamma=1$ when the sampling period denoted by T_s goes to zero. Therefore, the first derivative of $e(t)$ with respect to $\hat{\theta}_G$ may be numerically computed by

$$\frac{\partial e\left(t,\hat{d}\right)}{\hat{\theta}_G(t-1)}=\left[-\hat{\varphi}^T(t),\sum_{j=1}^{n_b}\hat{b}_j z^{-\hat{d}}\Delta u(t-j)\right]^T \tag{12}$$

Let $\phi(t)=-\dfrac{\partial e\left(t,\hat{d}\right)}{\hat{\theta}_G(t-1)}$ be a generalized observation vector, Eq. (10) can be rewritten as

$$\hat{\theta}_G(t)=\hat{\theta}_G(t-1)+\mu(t)\phi(t)e\left(t,\hat{d}\right) \tag{13}$$

To guarantee the convergence of $\hat{\theta}_G(t)$, $\mu(t)$ as discussed by Peter C. Young (2011) must satisfy

$$0<\mu(t)<\frac{2}{\phi(t)^T\phi(t)} \tag{14}$$

Owing to the fact that a SG-type algorithm does not involve the computation on covariance matrices, it has less computation load, but its convergence rate may be slow. Therefore, we introduce a variable forgetting factor to improve its tracking performance, i.e. letting

$$\mu(t)=\frac{1}{r(t)} \tag{15}$$

$$r(t)=\lambda(t)r(t-1)+\left\|\phi(t)\right\|_{\mathrm{F}}^2,\ r(0)=1 \tag{16}$$

where $\lambda(t)\in(0,1]$ is a variable forgetting factor (VFF) defined by

$$\lambda(t)=\frac{1}{\exp\left(\sum_{i=0}^{M}\left(\left\|\Delta\hat{\theta}_G(t-1-i)\right\|_F^2\times\left\|e(t-i)\right\|_F^2\right)\right)} \tag{17}$$

$$\lambda(t)\geq\lambda_{\min} \tag{18}$$

where M is a window used to alleviate the influence from disturbance or noise, and $\lambda_{\min}$ is a lower limit of the VFF for implementation.

The key idea behind using the VFF is to set a smaller forgetting factor to obtain a faster convergence rate when the prediction error is large, and when the algorithm converges to the true value, the VFF will become a larger value to ensure good accuracy as discussed by Ljung (2002).

However, the information vector $\hat{\varphi}(t)$ and $\phi(t)$ contains the unknown inner variables $x(t-i)$. A feasible solution is therefore given based on using an auxiliary model as discussed by Peter C. Young (2011). Hence, the proposed negative-gradient algorithm, named as AMVFSG, is summarized as below.

Step1. Take $\hat{\theta}_G(0)$ as a nonzero vector with very small elements. e.g. $\hat{\theta}_G(0)=10^{-3}I_{n\times1}$, $r(0)=\delta$ where δ is a small positive number, e.g. $r(0)=1$, M an suitable window length, e.g. $M\in[0,\ 10]$, and $\lambda_{\min}\in(0,1]$.

Step2. Construct the observation vector $\hat{\varphi}(t)$, and the generalized observation vector $\phi(t)$ in terms of the sampled data.

Step3. Compute $\hat{\theta}_G$ by

$$e(t)=y(t)-\hat{\varphi}^T(t)\hat{\theta}(t-1) \tag{19}$$

$$\lambda(t)=\frac{1}{\exp\left(\sum_{j=0}^{M}\left(\left\|\Delta\hat{\theta}_G(t-1-j)\right\|_F^2\times\left\|e(t-j)\right\|_F^2\right)\right)} \tag{20}$$

$$\lambda(t)\geq\lambda_{\min} \tag{21}$$

$$r(t)=\lambda(t)r(t-1)+\left\|\phi(i)\right\|_F^2 \tag{22}$$

$$\hat{\theta}_G(t)=\hat{\theta}_G(t-1)+\frac{\phi(t)}{r(t)}e(t) \tag{23}$$

$$\begin{aligned}\hat{\varphi}(t)=[&-\hat{x}(t-1),\cdots,-\hat{x}(t-n_a),\\ &u(t-1-\hat{d}),\cdots,u(t-n_b-\hat{d})]^T\end{aligned} \tag{24}$$

$$\phi(t)=\left[\hat{\varphi}^T(t),-\sum_{i=1}^{n_b}\hat{b}_i z^{-\hat{d}}\Delta u(t-i)\right]^T \tag{25}$$

$$\hat{x}(t)=\hat{\varphi}^T(t)\hat{\theta}(t) \tag{26}$$

Step4. Increase t by 1 and return to Step 2, until a specified convergence condition $\left\|\hat{\theta}_G(t)-\hat{\theta}_G(t-1)\right\|_2^2\leq\varepsilon$ is satisfied or $t=N$.

4. CONVERGENCE ANALYSIS

In this section, using the stochastic process theory and the martingale hyper convergence theorem, we focus on analyzing the convergence property of the proposed AMVFSG algorithm. The following lemma is given first to establish the main convergence results.

Lemma 1. For the system described in Eq. (6), if the generalized information vector $\phi(t)$ is constructed by persistent excitation, that is, there exist constants $0<\alpha\leq\beta<\infty$ and $N\geq n$ such that the following strong persistent excitation (SPE) condition holds

$$\alpha I\leq\frac{1}{N}\sum_{j=0}^{N-1}\phi(t+j)\phi^T(t+j)\leq\beta I,\quad a.s.,\ \mathrm{t}>0 \tag{27}$$

where $0<\alpha\leq\beta<\infty$, and $N\geq n$, while $r(0)$ is chosen to satisfy

$$\frac{n\alpha}{1-\lambda}\leq r(0)\leq\frac{nN\beta}{1-\lambda} \tag{28}$$

there stands

$$\frac{\lambda^{N-1}}{1-\lambda}n\alpha\leq r(t)\leq\frac{nN\beta}{1-\lambda} \tag{29}$$

Proof: Taking the Frobenius norm on (27), we have

$$nN\alpha\leq\left\|\phi(t)\right\|_F^2\leq\delta=nN\beta \tag{30}$$

Taking into account Eq. (22), it follows that

$$\begin{aligned}Nr(t)&=N\sum_{i=1}^{t}\lambda^{t-i}\left\|\phi(i)\right\|_F^2+N\lambda^t r(0)\\ &\geq\sum_{i=1}^{t-N+1}\lambda^{t-i}tr\left[\sum_{k=0}^{N-1}\phi(i+k)\phi^T(i+k)\right]+N\lambda^t r(0)\\ &\geq\sum_{i=1}^{t-N+1}\lambda^{t-i}nN\alpha+N\lambda^t r(0)\\ &=\frac{\lambda^{N-1}}{1-\lambda}nN\alpha+\left[r(0)-\frac{n\alpha}{1-\lambda}\right]\lambda^t N\end{aligned} \tag{31}$$

Thus, there is

$$r(t)\geq\frac{\lambda^{N-1}}{1-\lambda}n\alpha+\left[r(0)-\frac{n\alpha}{1-\lambda}\right]\lambda^t \tag{32}$$

Note that

$$\begin{aligned}r(t)&=\sum_{i=1}^{t}\lambda^{t-i}\left\|\phi(i)\right\|_F^2+\lambda^t r(0)\\ &\leq\sum_{i=1}^{t}\lambda^{t-i}tr\left[\sum_{k=0}^{N-1}\phi(i+k)\phi^T(i+k)\right]+\lambda^t r(0)\\ &\leq\sum_{i=1}^{t}\lambda^{t-i}nN\beta+\lambda^t r(0)\\ &=\frac{nN\beta}{1-\lambda}+\left[r(0)-\frac{nN\beta}{1-\lambda}\right]\lambda^t\end{aligned} \tag{33}$$

Combining Eq. (28), there stands

$$\frac{\lambda^{N-1}}{1-\lambda}n\alpha\leq r(t)\leq\frac{nN\beta}{1-\lambda} \tag{34}$$

This completes the proof of Lemma 1.

Theorem 1. For the system described in Eq. (6), with assumptions A1 and A2 and the SPE condition in Eq. (27), the parameter estimation error given by the AMVFSG algorithm in Eq. (19)-(26) satisfies

$$\lim_{t\to\infty}\sup E\left[\left\|\hat{\theta}_G(t)-\theta_G\right\|_F^2\right]\le\frac{n\left[2N^3-4N^2+7N+1\right](1-\lambda)\beta^2}{3\lambda^{N-1}\alpha^3}\sigma_v^2$$

Proof: Define the parameter estimation error vector

$$\tilde{\theta}_G(t)=\hat{\theta}_G(t)-\theta_G \tag{35}$$

Using (8) and (23) we have

$$\begin{aligned}\tilde{\theta}_G(t)&=\hat{\theta}_G(t-1)-\theta_G+\frac{\phi(t)}{r(t)}e(t)\\&=\tilde{\theta}_G(t-1)+\frac{\phi(t)}{r(t)}\left\{\left[\varphi^T(t)-\hat{\varphi}^T(t)\right]\theta+v(t)-\hat{\varphi}^T(t)\tilde{\theta}(t-1)\right\}\\&=\tilde{\theta}_G(t-1)+\frac{\phi(t)}{r(t)}\left[\triangle(t)-\tilde{y}(t)+v(t)\right]\end{aligned}\tag{36}$$

where

$$\tilde{y}(t)=\hat{\varphi}^T(t)\tilde{\theta}(t-1) \tag{37}$$

$$\triangle(t)=\left[\varphi(t)-\hat{\varphi}(t)\right]^T\theta \tag{38}$$

Taking the Frobenius norm on Eq. (36), we have

$$\begin{aligned}\left\|\tilde{\theta}_G(t)\right\|_F^2&=\left\|\tilde{\theta}_G(t-1)+\frac{\phi(t)}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]\right\|_F^2\\&=\left\|\tilde{\theta}_G(t-1)\right\|_F^2+2\phi(t)^T\frac{\tilde{\theta}_G(t-1)}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]\\&\quad+\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]^2\end{aligned}\tag{39}$$

Note that

$$\phi(t)^T\tilde{\theta}_G(t-1)=\tilde{y}(t)+\tilde{\psi}(t) \tag{40}$$

where $\tilde{\psi}(t)=-\sum_{i=1}^{n_b}\left[\hat{b}_i\left(\hat{d}-d\right)z^{-\hat{d}}\Delta u(t-i)\right]$

Define a non-negative definite function $T(t)=\left\|\tilde{\theta}_G(t)\right\|_F^2$, we have

$$\begin{aligned}T(t)&=T(t-1)-\left[\frac{2}{r(t)}-\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\right]\tilde{y}^2(t)\\&\quad+2\left[\frac{1}{r(t)}-\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\right]\tilde{y}(t)\left[\triangle(t)+v(t)\right]\\&\quad+\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\left[\triangle^2(t)+v^2(t)+2\triangle(t)v(t)\right]\\&\quad+2\frac{\tilde{\psi}}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]\end{aligned}\tag{41}$$

Using Eq.(22), it follows

$$\begin{aligned}T(t)&=T(t-1)-\frac{r(t)+\lambda(t)r(t-1)}{r^2(t)}\tilde{y}^2(t)\\&\quad+\frac{2\lambda(t)r(t-1)}{r^2(t)}\tilde{y}(t)\left[\triangle(t)+v(t)\right]\\&\quad+\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\left[\triangle^2(t)+v^2(t)+2\triangle(t)v(t)\right]\\&\quad+2\frac{\tilde{\psi}}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]\end{aligned}\tag{42}$$

Due to $\frac{r(t)+\lambda(t)r(t-1)}{r^2(t)}=\frac{1}{r(t)}\left[1+\frac{\lambda(t)r(t-1)}{r(t)}\right]\ge\frac{1}{r(t)}$, there exists

$$\begin{aligned}T(t)&\le T(t-1)-\frac{1}{r(t)}\tilde{y}^2(t)\\&\quad+\frac{2r(t-1)}{r^2(t)}\tilde{y}(t)\left[\triangle(t)+v(t)\right]\\&\quad+\frac{\left\|\phi(t)\right\|_F^2}{r^2(t)}\left[\triangle^2(t)+v^2(t)+2\triangle(t)v(t)\right]\\&\quad+2\frac{\tilde{\psi}}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right]\end{aligned}\tag{43}$$

Assume that $\triangle(t)$ is bounded by $0\le\triangle^2(t)\le\vartheta$. Because $\tilde{\theta}_G(t-1)$, $\tilde{y}(t)$, $\phi(t)$, $r(t)$, $\tilde{\psi}$ and $\triangle(t)$ are uncorrelated with $v(t)$, by taking the expectation of Eq.(43) with respect to F_{t-1}, where $\{F_{t-1}\}$ is the algebra sequence generated by $\{v(t-1)\}$, we have

$$E\left[T(t)|F_{t-1}\right]-T(t-1)\le-\frac{1}{r(t)}\left[\tilde{y}^2(t)-\frac{\delta}{r(t)}\left[\sigma_v^2+\vartheta\right]\right] \tag{44}$$

Define $f(t)=\frac{1}{r(t)}\left[\tilde{y}^2(t)-\frac{\delta}{r(t)}\left[\sigma_v^2+\vartheta\right]\right]$ and $\eta_t=\frac{\delta}{r(t)}\left[\sigma_v^2+\vartheta\right]$, we have

$$\eta_{\max}=\sup_t\left[\eta_t\right]\le\frac{\delta(1-\lambda)}{\lambda^{N-1}n\alpha}\left[\sigma_v^2+\vartheta\right] \tag{45}$$

Define $R_t=\left[\tilde{\theta}_G(t):\tilde{y}^2(t)\le\eta_{\max}\right]$ in terms of its neighbourhood bounded by ε,

$$N_\varepsilon(R_t)=\left[\tilde{\theta}_G(t):\tilde{y}^2(t)\le\eta_{\max}+\varepsilon\right] \tag{46}$$

we obtain for $\tilde{\theta}_G(t)\in N_\varepsilon^c(R_t)$ that

$$f(t)\ge\frac{1}{r(t)}\varepsilon\ge\frac{(1-\lambda)}{\lambda^{N-1}n\alpha}\varepsilon>0 \tag{47}$$

Using the Martingale Hyper Convergence Theorem (MHCT) (Ding and Yang, 1999), we ensure $\tilde{\theta}_G(t)\in N_\varepsilon(R_t)$. Because ε is arbitrary real number, we have

$$\tilde{\theta}_G(t)=\left[\tilde{\theta}_G(t):\tilde{y}^2(t)\le\frac{\delta(1-\lambda)}{\lambda^{N-1}n\alpha}\left[\sigma_v^2+\vartheta\right]\right]\in R_t \tag{48}$$

Notice that the auxiliary model satisfies (Peter C. Young, 2011, Ding and Chen, 2005)

$$\lim_{t\to\infty}\frac{1}{\mathrm{t}}\sum_{i=1}^{t}\left[\hat{x}(i)-x(i)\right]^2=0 \tag{49}$$

$$\lim_{t\to\infty}\frac{1}{t}\tilde{\theta}_G(t)=\lim_{t\to\infty}\frac{1}{t}\left[\tilde{\theta}_G(t):\tilde{y}^2(t)\le\frac{\delta}{r(t)}\sigma_v^2\right]\in R_\infty \tag{50}$$

Let

$$\Delta\tilde{\theta}_G(t)=\frac{\phi(t)}{r(t)}\left[-\tilde{y}(t)+\triangle(t)+v(t)\right] \tag{51}$$

Using Eq.(36) and Eq.(40), we obtain

$$\tilde{\theta}_G(t+i) = \tilde{\theta}_G(t-1) + \sum_{k=0}^{i} \Delta\tilde{\theta}_G(t+k) \tag{52}$$

$$\begin{aligned}\phi(t+i)\phi^T(t+i)\tilde{\theta}_G(t-1) &= \phi(t+i)\left[\tilde{y}(t+i)+\tilde{\psi}(t+i)\right] \\ &\quad - \phi(t+i)\phi^T(t+i)\sum_{k=0}^{i-1}\Delta\tilde{\theta}_G(t+k)\end{aligned} \tag{53}$$

$$\begin{aligned}&\left[\sum_{i=0}^{N-1}\phi(t+i)\phi^T(t+i)\right]\tilde{\theta}_G(t-1) \\ &= \sum_{i=0}^{N-1}\phi(t+i)\left[\tilde{y}(t+i)+\tilde{\psi}(t+i)\right] \\ &\quad - \sum_{i=0}^{N-1}\left[\phi(t+i)\phi^T(t+i)\sum_{k=0}^{i-1}\Delta\tilde{\theta}_G(t+k)\right]\end{aligned} \tag{54}$$

Taking the Frobenius norm on Eq. (54) and using Eq.(30), we have

$$\begin{aligned}\left\|\tilde{\theta}_G(t-1)\right\|_F^2 \le \frac{2}{(N\alpha)^2}&\left\{\left\|\sum_{i=0}^{N-1}\phi(t+i)\left[\tilde{y}(t+i)+\tilde{\psi}(t+i)\right]\right\|_F^2\right. \\ &\left. +\delta^2\left[\sum_{i=0}^{N-1}\sum_{k=0}^{i-1}\left\|\Delta\tilde{\theta}_G(t+k)\right\|_F\right]^2\right\}\end{aligned} \tag{55}$$

Taking the expectation of Eq. (55), we obtain

$$\begin{aligned}&\limsup_{t\to\infty} E\left[\left\|\hat{\theta}_G(t)-\theta_G\right\|_F^2\right] \\ &\le \frac{2}{(N\alpha)^2}\limsup_{t\to\infty} E\left\{N\delta\tilde{y}^2(t) + \frac{\delta^2(N-1)(2N^2-2N-1)}{12r(t)}\left[\tilde{y}^2(t)+\sigma_v^2\right]\right\} \\ &\le \frac{2}{(N\alpha)^2}\limsup_{t\to\infty} E\left\{\frac{N\delta^2}{r(t)}\sigma_v^2 + \frac{\delta^2(2N^3-4N^2+N+1)}{12r(t)}\left[1+\frac{\delta}{r(t)}\right]\sigma_v^2\right\} \\ &\le \frac{1}{(N\alpha)^2}\limsup_{t\to\infty} E\left\{\frac{\left[2N^3-4N^2+7N+1\right]\delta^2}{3r(t)}\sigma_v^2\right\} \\ &\le \frac{n\left[2N^3-4N^2+7N+1\right](1-\lambda)\beta^2}{3\lambda^{N-1}\alpha^3}\sigma_v^2\end{aligned} \tag{56}$$

This completes the proof of Theorem 1.

5. ILLUSTRATION

Let $\hat{\theta}_G(n)$ be the estimated value of θ_G in the n-th test of the total M Monte Carlo tests. Denote by *err* the relative error of each step, $err = \left\|\hat{\theta}_G(n)-\theta_G\right\| / \left\|\theta_G\right\|$. The noise level is evaluated in terms of the noise-to-signal ratio (NSR),

$$\text{NSR} = \frac{mean(abs(noise))}{mean(abs(signal))} \tag{57}$$

Example 1: Consider the time delay system studied by Na, Ren, and Xia (2014),

$$\ddot{y} + 7.5\dot{y} + 12.5y = 4\dot{u}(t-0.4) + 12.5u(t-0.4) + v(t)$$

With a sampling interval of 0.1s, the discrete-time model is below

$$y(t) = \frac{0.317z^{-1} + 0.2398z^{-2}}{1-1.385z^{-1}+0.4724z^{-2}} z^{-4}u(t) + v(t)$$

For illustration, the input excitation is taken as the pseudorandom binary sequence (PRBS) with an amplitude of 1.0 and $v(t)$ is taken as a white noise sequence with zeromean and variance $\sigma_v^2 = 0.1$, which causes NSR=9.78%. In the proposed algorithm take $\hat{\theta}_G(0) = 10^{-3} I_n$, $r(0) = 1$, $\lambda_{\min} = 0.1$, $M = 3$ and the identification data $t \in [10, 1000]$. The parameter estimates and error criteria with respect to t are shown in Fig. 1. It is seen that the proposed AMVFSG algorithm gives obviously improved accuracy and faster convergence rate compared to the results given in the cited reference, which used an input excitation composed of multiple piecewise sinusoidals.

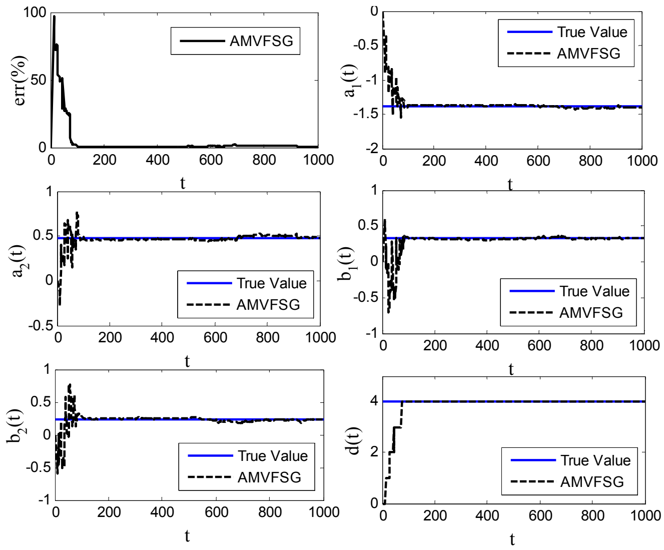

Fig.1 Parameter estimation results and errors for NSR=9.78%

To further demonstrate the identification effectiveness, one hundred Monte Carlo tests are performed with $NSR = 9.78\%$ and the total iterative steps of $N = 1000$ is adopted for each test. The computation results are listed in Table 1 where the estimation result for each parameter is shown by the mean value along with the standard deviation in parentheses, which demonstrate the effectiveness of the proposed algorithm.

Table 1. Identification results under one hundred Monte Carlo tests for NSR=9.78%

t	a_1	a_2	b_1	b_2	d	round(d)	err (%)
50	-1.2034 (±0.0716)	0.4670 (±0.0649)	-0.1447 (±0.1077)	0.1178 (±0.1142)	2.8633 (±0.2004)	3.0200 (±0.1407)	26.3001 (±2.8981)
100	-1.4008 (±0.1000)	0.4767 (±0.0812)	0.2778 (±0.1365)	0.2973 (±0.0920)	3.6181 (±0.1561)	3.8700 (±0.3380)	6.0583 (±7.8653)
500	-1.3793 (±0.0182)	0.4625 (±0.0239)	0.3250 (±0.0137)	0.2390 (±0.0184)	3.6903 (±0.1409)	4.0000 (±0.0000)	0.8193 (±0.4186)
1000	-1.3828 (±0.0132)	0.4685 (±0.0166)	0.3230 (±0.0107)	0.2397 (±0.0164)	3.7033 (±0.1414)	4.0000 (±0.0000)	0.6240 (±0.2770)
True	-1.3850	0.4724	0.3170	0.2398	4.0000	4.0000	

6. CONCLUSION

For the identification of industrial processes with time delay, a negative-gradient based iterative identification method has been proposed in this paper to obtain an output error model with a delay parameter. The key idea lies with using a feasible auxiliary model together with a variable forgetting factor to establish an iterative identification algorithm, which facilitates the convergence rate and identification accuracy. A strict proof on the convergence has been given. The application to an illustrative example from the literature has well demonstrated the effectiveness and merits 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 the Fundamental Research Funds for the Central Universities of China.

REFERENCES

Ahmed S., Huang, B., & Shah, S.L. (2007). Novel identification method from step response. Control Engineering Practice, 15, 545-556.

Atherton, D.P. (2006). Relay autotuning: An overview and alternative approach. Ind. Eng. Chem. Res., 45, 4075-4080.

Bedoui, S., Abderrahim, K., & Ltaief, M. (2013). Online identification of multivariable discrete time delay systems using a recursive least square algorithm. Mathematical Problems in Engineering, 1-18.

Bedoui, S., Ltaief, M., & Abderrahim, K. (2012).New results ondiscrete-time delay systems identification. International Journal of Automation and Computing, 9(6), 570-577.

Elnaggar, A., Dumont, G.A., & Elshafei, A.L. (1990). New method for delay estimation. Proceedings of the 29th IEEE Conference on Decision and Control, 3, 1929-1930.

Ferretti, G., Maffezzoni, C., & Scattolini, R. (1991). Recursive estimation of time delay in sampled systems. Automatica, 27(4), 653-661.

Gawthrop, P.J., Nihtila, M.T., & Rad, A.B. (1980). Recursive parameter estimation of continuous systems with unknown time-delay. Control Theory and Advanced Technology, 5(3), 227-248.

Gomez, O., Orlov, Y., & Kolmanovsky, I.V. (2007). Online identification of SISO linear time-invariant delay systems from output measurements. Automatica, 43(12), 2060-2069.

Goodwin, G.C., & Sin, K.S. (1984). Adaptive Filtering, Prediction and Control. Prentice–Hall, Englewood Cliffs, NJ.

Hang, C.C., Åström, K.J., & Wang, Q.G. (2002). Relay feedback auto-tuning of process controllers− A tutorial review. Journal of Process Control 12, 143-163.

Liu, T., & Gao, F. (2009). A generalized relay identification method for time delay and non-minimum phase processes. Automatica, 45 (4), 1072-1079.

Liu, T., & Gao, F. (2012). Industrial Process Identification and Control Design: Step-test and Relay-experiment-based Methods. Springer, London, UK.

Liu, T., & Gao, F. (2010). A frequency domain step response identification method for continuous-time processes with time delay. Journal of Process Control, 20(7), 800-809.

Liu, T., Wang, Q., & Huang, H. (2013). A tutorial review on process identification from step or relay feedback test. Journal of Process Control, 23(10), 1597-1623.

Liu, T., & Wang, Y. (2012) A synthetic approach for robust constrained iterative learning control of piecewise affine batch processes. Automatica, 48 (11), 2762-2775.

Ljung, L. (2002) .System Identification: Theory for the User, second edition. Prentice-Hall, Englewood Cliffs, New Jersey.

Loxton, R., Teo, K.L., & Rehbock, V. (2010). An optimization approach to state-delay identification. IEEE Transactions on Automatic Control, 55(9), 2113-2119.

Na, J., Ren, X.M., & Xia, Y.Q. (2014). Adaptive parameter identification of linear SISO systems with unknown time-delay. Systems & Control Letters, 66, 43-50.

Peter C. Young. (2011). Recursive Estimation and Time-Series Analysis: An introduction for the student and practitioner. Springer, London, UK.

Rad, A.B., Lo,W.L., & Tsang, K.M. (2003). Simultaneous on line identification of rational dynamics and time delay: A correlationbased approach. IEEE Transactions on Control Systems Technology, 11(6), 957-959.

Ren, X.M., Rad, A.B., & Chan, P.T. Lo, W.L. (2005). Online identification of continuous-time systems with unknown time delay. IEEE Transactions on Automatic Control, 50(9), 1418-1422.

Richard, J. (2003). Time-delay systems: an overview of some recent advances and open problems. Automatica, 39(10), 1667- 1694.

Söderström, T., & Stoica, P. (1989). System Identification. Prentice Hall, New York, USA.

Wang, Q.G., Liu, M., Hang, C.C., Zhang, Y., Zhang, Y., & Zheng, W.X. (2008). Integral identification of continuous-time delay systems in the presence of unknown initial conditions and disturbances from step tests, Ind. Eng. Chem. Res. 47 4929-4936.

Ding, F., & Chen. T. (2005). Parameter estimationof dual-rate stochastic systems by using an auxiliary model. IEEE Transactions on Automatic Control, 50(9), 1436-1441.

Ding, F., & Yang. J.B (1999). Remarks on Martingale hyperconvergence theorem and the convergence of forgetting factor leasrt squares algorithm. Control Theory and Applications, 16(4), 1-10.