amcs

RECURSIVE IDENTIFICATION ALGORITHM FOR DYNAMIC SYSTEMS WITH OUTPUT BACKLASH AND ITS CONVERGENCE

RUILI DONG *, QINGYUAN TAN **, YONGHONG TAN *

* College of Mechanical and Electronic Technology
Shanghai Normal University, ♯100 Guilin Road, Shanghai, 200234, China
e-mail: `tanyongh@yahoo.com.cn`

** School of Electronic, Information and Electrical Engineering
Shanghai Jiaotong University, Shanghai 200030, China

This paper proposes a recursive identification method for systems with output backlash that can be described by a pseudo-Wiener model. In this method, a novel description of the nonlinear part of the system, i.e., backlash, is developed. In this case, the nonlinear system is decomposed into a piecewise linearized model. Then, a modified recursive general identification algorithm (MRGIA) is employed to estimate the parameters of the proposed model. Furthermore, the convergence of the MRGIA for the pseudo-Wiener system with backlash is analysed. Finally, a numerical example is presented.

Keywords: nonlinear system, backlash, recursive identification, pseudo-Wiener model.

1. Introduction

Nonlinear systems with output backlash shown in Fig.1 usually exist in actuators of servo control systems, mechanical transmission systems, hydraulic control valves, etc. For example, in a mechanical transmission system, the space between the gear teeth often leads to the phenomenon of backlash. The existence of output backlash often leads to the deterioration of the system performance such as vibration, oscillation or even instability, supposing closed-loop control is implemented. Hence, in order to improve the system performance, some compensation for the effect of backlash should be considered.

Since most methods to compensate for the effect of backlash are model based (Chandler *et al.*, 2000; Selmic and Lewis, 2001; Toyozawa *et al.*, 2004; Campos *et al.*, 2000; Nordin and Gutman, 2002), the accuracy of the model to describe the behavior of systems with output backlash is very important for the performance of the compensation. In many cases, the operation environment may be changed with time due to the wear and tear of the components of the system. Thus, on-line identification of systems with backlash is significant for real-time compensation. Toyozawa *et al.* (2004) presented a method based on Volterra kernels to identify a backlash system. However, it is rather complicated to implement this algorithm
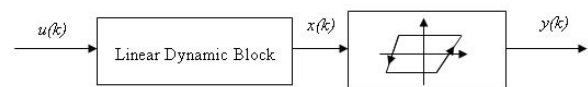

Fig. 1. Dynamic system with output backlash nonlinearity.
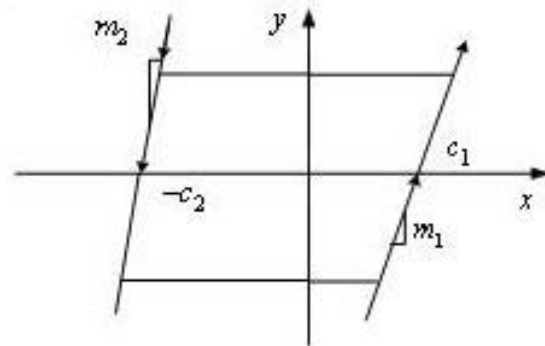

Fig. 2. Description of backlash.

in real-time.

It is known that a linear dynamic subsystem cascaded with an output nonlinear smooth function is usually called the Wiener system. Up to now, there have appeared in the literature some works concerning the identification of

the Wiener model with continuous static nonlinear functions (Wigren, 1997; Pearson and Pottmann, 2000; Grebicki, 1998; 2001; Celka *et al.,* 2001; Hu and Chen, 2005). Voros (2001) and Chen (2006) applied parametric approaches to the identification of Wiener models with discontinuous nonlinear functions. However, their research results only concern Wiener models containing discontinuous functions with one-to-one mappings. Moreover, those works did not discuss the convergence of identification algorithms for such systems. As backlash is a non-smooth function with multi-valued mapping, it is difficult to utilize the conventional identification methods to identify systems with output backlash since those methods can only be applied to systems with one-to-one mapping. Therefore, the identification of this non-smooth nonlinearity with multi-valued mapping in real-time is a challenge.

In this paper, a pseudo-Wiener system with backlash is defined, i.e., the static nonlinear function in a Wiener model is replaced by a non-smooth backlash function. Then, the so-called key term separation principle (Voros, 1995; 2001) is used to decompose the pseudo-Wiener model with backlash into a piecewise model with linear coefficients combining with nonlinear variables. Then, a modified recursive general identification algorithm (MRGIA) is applied to estimate all the unknown parameters of the model. After that, convergence analysis of the MRGIA for the pseudo-Wiener model with backlash in a noise-free case is presented. Finally, a numerical example is demonstrated.

2. Model of dynamic systems with output backlash nonlinearities

The model to describe the dynamic systems with backlash can be separated into two parts. One is linear and the other nonlinear (backlash nonlinearity). It is assumed that the linear subsystem is a stable and minimum phase system.

In this section, a pseudo-Wiener model is proposed based on the key terms separation principle. Thus, a piecewise linearized model is derived. Both parts of the model will be described in the following.

2.1. Description of the linear part of the model.
In Fig. 1, the linear dynamic block in discrete time can be described by

$$x(k) = -\sum_{i=1}^{n_a} a_i x(k-i) + \sum_{j=0}^{n_b} b_j u(k-d-j), \quad (1)$$

where n_a and n_b are the orders of the linear part of the system, d is the time-delay, $a_1, \ldots, a_{n_a}$ and $b_0, \ldots, b_{n_b}$ are the coefficients of the linear part of the model, $u(k)$ and $x(k)$ are the input and the output of the linear block, respectively. Suppose that n_a and n_b as well as d are given. Note that $x(k)$ cannot be measured directly.

2.2. Backlash model and a pseudo-Wiener model.
On the other hand, the nonlinear part of the system shown in Fig. 1 is backlash. Usually, it can be described by the architecture shown in Fig. 2. Suppose that the absolute values of the slopes of backlash are m_1 and m_2 ($0 < m_1 < \infty$ and $0 < m_2 < \infty$), respectively. Moreover, the absolute values of the dead zones in backlash are defined as c_1 and c_2 ($0 < c_1 < \infty$ and $0 < c_2 < \infty$), respectively. Assume that backlash can be described by non-smooth mapping with memory, i.e.,

$$y(k) = f(x(k), y(k-1)), \quad (2)$$

where $x(k)$ and $y(k)$ are the input and the output of backlash, respectively.

Define the internal variable, i.e., $m(k)$, as

$$m(k) = m_1 + (m_2 - m_1)g_3(k), \quad (3)$$

where

$$g_3(k) = \begin{cases} 0, & \Delta x(k) \geq 0, \\ 1, & \Delta x(k) < 0, \end{cases} \quad (4)$$

where $\Delta x(k) = x(k) - x(k-1)$.

According to the properties of backlash shown in Fig. 2, define $y_1(k)$ as an internal variable, i.e.,

$$\begin{aligned} y_1(k) = m(k)\{&x(k) - 0.5c_1[\text{sgn}(\Delta x(k)) + 1]g_1(k) \\ &+ 0.5c_2[-\text{sgn}(\Delta x(k)) + 1]g_2(k)\} \end{aligned} \quad (5)$$

where $\text{sgn}(\cdot)$ is the sign function; both $g_1(k)$ and $g_2(k)$ are the switching functions, respectively, defined by

$$g_1(k) = \begin{cases} 0, & \frac{y(k-1)}{m_1} + c_1 \geq x(k), \\ 1, & \frac{y(k-1)}{m_1} + c_1 < x(k) \end{cases} \quad (6)$$

and

$$g_2(k) = \begin{cases} 0, & \frac{y(k-1)}{m_2} - c_2 \leq x(k), \\ 1, & \frac{y(k-1)}{m_2} - c_2 > x(k), \end{cases} \quad (7)$$

where the region, i.e., $(\frac{y(k-1)}{m_2} - c_2, \frac{y(k-1)}{m_1} + c_1)$, can be defined as the memory zone of the backlash.

The switching functions, i.e., $g_1(k)$ and $g_2(k)$, can be switched between the two zones of both increasing and decreasing linear segments so as to obtain a one-to-one mapping within linear zones. The output $y(k)$ can be rewritten as

$$\begin{aligned} y(k) = &y_1(k) \\ &+ [y(k-1) - y_1(k)](g_1(k) - 1)(g_2(k) - 1). \end{aligned} \quad (8)$$

By substituting Eqn. (3) for the separated term into Eqn. (5) based on the key term separation principle,

$$\begin{aligned} y_1(k) = &m_1 x(k) + (m_2 - m_1)g_3(k)x(k) \\ &- 0.5c_1 m(k)[\text{sgn}(\Delta x(k)) + 1]g_1(k) \\ &+ 0.5c_2 m(k)[-\text{sgn}(\Delta x(k)) + 1]g_2(k). \end{aligned} \quad (9)$$

Re-substitute (9) into (8) to arrive at

$$
\begin{aligned}
y(k) &= m_1 x(k) + (m_2 - m_1)g_3(k)x(k) \\
&\quad - 0.5c_1 m(k)[\mathrm{sgn}(\Delta x(k)) + 1]g_1(k) \\
&\quad + 0.5c_2 m(k)[-\mathrm{sgn}(\Delta x(k)) + 1]g_2(k) \\
&\quad + [y(k-1) - y_1(k)](g_1(k) - 1)(g_2(k) - 1).
\end{aligned} \tag{10}
$$

Hence, the corresponding parameters of backlash can be determined by (3), (9) and (10). It is known that the switching functions $g_1(k)$ and $g_2(k)$ cannot be calculated directly. However, the switching functions $g_1(k)$ and $g_2(k)$, the internal variables, i.e., $m(k)$ and $y_1(k)$, can be estimated based on the estimated parameters of the model obtained in the previous sampling period.

Note that it is not possible for m_1 to be zero. Hence, substituting (1) into (10) yields

$$
\begin{aligned}
y(k) &= -\sum_{i=1}^{n_a} a_i m_1 x(k-i) + \sum_{j=0}^{n_b} m_1 b_j u(k-d-j) \\
&\quad + (m_2 - m_1)g_3(k)x(k) \\
&\quad - 0.5c_1 m(k)[\mathrm{sgn}(\Delta x(k)) + 1]g_1(k) \\
&\quad + 0.5c_2 m(k)[-\mathrm{sgn}(\Delta x(k)) + 1]g_2(k) \\
&\quad + [y(k-1) - y_1(k)](g_1(k) - 1)(g_2(k) - 1),
\end{aligned} \tag{11}
$$

where m_1 is set to unity for normalization. Thus, (1), (3), (9) and (11) are composed of the corresponding pseudo-Wiener model with backlash. In this model, $x(k)$ can be predicted based on the previously estimated parameters.

3. Recursive identification algorithm

In this section, a recursive identification approach is proposed to estimate the parameters of the above-mentioned pseudo-Wiener model with backlash. As the internal variables, i.e., $m(k)$, $y_1(k)$ and $x(k)$, cannot be measured directly, they will be estimated during the identification procedure.

Define

$$
\begin{aligned}
y_c(k) &= y(k) \\
&\quad + [y_1(k) - y(k-1)](g_1(k) - 1)(g_2(k) - 1)
\end{aligned} \tag{12}
$$

as the adjusted output of the model. Equation (12) can also be re-arranged as

$$
y_c(k) = h^T(k)\theta, \tag{13}
$$

where $h(k)$ denotes the data vector, i.e.,

$$
\begin{aligned}
h(k) &= [-x(k-1), \cdots, -x(k-n_a), u(k-d), \ldots, \\
&\quad u(k-d-n_b)g_3(k)x(k) - 0.5m(k)[\mathrm{sgn}(\Delta x(k)) \\
&\quad + 1]g_1(k), 0.5m(k)[-\mathrm{sgn}(\Delta x(k)) + 1]g_2(k)]^T
\end{aligned} \tag{14}
$$

and θ represents the parameter vector, i.e.,

$$
\theta = [a_1, \ldots, a_{n_a}, b_0, b_1, \ldots, m_2 - 1, c_1, c_2]^T. \tag{15}
$$

Note that the internal variables $m(k)$, $y_1(k)$ and $x(k)$ as well as the switching functions $g_1(k)$ and $g_2(k)$ all depend on the parameter vector θ. Therefore, $h(k) = h(k|\theta)$ and $y_c(k) = y_c(k|\theta)$ are defined.

Then, the RGIA (Fang, 2004) is applied to the estimation of the parameters of the proposed model. It is known that the RGIA method is usually available for linear and smooth dynamic systems. To handle the identification of a nonlinear and non-smooth system, the algorithm should be correspondingly modified. Suppose this algorithm is to minimize the following quadratic criterion:

$$
\hat{\theta}(k) = \arg\min_{\theta} \sum_{k=1}^{n} Q(k)[\hat{y}_c(k) - \hat{h}^T(k)\hat{\theta}(k-1)]^2, \tag{16}
$$

where the data vector $h(k)$ and the adjusted output $y_c(k)$ are respectively replaced by the corresponding $\hat{h}(k)$ and $\hat{y}_c(k)$, $Q(k) > 0$ is a weighted factor, and $Q(k) = \hat{\Sigma}^{-1}(k)$; n is the length of the data set. The MRGIA with the estimation of internal variables is summarized as follows:

$$
\hat{e}(k) = \hat{y}_c(k) - \hat{h}^T(k)\hat{\theta}(k-1), \tag{17}
$$

$$
\hat{\theta}(k) = \hat{\theta}(k-1) - K(k)\Phi(k)\hat{e}(k-1), \tag{18}
$$

$$
\begin{aligned}
S(k) &= S(k-1)(1 - \Phi(k)) \\
&\quad + \Phi(k)(\hat{h}^T(k)P(k-1)\hat{h}(k) + \mu(k)\Sigma(k)),
\end{aligned} \tag{19}
$$

$$
\begin{aligned}
K(k) &= K(k-1)(1 - \Phi(k)) \\
&\quad + \Phi(k)P(k-1)\hat{h}(k)S^{-1}(k),
\end{aligned} \tag{20}
$$

$$
\begin{aligned}
P(k) &= P(k-1)(1 - \Phi(k)) \\
&\quad + \Phi(k)\{\frac{1}{\mu(k)}\}[I - K(k)\hat{h}^T(k)]P(k-1) \\
&\quad \times [I - K(k)\hat{h}^T(k)]^T + K(k)\hat{\Sigma}(k)K^T(k)
\end{aligned} \tag{21}
$$

$$
\hat{\Sigma}(k) = \hat{\Sigma}(k-1) + \rho(k)\Phi(k)[\hat{e}^2(k) - \hat{\Sigma}(k-1)], \tag{22}
$$

$$
\Phi(k) = \begin{cases} 1, & (\hat{g}_1(k) - 1)(\hat{g}_2(k) - 1) = 0, \\ 0, & (\hat{g}_1(k) - 1)(\hat{g}_2(k) - 1) = 1, \end{cases} \tag{23}
$$

$$
\begin{aligned}
\hat{\theta}(k-1) \\
= [\hat{a}_1(k-1), \ldots, \hat{a}_{n_a}(k-1), \\
\hat{b}_0(k-1), \ldots, \hat{b}_{n_b}(k-1), \hat{m}_2(k-1) \\
- 1, \hat{c}_1(k-1), \hat{c}_2(k-1)]^T,
\end{aligned}
\tag{24}
$$

$$
\begin{aligned}
\hat{h}(k) \\
= [\hat{x}(k-1), \ldots, \hat{x}(k-n_a), \\
u(k-d), \ldots, u(k-d-n_b), \hat{g}_3(k)\hat{x}(k), \\
- 0.5\hat{m}(k)[\mathrm{sgn}(\Delta\hat{x}(k)) + 1]\hat{g}_1(k), \\
0.5\hat{m}(k)[-\mathrm{sgn}(\Delta\hat{x}(k)) + 1]\hat{g}_2(k)]^T,
\end{aligned}
\tag{25}
$$

$$
\hat{m}(k) = 1 + (\hat{m}_2(k-1) - 1)\hat{g}_3(k), \tag{26}
$$

$$
\hat{x}(k) = -\sum_{i=1}^{n_a} \hat{a}_i \hat{x}(k-i) + \sum_{j=0}^{n_b} \hat{b}_j u(k-d-j), \tag{27}
$$

$$
\begin{aligned}
\hat{y}_1(k) \\
= \hat{x}(k) + (\hat{m}_2(k-1) - 1)\hat{g}_3(k)\hat{x}(k) \\
- 0.5\hat{c}_1(k-1)\hat{m}(k-1)[\mathrm{sgn}(\Delta\hat{x}(k)) + 1]\hat{g}_1(k) \\
+ 0.5\hat{c}_2(k-1)\hat{m}(k-1)[-\mathrm{sgn}(\Delta\hat{x}(k)) + 1]\hat{g}_2(k),
\end{aligned}
\tag{28}
$$

$$
\hat{y}_c(k) = y(k) + [\hat{y}_1(k) - y(k-1)](\hat{g}_1(k) - 1)(\hat{g}_2(k) - 1), \tag{29}
$$

where $\rho(k)$ is the convergence factor, which not only satisfies the conditions given by (Fang, 2004) but is also located within (0, 1), and

$$
\mu(k) = \frac{\rho(k-1)}{\rho(k)}[1 - \rho(k)], \tag{30}
$$

$$
P(0) = \lambda I, \quad 0 < \lambda < \infty, \tag{31}
$$

$$
\hat{g}_1(k) = \begin{cases} 0, & y(k-1) + \hat{c}_1(k-1) \geq \hat{x}(k), \\ 1, & y(k-1) + \hat{c}_1(k-1) < \hat{x}(k), \end{cases} \tag{32}
$$

$$
\hat{g}_2(k) = \begin{cases} 0, & \frac{y(k-1)}{\hat{m}_2(k-1)} - \hat{c}_2(k-1) \leq \hat{x}(k), \\ 1, & \frac{y(k-1)}{\hat{m}_2(k-1)} - \hat{c}_2(k-1) > \hat{x}(k), \end{cases} \tag{33}
$$

and

$$
\hat{g}_3(k) = \begin{cases} 0, & \Delta\hat{x}(k) \geq 0, \\ 1, & \Delta\hat{x}(k) < 0. \end{cases} \tag{34}
$$

In order to reduce the large estimated error during the initial procedure of the identification, the covariance matrix $P(k)$ can be calculated by Eqn. (21) (Fang, 2004).

Remark 1. As $y_c(k)$ in this system cannot be measured directly, the estimate $\hat{y}_c(k)$ is implemented for approximation. Since the estimated error of $\hat{y}_1(k)$ can be convergent to zero as the identification procedure is going on, the error caused by $\hat{y}_c(k)$ depends on the switching error, i.e., $(\hat{g}_1(k) - 1)(\hat{g}_2(k) - 1)$, when backlash switches between the linear and memory zones.

Remark 2. In order to reduce the switching error, the input signal should be kept as small as possible, which ensures the estimated errors and estimated variables to be bounded. However, the amplitude of the input signal should be distributed within both the memory and linear zones of backlash so that the system can be fully excited, which means the amplitude of the input signal will cover all the equilibrium points of the system and the input signal should satisfy the persistently exciting conditions at each equilibrium point.

As backlash in the pseudo-Wiener model has memory zones, the output of backlash will be kept constant in those zones. Hence, $\hat{\theta}(k)$ as well as $P(k)$, $S(k)$, $\Sigma(k)$ and $K(k)$ are not updated when the estimate is implemented in the memory zones of backlash. The memory zones of backlash are its transient dead zones. Thus, (23) is used to implement those functions.

Furthermore, $y(k)$, the measured output, is introduced to adjust both the linear and memory zones of the backlash model since the change of the sign of $y(k)$ is the same as the change of the sign of the input in linear zones. Also, $y(k)$ is equal to $y(k-1)$ in the memory zones of backlash. In order to determine the linear and memory zones of the backlash model, the switching functions $g_4(k)$ and $g_5(k)$ are defined as

$$
g_4(k) = \begin{cases} 1, & y(k) - y(k-1) > 0, \\ 0, & \text{otherwise} \end{cases} \tag{35}
$$

and

$$
g_5(k) = \begin{cases} 1, & y(k-1) - y(k) > 0, \\ 0, & \text{otherwise}. \end{cases} \tag{36}
$$

If the output of the system is affected by a sequence of the zero mean white noise, the switching functions $g_4(k)$ and $g_5(k)$ can be respectively modified as

$$
g_4(k) = \begin{cases} 1, & y(k) - y(k-1) \geq \varepsilon, \\ 0, & \text{otherwise} \end{cases} \tag{37}
$$

and

$$
g_5(k) = \begin{cases} 1, & y(k-1) - y(k) \geq \varepsilon, \\ 0, & \text{otherwise}, \end{cases} \tag{38}
$$

where ε is the estimated mean value of the noise.

Hence, Eqns. (9), (11), (12) and (23) are respectively rewritten as

$$
\begin{aligned}
y_1(k) \\
= m_1 x(k) + (m_2 - m_1)g_3(k)x(k) \\
- 0.5c_1 m(k)[\mathrm{sgn}(\Delta x(k)) + 1]g_1(k)g_4(k) \\
+ 0.5c_2 m(k)[-\mathrm{sgn}(\Delta x(k)) + 1]g_2(k)g_5(k),
\end{aligned}
\tag{39}
$$

$y(k)$

$$= -\sum_{i=1}^{n_a} a_i m_1 x(k-i) + \sum_{j=0}^{n_b} m_1 b_j u(k-d-j)$$

$$+ (m_2 - m_1) g_3(k) x(k) - 0.5 c_1 m(k) \qquad (40)$$

$$\times [\text{sgn}(\Delta x(k)) + 1] g_1(k) g_4(k)$$

$$+ 0.5 c_2 m(k) [-\text{sgn}(\Delta x(k)) + 1] g_2(k) g_5(k)$$

$$+ [y(k-1) - y1(k)]$$

$$\times (g_1(k) g_4(k) - 1)(g_2(k) g_5(k) - 1),$$

$$y_c(k) = y(k) + [y1(k) - y(k-1)]$$

$$\times (g_1(k) g_4(k) - 1)(g_2(k) g_5(k) - 1), \qquad (41)$$

and

$$\Phi(k) = \begin{cases} 1, & (\hat{g}_1(k)\hat{g}_4(k) - 1)(\hat{g}_2(k)\hat{g}_5(k) - 1) = 0, \\ 0, & (\hat{g}_1(k)\hat{g}_4(k) - 1)(\hat{g}_2(k)\hat{g}_5(k) - 1) = 1. \end{cases}$$

$$(42)$$

Equations (39)–(42) will greatly reduce the switching errors compared with Eqns. (9), (11), (12) and (23) since the output of backlash can be measured in this case.

Definition 1. (Boutayeb and Darouach, 1995) Let $\hat{h}^T(k)[\hat{\theta}(k-1) - \theta] = -\alpha(k)\hat{e}(k)$, where θ is the true value vector of the model parameters. Moreover, $\alpha(k)$ implies that the identification is affected by the estimated errors of the internal variables.

As the linear part of the system is BIBO stable, $x(k)$ is bounded for any bounded input $u(k)$. Hence $y(k)$ is also bounded based on Eqns.(3), (9) and (11). That means the pseudo-Wiener model with backlash is BIBO stable. Thus, a recursive approach can be applied to identification (Fang, 2004; Ljung, 1977a; Ljung, 1977b).

In the following, convergence of the algorithm of recursive identification for systems with output backlash in the case of the absence of noise will be analysed.

Lemma 1. *Assume that an input signal $u(k)$ is bounded but can fully excite the system. Then it satisfies*

$$\gamma \leq \frac{1}{n} \sum_{k=1}^{n} \hat{h}^T(k)\hat{h}(k) \leq \beta,$$

where both γ and β are finite positive real numbers.

Proof. Assume that the pseudo-Wiener model with backlash is BIBO stable. It is obvious that the input signal $u(k)$ is bounded. If Eqns. (1), (3) and (9) hold, the elements of the estimated data vector, i.e., $\hat{h}(k)$, are bounded. This implies that there exist two positive real numbers γ and β, which lead to $\gamma \leq \frac{1}{n}\sum_{k=1}^{n} \hat{h}^T(k)\hat{h}(k) \leq \beta$. ∎

Theorem 1. *For the algorithm described by Eqns. (17)–(29), which is applied to parameter estimation of the pseudo-Wiener model given by Eqns. (1), (3), (9) and (11), assume that the following conditions are satisfied:*

1. The condition of Lemma 1 is met.

2. $P(k)$ is a positive definite matrix and $\lim_{k \to \infty} \lambda_{\min}(\rho(k)P^{-1}(k)) \to \varpi$, where ϖ is a positive number.

3. Moreover,

$$\lim_{k \to \infty} \lambda_{\max} P(k) \to 0, \qquad (43)$$

$$\limsup_{k \to \infty} \frac{\lambda_{\max} P(k)}{\lambda_{\min} P(k)} < \infty, \qquad (44)$$

and

$$1 - \sqrt{\triangle_1(k)} \leq \alpha(k) \leq 1 + \sqrt{\triangle_1(k)}. \qquad (45)$$

Then the estimated parameter vector $\hat{\theta}(k)$ will converge to θ, the true value of the parameter vector, as $k \to \infty$.

In the theorem, $\lambda_{\max} P(k)$ and $\lambda_{\min} P(k)$ are the maximum and minimum eigenvalues of $P(k)$, respectively. In the formula (45), $\triangle_1(k)$ is equal to

$$1 - \frac{h^T(k)P(k-1)h(k)}{h^T(k)P(k-1)h(k) + \Sigma(k)}.$$

The proof of this theorem is given in Appendix.

Remark 3. If the estimated errors of the internal variables converge to zeros, $\alpha(k)$ would converge to unity. Also, the approach is available only if $\alpha(k)$ satisfies the conditions given by (45).

Hence, choosing the proper input signal and embedding the switching functions will help to decrease the switching errors between the model zones. It will largely reduce the estimated errors of the internal variables. Moreover, from Section 2, it is known that the absolute values of the dead zones of backlash, i.e., c_1 and c_2, are not equal to zero. Therefore, the initial values of c_1 and c_2 are set to very small positive numbers.

4. Numerical examples

In this section, the proposed approach is used to identify a pseudo-Wiener system with backlash based only on the measured system input and output. Suppose that the parameters of backlash in the system are $m_1 = 1, m_2 = 1.5$, $c_1 = 0.5, c_2 = 1$, and the linear part of the system is described by

$$x(k) = -0.45x(k-1) - 0.55x(k-2) + 0.8u(k-1). \quad (46)$$

This implies $a_1 = 0.45$, $a_2 = 0.55$, and $b_0 = 0.8$. Select a uniformly bounded distributed random input sequence, i.e., $u(k)$, to ensure most of the output data are within the linear zones of backlash, especially in the initialized period of the identification. Hence, the corresponding input sequence of the bounded white noise with the variance $\sigma^2 = 1.93$ is chosen in this example. The initial variables are respectively chosen as
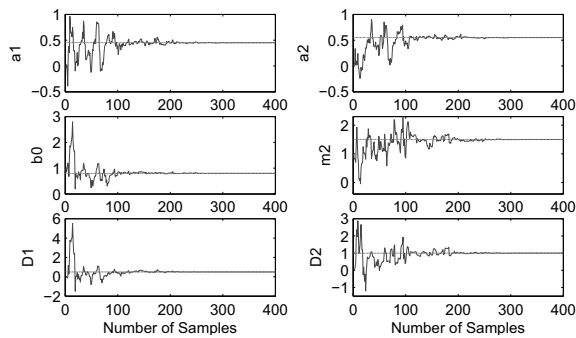
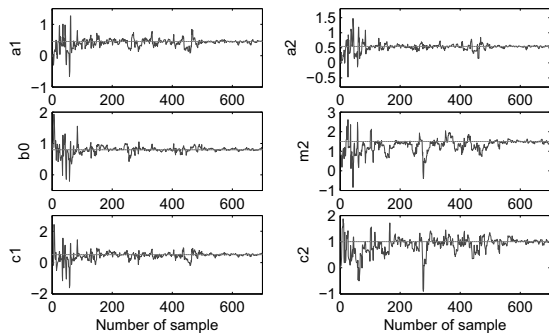Fig. 3. Parameter estimation of the model (noise-free case).

Fig. 4. Parameter estimation of the model (noisy case).

Fig. 5. Auto-correlation error of the estimated parameters (noisy case).

5. Conclusions

In this paper, a pseudo-Wiener model with backlash was proposed for on-line identification of dynamic systems with backlash. In this scheme, a separate technique was applied to decompose nonlinear backlash into a group of piecewise linearized dynamic functions. Thus, the MR-GIA can be employed for the estimation of model parameters. Then, the convergence analysis of the MRGIA for the identification of the presented non-smooth nonlinear system was discussed.

Numerical examples were presented to illustrate the performance of the proposed identification method. The properties of on-line identification as well as the technique to handle the estimate of the non-smooth nonlinear systems are promising for practical engineering.

$P(0) = 1.0 \times 10^6 I$, where I is the identity matrix, $\theta(0) = [0, 0, 1, 0, 0, 0]^T$ and $\Sigma(0) = 1$. The convergence factor $\rho(k)$ is set to $(k + 1)^{-0.25}$. Only when the mean square error (MSE) becomes less than the expected value, the iteration stops. The obtained estimation results are shown in Fig. 3. If the systems contains a measurement disturbance which is a Gaussian white noise, the proposed technique can also be used to handle parameter estimation of noisy systems. Suppose that the system parameters as well as the initial parameter values and the input signal are the same as in the noise-free case shown above. In this case, the signal to noise ratio (SNR), which means the square root of the ratio of the output and noise variance, is 43. The results of the corresponding parameter estimation for the noisy system are shown in Fig. 4. The maximum relative error of the parameters is 2.62×10^{-2}. In Fig. 5, $\sigma(a_1), \sigma(a_2), \sigma(b_0), \sigma(m_2), \sigma(c_1)$ and $\sigma(c_2)$ are the auto-correlation errors of the estimated parameters. It can be seen that, after 600 steps, all the auto-correlation errors gradually converge to zeros.

From the presented numerical examples, it is known that the proposed on-line identification method is promising for the modeling of nonlinear dynamic systems with output backlash.
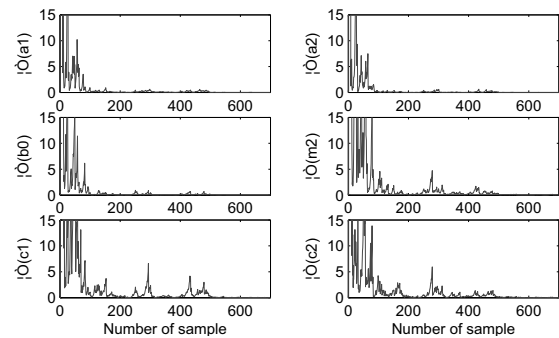
References

Boutayeb, M. and Darouach, M. (1995). Recursive identification method for MISO Wiener–Hammerstein model, *IEEE Transactions on Automatic Control* **40**(2): 287–291.

Campos, J., Lewis, F. L. and Selmic, R. (2000). Backlash compensation in discrete time nonlinear systems using dynamic inversion by neural networks, *International Conference on Automation and Robotics Research, San Francisco, CA, USA*, pp. 1289–1295.

Celka, P., Bershad, N. J. and Vesin, J. M. (2001). Stochastic gradient identification of polynomial Wiener systems: Analysis and application, *IEEE Transactions on Signal Processing* **49**(2): 301–313.

Chandler, H. W., George, S. D., Liddle, J. and Nixon, S. A. (2000). Backlash: A feature of the cyclic deformation of ceramic pastes, *Journal of the European Ceramic Society* **20**(11): 1699–1705.

Chen, H. F. (2006). Recursive identification algorithms for Wiener model with discontinuous piecewise-linear linear function, *IEEE Transactions on Automatic Control* **51**(3): 390–400.

Fang, C. Z. (2004). *System Identification*, Qinghua University, Beijing, (in Chinese).

Greblicki, W. (1998). Continuous-time Wiener system identification, *IEEE Transactions on Automatic Control* **43**(10): 1488–1493.

Greblicki, W. (2001). Recursive identification of Wiener systems, *International Journal of Applied Mathematics and Computer Science* **11**(4): 977–991.

Hu, X. L. and Chen, H. F. (2005). Strong consistence of recursive identification for Wiener systems, *Automatica* **41**(11): 1905–1916.

Ljung, L. (1977a). Analysis of recursive stochastic algorithms, *IEEE Transactions on Automatic Control* **AC-22**(4): 551–575.

Ljung, L. (1977b). On positive real transfer functions and the convergence of some recursive schemes, *IEEE Transactions on Automatic Control* **AC-22**(4): 539–551.

Nordin, M. and Gutman, P.-O. (2002). Controlling mechanical systems with backlash—A survey, *Automatica* **38**(10): 1633–1649.

Pearson, R. K. and Pottmann, M. (2000). Gray-box identification of block-oriented nonlinear models, *Journal of Process Control* **10**(4): 301–315.

Selmic, R. R. and Lewis, F. L. (2001). Neural net backlash compensation with Hebbian tuning using dynamic inversion, *Automatica* **37**(8): 1269–1277.

Toyozawa, Y., Kashiwag, H. and Harada, H. (2004). Identification of Volterra kernels of nonlinear system having backlash type nonlinearity, *30th Annual Conference of the IEEE Industrial Electronics Society, Busan, Korea*, pp. 2798–2802.

Voros, J. (1995). Identification of nonlinear dynamic system using extended Hammerstein and Wiener models, *Control-Theory Advanced Technology* **10**(4): 1203–1212.

Voros, J. (2001). Parameter identification of Wiener systems with discontinuous nonlinearities, *System and Control Letters* **44**(5): 363–372.

Wigren, T. (1997). Circle criteria in recursive identification, *IEEE Transactions on Automatic Control* **42**(7): 975–979.

Ruili Dong received both a B.S. degree in mechanical engineering and an M.S. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China. In 2009, she obtained a Ph.D. degree from the Department of Automation, Shanghai Jiaotong University, China. Presently, she holds a lecturer position at the College of Mechanical and Electronic Engineering, Shanghai Normal University. Her research interests are in modeling and control for nonlinear dynamic systems.

Qingyuan Tan is now a student at the Department of Instrumentation Engineering, School of Electronic, Information and Electrical Engineering, Shanghai Jiaotong University, China. His research interests cover process sensing, signal processing, modeling and control.

Yonghong Tan received his B.S. and M.S. degrees from the South China University of Technology, Guangzhou, China. In 1996, he obtained the Ph.D. degree in electrical engineering from the University of Ghent, Belgium. Then, he was a post-doctoral fellow at Simon Fraser University, Canada. Dr. Tan has been a visiting professor, respectively, at Colorado State University, USA, and Concordia University, Canada, and a professor at the Guilin University of Electronic Technology and the University of Electronic Science and Technology of China. Presently, he is a full professor at the College of Mechanical and Electronic Engineering, Shanghai Normal University, China. His research interests are modeling and control of nonlinear systems, biomedical signal processing and intelligent control. (Y. Tan is the correspondence author for this paper; e-mail: `tanyongh@yahoo.com.cn`, phone: +86 21 57122 599.)

Appendix

Consider two cases.

Case 1. Suppose that $\Phi(k)$ is equal to unity. From (19)–(22), we obtain

$$P(k)\hat{h}(k) = \Sigma(k)P(k-1)\hat{h}(k)S^{-1}(k) \qquad (47)$$

and

$$\rho(k)P^{-1}(k) = \rho(k-1)[1 - \rho(k)]P^{-1}(k-1) + \hat{h}(k)\frac{\rho(k)}{\Sigma(k)}\hat{h}^T(k). \qquad (48)$$

It is known that $\rho(k)P^{-1}(k)$ is a positive definite matrix. It not only converges to a nonzero matrix but is also a bounded matrix as $k \rightarrow \infty$ (Fang, 2004; Ljung, 1977a; 1977b). Hence, a corresponding quadratic function can be defined as

$$V(k) = \tilde{\theta}^T(k)\rho(k)P^{-1}(k)\tilde{\theta}(k), \qquad (49)$$

where $\tilde{\theta}(k) = \hat{\theta}(k) - \theta$. Subtracting θ from both the sides of Eq. (18) yields

$$\tilde{\theta}(k) = \tilde{\theta}(k-1) + K(k)\hat{e}(k). \tag{50}$$

Then substituting (50) into (49) gives

$$\begin{aligned} V(k) = (\tilde{\theta}(k-1) + K(k)\hat{e}(k))^T \rho(k) P^{-1}(k) \\ \times \tilde{\theta}(k-1) + K(k)\hat{e}(k)). \end{aligned} \tag{51}$$

Equation (51) can also be rewritten as

$$\begin{aligned} V(k) \\ = \tilde{\theta}^T(k-1)\rho(k)P^{-1}(k)\tilde{\theta}(k-1) \\ + 2\tilde{\theta}^T(k-1)\rho(k)P^{-1}(k)K(k)\hat{e}(k) \\ + K^T(k)\hat{e}(k)\rho(k)P^{-1}(k)K(k)\hat{e}(k). \end{aligned} \tag{52}$$

Furthermore, substitute (20), (47) and (48) into (52), which results in

$$\begin{aligned} V(k) \\ = \tilde{\theta}^T(k-1)\rho(k-1)[1 - \rho(k)]P^{-1}(k-1)\tilde{\theta}(k-1) \\ + \tilde{\theta}^T(k-1)\hat{h}(k)\frac{\rho(k)}{\Sigma(k)}\hat{h}^T(k)\tilde{\theta}(k-1) \\ + \frac{2\tilde{\theta}^T(k-1)\rho(k)\hat{h}(k)\hat{e}(k)}{\Sigma(k)} \\ + \frac{\hat{h}^T(k)P(k)\hat{h}(k)\rho(k)\hat{e}^2(k)}{\Sigma^2(k)}. \end{aligned} \tag{53}$$

From (47) and (49), Eqn. (53) can be rewritten as

$$\begin{aligned} V(k) = (1 - \rho(k))V(k-1) \\ + \tilde{\theta}^T(k-1)\hat{h}(k)\frac{\rho(k)}{\Sigma(k)}\hat{h}^T(k)\tilde{\theta}(k-1) \\ + \frac{2\tilde{\theta}^T(k-1)\rho(k)}{\Sigma(k)} + \frac{\hat{e}(k)\hat{h}(k)}{\Sigma(k)} \\ + \frac{\hat{h}^T(k)P(k-1)\hat{h}(k)\rho(k)\hat{e}^2(k)}{\Sigma(k)S(k)}. \end{aligned} \tag{54}$$

It is easy to see that

$$1 - \rho(k) < 1. \tag{55}$$

Based on the definition in Section 3, we get

$$\hat{h}^T(k)\tilde{\theta}(k-1) = -\alpha(k)\hat{e}(k). \tag{56}$$

From (21), (54), (55) and (56), we get

$$\begin{aligned} V(k) \\ < V(k-1) + \frac{\rho(k)\alpha^2(k)e^2(k)}{\Sigma(k)} - \frac{2\rho(k)}{\Sigma(k)} \\ - \frac{\hat{e}^2(k)\alpha(k)}{\Sigma(k)} + \frac{\hat{h}^T(k)P(k-1)\hat{h}(k)\rho(k)\hat{e}^2(k)}{\Sigma(k)(\hat{h}^T(k)P(k-1)\hat{h}(k) + \Sigma(k))}. \end{aligned} \tag{57}$$

Define

$$\begin{aligned} t(k) = \frac{\rho(k)\alpha^2(k)e^2(k)}{\Sigma(k)} - \frac{2\rho(k)\hat{e}^2(k)\alpha(k)}{\Sigma(k)} \\ + \frac{\hat{h}^T(k)P(k-1)\hat{h}(k)\rho(k)\hat{e}^2(k)}{\Sigma(k)(\hat{h}^T(k)P(k-1)\hat{h}(k) + \Sigma(k))}. \end{aligned} \tag{58}$$

Also, assume that (43) and (44) hold and $\alpha(k)$ satisfies (45). Thus, this leads to $t(k) \leq 0$, or

$$V(k) - V(k-1) < 0. \tag{59}$$

In this case, $\rho(k)P^{-1}(k)$ is a positive matrix. Therefore,

$$V(k) \geq 0, \quad k = 1, 2, \dots. \tag{60}$$

According to (59) and (60), based on the Lyapunov stability theorem, it can be concluded that $V(k)$ converges to zero as $k \to \infty$. Then, we obtain

$$\left\| \tilde{\theta}(k) \right\|^2 \leq \frac{V(k)}{\lambda_{\min}(\rho(k)P^{-1}(k))}. \tag{61}$$

Hence, we have

$$\lim_{k \to \infty} \hat{\theta}(k) = \theta. \tag{62}$$

Case 2. Suppose that $\Phi(k)$ is equal to zero. From (51), it follows that

$$V(k) = (\tilde{\theta}(k-1))^T \rho(k)P^{-1}(k)(\tilde{\theta}(k-1)). \tag{63}$$

Based on (49), $P^{-1}(k)$ is equal to $P^{-1}(k-1)$ and $\rho(k)$ decreases as k increases, so this yields

$$V(k) < V(k-1). \tag{64}$$

Thus, it can be seen that, although there exist dead zones in the system, based on (64), the formula (62) can still be obtained.