Journal of Process Control 24 (2014) 1–12

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Detection of asymmetric control valve stiction from oscillatory data using an extended Hammerstein system identification method☆

Jiandong Wang[a,*], Qinghua Zhang[b]

[a] College of Engineering, Peking University, Beijing, China
[b] INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France

ARTICLE INFO

Article history:
Received 11 March 2013
Received in revised form 7 October 2013
Accepted 14 October 2013
Available online 20 November 2013

Keywords:
Control valve stiction
Oscillation
Hammerstein system identification
Cyclo-stationarity

ABSTRACT

The study in this paper is motivated by the detection of control valves with asymmetric stiction resulting in oscillations in feedback control loops. The joint characterization of the control valve and the controlled process is formulated as the identification of a class of extended Hammerstein systems. The input nonlinearity is described by a point-slope-based hysteretic model with two possibly asymmetric ascent and descent paths. An iterative identification method is proposed, based on the idea of separating the ascent and descent paths subject to the oscillatory input and output. The structure of the formulated extended Hammerstein system is shown to be identifiable, and the oscillatory signals in feedback control loops are proved to be informative by exploiting the cyclo-stationarity of these oscillatory signals. Numerical, experimental and industrial examples are provided to illustrate the effectiveness of the proposed identification method.

1. Introduction

Oscillation is one of the most common abnormal phenomena in process industries, and may lead to fluctuation of product quality, increased consumption of energy and raw materials, and compromised stability and safety [32]. One of the common causes for oscillation is the control valve nonlinearity [9]. Industrial surveys in [5,11,28] indicated that about 20–30% of control loops in various process industries oscillate due to control valve nonlinearities, e.g., stiction, backlash and dead band. Among various nonlinearities, stiction is the most commonly encountered one. Hence, it is practically very important to detect the presence of control valve stiction and to compensate its negative effect of resulting in oscillations in a single control loop or even loops throughout a whole industrial plant. To deal with the stiction problem, the very first step is to detect whether a control valve is sticky and to quantify the severity of the stiction if it exists.

The oscillation caused by control valve stiction is illustrated as follows. In a feedback control loop depicted in Fig. 1, $G(s)$ is a continuous-time process to be controlled, $f(\cdot)$ a control valve, $C(z)$ a discrete-time controller, H the zero-order holder and S the sampler. The signals y, u, r, v and x are the process output, the controller output, the setpoint, the process noise and the valve position, respectively. If the control valve $f(\cdot)$ is sticky, then its stiction may lead to oscillations in the loop even if the setpoint r stays with a constant value. Such a typical example is shown in Fig. 2. The data is obtained from an experiment performed at a laboratory at the Peking University. The experiment configuration is the same as that in Fig. 1. In the experiment, the process $G(s)$ is a water tank system and the water level of the tank is selected as the process output y. The opening of the outlet valve is fixed, while the inlet flow is controlled by a control valve. The controller $C(z)$ is the discretized counterpart of a continuous-time proportional-integral (PI) controller $C(s) = 1 + (1/10s)$, with the sampling period 0.5 s. There is no other disturbance entering into the tank, and the oscillations appear due to the control valve stiction, which is clearly revealed by plotting the measured valve position x against the controller output u in Fig. 3. With the measurement of x, it is straightforward to detect the presence of stiction. In practice, however, many industrial control valves do not provide the measurement of valve position so that the detection of control valve stiction is not an easy task.

The detection of control valve stiction have been very active research topics recently; see the collected results in [22]. The existing methods can be classified into invasive and non-invasive methods. Their main difference is whether the methods reply upon special field tests or only require normal operating data. The invasive method can provide conclusive results; however, special field tests may disturb the normal operation of processes and are often time consuming for implementation. By contrast, the non-invasive method is easier for implementation and more suitable for a large

☆ This research was partially supported by the National Natural Science Foundation of China under grants No. 61074105 and No. 61061130559, and by the French ANR EBONSI project.

* Corresponding author. Tel.: +86 10 6275 3856.
E-mail addresses: jiandong@pku.edu.cn (J. Wang), qinghua.zhang@inria.fr (Q. Zhang).

0959-1524/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jprocont.2013.10.012

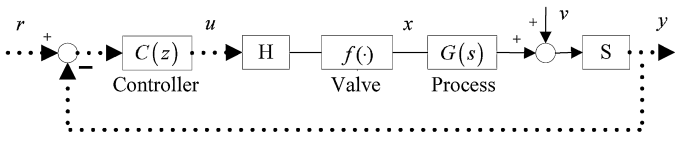

Fig. 1. The diagram of a feedback control loop.

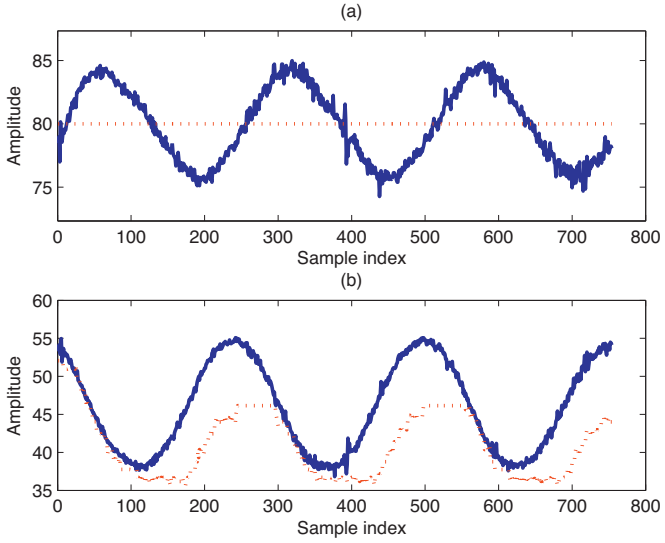

Fig. 2. Control-loop oscillation caused by control valve stiction in a laboratory experiment: (a) $y(t)$ (solid) and $r(t)$ (dash), (b) $u(t)$ (solid) and $x(t)$ (dash)).

number of control loops. In practice, a combined approach can be adopted, i.e., the non-invasive method provides a preliminary analysis to pick up control loops possibly having control valve stiction problems, for which the invasive method is applied to make a final validation. Among various non-invasive methods, the methods based on Hammerstein system identification perhaps have the best performance, as shown in Chapter 13 of [22], and has been studied by many researches [31,21,23,29,14]. In all of these cited works, the control valve stiction is represented by some special data-driven stiction models (see, e.g., Chapters 2 and 3 in [22]). However, the characters of many sticky control valves in practice may be much more complicated than these special models. For instance, some control valve demonstrates an asymmetric characteristic, namely,

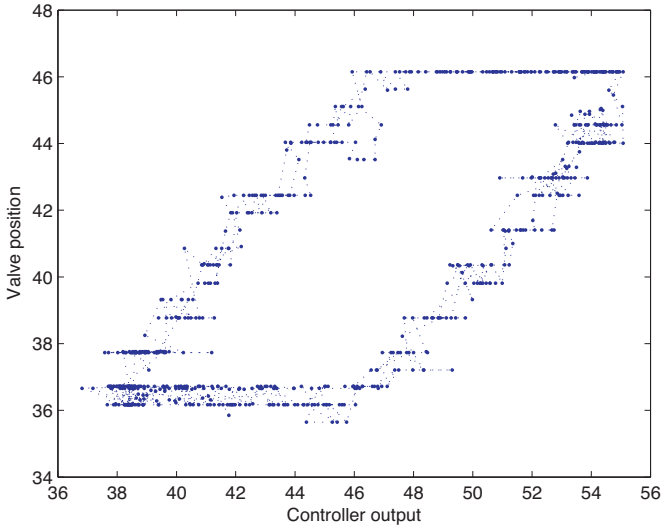

Fig. 3. The characteristic of the control valve in a laboratory experiment.

the stiction exists only at the occasions that the valve is opening (the controller output is increasing), while the valve position follows quickly with the controller output when it is decreasing. Some data-driven stiction models, such as the one proposed by Choudhury et al. [8], could be adapted by using two sets of stiction parameters for different directions of valve movements. However, many data-driven stiction models in the literature could not capture the complicated characters of control valves such as the asymmetric property and thus would lead to a large model mismatch (to be illustrated via an industrial example later in Section 5). Therefore, it is desirable to use a model with more flexible structure in the non-invasive method based on Hammerstein system identification. Note that for the purpose of detecting control valve stiction, it is adequate to use a model capable of describing the valve movements when u and y in Fig. 1 are oscillatory, and the model does not need to have the extrapolation ability like data-driven stiction models.

Hammerstein systems are usually composed of static (memoryless) nonlinearities and linear dynamic components. However, in order to detect the control valve stiction, Hammerstein systems needs to be extended: the input nonlinearity has a hysteretic behavior, instead of being memoryless. Such systems are referred to as *extended Hammerstein systems* in this paper. It is worthy to point out that some recent articles [2,6,7,10,13,17,18,35,38] studied the identification of Hammerstein systems with non-static input nonlinearities. However, these studies are not designed for describing control valve stiction based on oscillatory signals, and consequently are hardly applicable here.

In this paper, we first provide a novel formulation for a class of extended Hammerstein systems with hysteretic input nonlinearity. The novelty lies at the use of a point-slope-based hysteretic model for input nonlinearity, composed by two ascent and descent paths. The input nonlinearity model has a flexible structure to describe the asymmetric control valve stiction. Second, an iterative method is proposed to identify the extended Hammerstein systems. The basic idea is to separate the identification of the ascent and descent paths of the input nonlinearity under the assumption that the input and output signals are oscillatory with two extreme values. Finally, identifiability analysis is performed to show that the proposed extended Hammerstein system structure is identifiable, and the oscillatory signals in feedback control loops are informative owing to the cyclo-stationarity of these oscillatory signals.

Some preliminary results of this work have been presented in [39]. The present paper is more complete in the identifiability analysis and provides an improved formulation of the proposed method.

The rest of the paper is organized as follows. Section 2 formulates the identification problem for the extended Hammerstein systems. Section 3 provides an iterative identification method. The identifiability analysis is presented in Section 4. Section 5 gives numerical, experimental and industrial examples to illustrate the effectiveness of the proposed method. A conclusion is made in Section 6.

2. Problem formulation

Consider an extended Hammerstein system depicted in Fig. 4. The linear time-invariant (LTI) dynamic block is described by an auto-regression with extra input (ARX) model,

$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j x(t-j-n_k) + d + e(t). \tag{1}$$

Here $y(t)$ is the process output, $x(t)$ is the inner signal, and $e(t)$ is a zero-mean stationary white noise; $x(t)$ and $e(t)$ are not measurable. Symbols a_i's, b_j's and d are the unknown parameters to be estimated, while n_a, n_b and n_k are the structure parameters to be

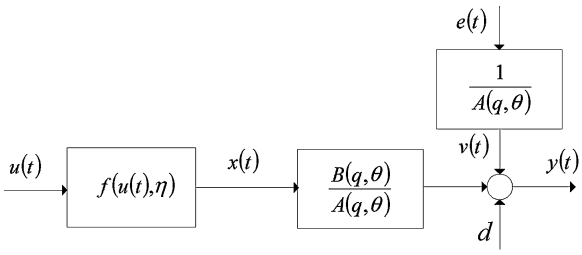

Fig. 4. The diagram of an extended Hammerstein system.

estimated. The unknown parameter d is resulted from the effect of non-zero means of $y(t)$ and $x(t)$. It is a well-known fact that the estimation of ARX models only requires the standard linear least-squares method, and a high-order ARX model is capable of approximating any LTI dynamic systems arbitrarily well [25] (page 336). A two-step estimation procedure, consisted of a high-order ARX model estimation and a subsequent model reduction, can yield very accurate model estimation even if the measurement noise is heavily colored (see, e.g., Chapter 6 of [40] and [33]). Hence, the ARX model is used here to describe the linear dynamics, while the subsequent model reduction step is omitted since the estimation of the linear part is not the main objective in this context.

The input nonlinearity $f(\cdot)$ is assumed to be hysteretic with two asymmetric ascent and descent paths. By adopting a so-called point-slope model [34] for each of the ascend and descend paths, we propose an input nonlinearity model as

$$x(t) = f(u(t)) = \begin{cases} \eta_A^T(u(t))g_A, & \text{for } h(t) = 1 \\ \eta_D^T(u(t))g_D, & \text{for } h(t) = 0 \end{cases}. \tag{2}$$

Here the direction signal $h(t)$ is generated to indicate that the input $u(t)$ is increasing or decreasing, i.e.,

$$h(t) = \begin{cases} 1, & \text{if } u(t) > u(t-1) \\ h(t-1), & \text{if } u(t) = u(t-1) \\ 0, & \text{if } u(t) < u(t-1) \end{cases}. \tag{3}$$

The other symbols in (2) are defined as

$$\begin{aligned} g_A^T &= \begin{bmatrix} g_1^A & g_2^A & \cdots & g_{p_A-1}^A & k_A \end{bmatrix}, \\ g_D^T &= \begin{bmatrix} g_1^D & g_2^D & \cdots & g_{p_D-1}^D & k_D \end{bmatrix}, \\ \eta_A^T(u(t)) &= \begin{bmatrix} \eta_1^A(u(t)) & \eta_2^A(u(t)) & \cdots & \eta_{p_A-1}^A(u(t)) & 1 \end{bmatrix}, \\ \eta_D^T(u(t)) &= \begin{bmatrix} \eta_1^D(u(t)) & \eta_2^D(u(t)) & \cdots & \eta_{p_D-1}^D(u(t)) & 1 \end{bmatrix}, \end{aligned}$$

with

$$\eta_i^A(u(t)) = \begin{cases} 0, & \text{if } u(t) < c_i^A \\ u(t) - c_i^A, & \text{if } u(t) \in [c_i^A, c_{i+1}^A) \\ c_{i+1}^A - c_i^A, & \text{if } u(t) \geq c_{i+1}^A \end{cases},$$

$$\eta_j^D(u(t)) = \begin{cases} 0, & \text{if } u(t) < c_j^D \\ u(t) - c_j^D, & \text{if } u(t) \in [c_j^D, c_{j+1}^D) \\ c_{j+1}^D - c_j^D, & \text{if } u(t) \geq c_{j+1}^D \end{cases},$$

for $i \in [1,\ p_A - 1]$ and $j \in [1,\ p_D - 1]$. Here $\begin{bmatrix} c_1^A & c_2^A & \cdots & c_{p_A}^A \end{bmatrix}$ for $c_1^A < c_2^A < \cdots < c_{p_A}^A$ are p_A knots of $u(t)$ for the ascent path, while $\begin{bmatrix} c_1^D & c_2^D & \cdots & c_{p_D}^D \end{bmatrix}$ for $c_1^D < c_2^D < \cdots < c_{p_D}^D$ are the p_D knots of $u(t)$ for the descent path. The parameter g_i^A in (2) for $i \in [1, p_A - 1]$ is the slope of the straight line connecting the knots c_i^A and c_{i+1}^A in the ascent path, while the parameter g_j^D for $j \in [1, p_D - 1]$ is the slope of the straight line connecting the knots c_j^D and c_{j+1}^D in the descent path. The parameters k_A and k_D in (2) are the values of $x(t)$ corresponding to a particular knot c_i^A on the ascent path and to another knot c_j^D on the descent path, respectively. Without the loss of generality, k_A and k_D are chosen as the values of $x(t)$ corresponding to the knots c_1^A and c_1^D, respectively.

Owing to stiction, both the process output and controller output in the control loop depicted in Fig. 1 are oscillatory as those in Fig. 2. As a result, the control valve moves along the ascent and descent paths as shown in Fig. 3, and thus shows a hysteretic nonlinearity. The point-slope representation in (2) is capable of describing such a hysteretic nonlinearity. The input nonlinearity model in (2) is not a data-driven stiction model like those in Chapters 2 and 3 in [22], because it lacks the extrapolation ability, and is used only to describe the valve movements in a given measurement set of signals in the control loop, in order to detect control valve stiction. Eq. (2) is linear-in-parameter for each of the ascent and descent paths that will lead to simple identification algorithms. Even though there are quite a few linear-in-parameter representations for the static input nonlinearity of standard Hammerstein systems, most of them including the commonly-used polynomials and some other types of piecewise-linear models (e.g., that in [12]) cannot be used here since the ascent and descent paths of sticky control valve may involve zero slopes at some portions of the paths.

The identification objective is to estimate structure parameters $(n_a,\ n_b,\ n_k,\ p_A,\ p_D)$ and knots $\begin{bmatrix} c_1^A & c_2^A & \cdots & c_{p_A}^A \end{bmatrix}$ and $\begin{bmatrix} c_1^D & c_2^D & \cdots & c_{p_D}^D \end{bmatrix}$, and the unknown parameters a_i's, b_j's, d and the parameter vectors g_A and g_D, based on the available data $\{u(t), y(t)\}_{t=1}^N$.

The signals $u(t)$ and $y(t)$ are assumed to satisfy the following condition, namely,

A1 $u(t)$ and $y(t)$ are cyclo-wise-sense stationary (CWSS), and $E(u(t))$ or $E(y(t))$ has only two extrema.

A discrete-time signal $s(t)$ is said to be CWSS, if there exists a positive integer p such that

$$\begin{aligned} E(s(t + lp)) &= E(s(t)), \\ R_s(t_1 + lp, t_2 + lp) &= R_s(t_1, t_2) := E(s(t_1)s^*(t_2)), \end{aligned} \tag{4}$$

for all integers t_1, t_2 and l [27]. Symbol $E(\cdot)$ is the expectation operation and superscript $*$ denotes the conjugate transpose. If $p = 1$, then $s(t)$ is wide-sense stationary (WSS).

The cyclo-stationarity of $u(t)$ and $y(t)$ comes essentially from the self-sustained oscillations arisen in the feedback control loop caused by control valve stiction, as depicted in Fig. 1. That is, even if the setpoint $r(t)$ is kept constant and the process noise $v(t)$ is absent, some non-zero initial condition can drive $u(t)$ and $y(t)$ into oscillations owing to the control valve nonlinearity $f(\cdot)$. This can be theoretically shown via the small-gain theorem and the describing function analysis [24] (Sections 5.4 and 7.2 therein). For instance, such an analysis for a particular data-driven stiction model has been performed in [8]. In practice, the process noise $v(t)$ is usually present; as a result, either $u(t)$ or $y(t)$ is a linear combination of a deterministic periodic sequence and a stationary random process, which leads to the cyclo-stationarity of $u(t)$ and $y(t)$ [15]. Thus, $E(u(t))$ and $E(y(t))$ are periodic sequences. In addition, $E(u(t))$ or $E(y(t))$ is assumed to have only two extreme values. Taking $E(u(t))$ as an illustration, if the two extreme values of $E(u(t))$ are respectively denoted as u_{min} for the minimum and u_{max} for the maximum, then $E(u(t))$ either goes from u_{min} to u_{max} in a non-decreasing manner, or from u_{max} to u_{min} in a non-increasing manner. The assumption A1 is satisfied by many practical oscillatory signals caused by control valve stiction. Let us take the data group from chemical industry

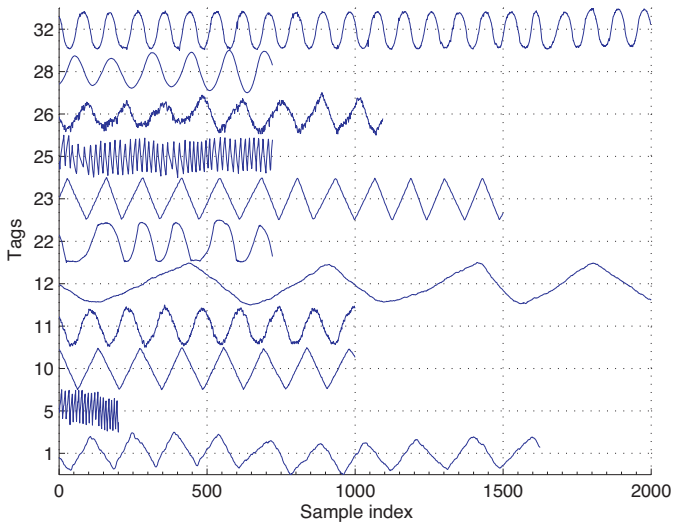

Fig. 5. The time trends of controller output $u(t)$'s from eleven feedback control loops with tag No. {1, 5, 10, 11, 12, 22, 23, 25, 26, 28, 32} in the international database.

(cdata.chemicals) in an international database [1] based on the book [22] as an example. There are seventy-six control loops in this data group, among which nineteen loops have been commented by data providers having the control valve stiction. In the nineteen loops, seventeen loops are in the feedback control mode. Among them, the controller outputs of eleven loops with tag No. {1, 5, 10, 11, 12, 22, 23, 25, 26, 28, 32}, shown in Fig. 5 here, (approximately) satisfy the assumption A1. Note that most of the examples in the articles using Hammerstein system identification methods to detect control valve stiction [31,21,23,29,14] satisfy the assumption A1 too.

3. Identification method

This section proposes a method to solve the identification problem formulated in Section 2.

Using the ARX model in (1) and the input nonlinearity model in (2), the extended Hammerstein system is described as

$$\begin{aligned} y(t) &= -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j x(t-j-n_k) + d + e(t) \\ &= -\sum_{i=1}^{n_a} a_i y(t-i) + d + e(t) + \sum_{j=1}^{n_b} b_j h(t-j-n_k)\eta_A^T \\ &\quad \times (u(t-j-n_k))g_A + \sum_{j=1}^{n_b} b_j(1-h(t-j-n_k))\eta_D^T \\ &\quad \times (u(t-j-n_k))g_D \end{aligned} \tag{5}$$

The estimation of the knots $\begin{bmatrix} c_1^A & c_2^A & \cdots & c_{p_A}^A \end{bmatrix}$ and $\begin{bmatrix} c_1^D & c_2^D & \cdots & c_{p_D}^D \end{bmatrix}$ is rather difficult, because they appear in (5) in a nonlinear manner. In order to make the identification method simple, we choose the knots to be uniformly distributed in the range $[\min(u(t)), \max(u(t))]$; a possible refinement of the knots will be discussed later at Remark #4. Thus, the identification objective is simplified to estimate the structure parameters (n_a, n_b, n_k, p_A, p_D) and the unknown parameters $\theta := \begin{bmatrix} a & b & d \end{bmatrix}^T := \begin{bmatrix} a_1 & \cdots & a_{n_a} & b_1 & \cdots & b_{n_b} & d \end{bmatrix}^T$ and (g_A, g_D), based on the available data $\{u(t), y(t)\}_{t=1}^N$. The direction signal $h(t)$ is usually not available at hand and has to be generated from the measurement of $u(t)$; however, the noise in $u(t)$ may deteriorate the generation of $h(t)$ so that filtering $u(t)$ is often necessary.

Eq. (5) is bilinear in the unknown parameter vectors θ and (g_A, g_D). The overparameterized method or the iterative method for identification of conventional Hammerstein systems [19] can be exploited to estimate these parameters. If the number of unknown parameters is large, which is the case for adopting a high-order ARX model to deal with process noise, the overparameterized method may suffer from the so-called dimension problem [3]. Hence, inspired by the study in [4] on the iterative method for Hammerstein systems, we propose an iterative method as follows:

Step 1. Choose the initial estimates as

$$\hat{a}(0) = \begin{bmatrix} \hat{a}_1 \\ \vdots \\ \hat{a}_{n_a} \end{bmatrix} = 0, \hat{b}(0) = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_{n_b} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \hat{d}(0) = 0.$$

Step 2. **For** $k = 1, 2, \ldots$ **until convergence**

By treating $\hat{a}(k-1)$, $\hat{b}(k-1)$, $\hat{d}(k-1)$ as known values, the parameter vectors g_A and g_D are estimated by solving an non-negative linear least-squares (LS) constraint problem, based on the linear regression model

$$\begin{aligned} y(t) = &-\sum_{i=1}^{n_a} \hat{a}_i y(t-i) + \hat{d} + \sum_{j=1}^{n_b} \hat{b}_j h(t-j-n_k)\eta_A^T(u(t-j-n_k))g_A \\ &+ \sum_{j=1}^{n_b} \hat{b}_j(1-h(t-j-n_k))\eta_D^T(u(t-j-n_k))g_D, \end{aligned}$$

subject to some non-negative constraints,

$$g_i^A \geq 0, \quad i = 1, 2, \cdots, p_A - 1,$$

$$g_j^A \geq 0, \quad j = 1, 2, \cdots, p_D - 1.$$

These constraints are arisen from a fact that the parameters g_i^A's and g_j^D's are the slopes of the straight lines in the ascent and descent paths, and thus they should take non-negative values. Note that the non-negative linear LS constraint problem can be solved effectively, e.g., via the Matlab function 'lsqnonneg'.

Next, let $\hat{g}_A(k)$ and $\hat{g}_D(k)$ be the estimates of g_A and g_D in the k-th iteration and treat them as known parameter vectors, estimate the parameters a, b and d based on another linear regression

$$\begin{aligned} y(t) = &-\sum_{i=1}^{n_a} a_i y(t-i) + d + \sum_{j=1}^{n_b} b_j \left[h(t-j-n_k)\eta_A^T(u(t-j-n_k)) \right. \\ &\left. \times \hat{g}_A(k) + (1-h(t-j-n_k))\eta_D^T(u(t-j-n_k))\hat{g}_D(k) \right]. \end{aligned}$$

The estimated parameter values are denoted by $\hat{a}(k)$, $\hat{b}(k)$ and $\hat{d}(k)$. Finally, normalize $\hat{b}(k)$ to have a unit Euclidean norm and set the sum of each element in $\hat{b}(k)$ to be positive if necessary to remove the gain ambiguity, i.e.,

$$\|\hat{b}(k)\| = 1, \sum_{j=1}^{n_b} \hat{b}_j > 0. \tag{6}$$

end

Remark #1: The convergence criterion for Step 2 is designed by monitoring the relative percentage improvement of the estimated parameters:

$$\frac{\|\hat{\vartheta}(k) - \hat{\vartheta}(k-1)\|}{\|\hat{\vartheta}(k)\|} < \mu, \tag{7}$$

[1] Available online at http://www.ualberta.ca/ bhuang/book2.htm.

where μ is a small positive real number close to zero, e.g., $\mu = 0.01$, and $\hat{\vartheta}(k)$ stands for the estimated parameter vector composed by $\hat{a}(k)$, $\hat{b}(k)$, $\hat{d}(k)$, $\hat{g}_A(k)$ and $\hat{g}_D(k)$. The proposed iterative method follows the same principle as that in [4]. Since the input $u(t)$ is not white noise as required in [4], the convergence therein cannot be generalized here at this moment; however, the divergent has not appeared to our experience.

Remark #2: The way to remove the gain ambiguity is certainly not unique. Taking the sum of b_j's in (6) is to avoid picking a zero single component by accident. However, the sum can be zero, too. To avoid this, an option is to drop this constraint of the sum, and keep only the normalized norm constraint. The resulted solution is unique up to the sign. By doing so, it may happen that in each iteration the sign jumps. This is not really a problem if the criterion defined in (7) is modified to

$$\frac{\min(\|\hat{\vartheta}(k) - \hat{\vartheta}(k-1)\|, \|\hat{\vartheta}(k) + \hat{\vartheta}(k-1)\|)}{\|\hat{\vartheta}(k)\|} < \mu.$$

Remark #3: The model structure parameters $(n_a, n_b, n_k, p_A, p_D)$ are determined via a grid search by minimizing the Akaike information criterion (AIC) (see, e.g., Section 11.5 of [30]),

$$W(n_a, n_b, n_k, p_A, p_D) = N \log V(n_a, n_b, n_k, p_A, p_D) + 2dim(\hat{\vartheta}(k)). \quad (8)$$

Here $dim(\hat{\vartheta}(k))$ is the number of unknown parameters in $\hat{\vartheta}(k)$, and the loss function $V(n_a, n_b, n_k, p_A, p_D)$ is based the model simulation error,

$$V(n_a, n_b, n_k, p_A, p_D) = \sum_{t=1}^{N} (y(t) - \hat{y}_s(t))^2$$

where

$$\hat{y}_s(t) = \frac{\sum_{j=1}^{n_b} \hat{b}_j q^{-j}}{1 + \sum_{i=1}^{n_a} \hat{a}_i q^{-i}} \hat{x}(t - n_k) + \frac{1}{1 + \sum_{i=1}^{n_a} \hat{a}_i q^{-i}} \hat{d}$$

and

$$\hat{x}(t) = \begin{cases} \eta_A^T(u(t))\hat{g}_A, & \text{for } h(t) = 1 \\ \eta_D^T(u(t))\hat{g}_D, & \text{for } h(t) = 0 \end{cases}.$$

That is, for each set of $\{n_a, n_b, n_k, p_A, p_D\}$, Steps 1 and 2 are implemented; thus, a number of models with different sets of $\{n_a, n_b, n_k, p_A, p_D\}$ are identified. The model having the minimum AIC $W(n_a, n_b, n_k, p_A, p_D)$ is selected as the final choice. The usage of AIC is owing to the parsimony principle, namely, a model having less number of parameters is preferred, if the improvement of $V(n_a, n_b, n_k, p_A, p_D)$ achieved by increasing the number of parameters is minor.

Remark #4: The slope parameters g_i^A's and g_j^D's are always bounded; as a result, some modeling error may be arisen if the control valve stiction contains some nonlinearity that is close to a discontinuity. Note that the stem movement of a physical control valve cannot be a pure discontinuity. Such a modeling error can be reduced by a refinement of the knots that are originally chosen to be uniformly distributed. If the estimated slope parameters are $\hat{g}_A = [\cdots, 0, \hat{g}_i^A, 0, \cdots]$ for some non-zero value $\hat{g}_i^A$, then an approximated discontinuity may exist between the two knots c_i^A and c_{i+1}^A, and more knots evenly spaced can be introduced in the interval $[c_i^A, c_{i+1}^A]$, e.g., four knots c_i^A, $c_i^A + (c_{i+1}^A - c_i^A)/3$, $c_i^A + (2(c_{i+1}^A - c_i^A))/3$ and c_{i+1}^A are obtained if two more knots are added. The knots in the descent path can be refined similarly. With the refined knots, the extended Hammerstein system is re-identified, and the modeling error could be reduced, leading to a smaller AIC $W(n_a, n_b, n_k, p_A, p_D)$ in (8). Similar to the model structure parameters n_a, n_b, n_k, p_A and p_D, the knots $c_1^A, \cdots, c_{p_A}^A$ and $c_1^D, \cdots, c_{p_D}^D$ are fixed before Steps 1 and 2 so that the refinement of knots is performed outside the iterations in Step 2.

4. Identifiability analysis

This section investigates the identifiability problem for the extended Hammerstein systems. A relevant well-known fact is that if the setpoint is kept constant, then the data set in the feedback control loop without the input nonlinearity solely driven by the process noise $v(t)$ (see Fig. 1) is not informative [25]. In this context, the setpoint is indeed a constant; thus, it is doubtful if the oscillatory signals $u(t)$ and $y(t)$ caused by control valve stiction are informative enough to make the the extended Hammerstein system identifiable. Even though a few researchers have successfully identified the symmetric stiction model for control valves and the linear dynamic model [31,21,23,29,14], the study on the identifiability problem is very limited. Section 11.5 of [22] provided a preliminary identifiability analysis based on an assumption that the inner signal $x(t)$ is independent to the noise $v(t)$, which is inconsistent with a fact that $x(t)$ and $v(t)$ are connected via the feedback loop. Therefore, the identifiability problem still is a fundamental open problem, and is to be resolved here.

As stated in [25] (Page 112 therein), the identifiability problem involves "aspects on whether the data set (the experimental conditions) is informative enough to distinguish between different models as well properties of the model structure itself".

First, let us investigate the model structure of the extended Hammerstein system in (5). Depending on the input $u(t)$ is increasing or decreasing, the direction signal $h(t)$ in (3) takes the value 1 or 0 so that the extended Hammerstein system indeed switches between two standard Hammerstein systems:

$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j \eta_A^T(u(t-j-n_k))g_A + d + e(t),$$
$$y(t) = -\sum_{i=1}^{n_a} a_i y(t-i) + \sum_{j=1}^{n_b} b_j \eta_D^T(u(t-j-n_k))g_D + d + e(t).$$

By incorporating the constraints in (6), the gain ambiguity between b_j and g_A or g_D is removed. Thus, each standard Hammerstein system can be regarded as an ARX model with multiple inputs, whose model structure is well-known to be injective [25]. Hence, the model structure of the extended Hammerstein system is injective, i.e., different values of parameters a_i, b_j, d, g_A and g_D will not yield an equal extended Hammerstein process system.

Next, the identifiability analysis is on whether the data set is informative. The analysis is performed for a general extended Hammerstein system,

$$y(t) = G(q, \theta)f(u(t), \theta_f) + H(q, \theta)e(t) \quad (9)$$

That is, the choice of the ARX model in (1) for the LTI dynamic block is not necessary for the identifiability analysis. In (9), $f(u(t), \theta_f)$ is the input nonlinearity described by the hysteretic model in (2), $G(q, \theta)$ stands for the discrete-time LTI process model, and $H(q, \theta)$ is the discrete-time LTI noise model driven by a zero-mean stationary white noise $e(t)$. The objective of is to investigate whether the data set for the closed-loop identification of extended Hammerstein system is informative enough to give unique values of θ_f and θ equal to their true values $\theta_{f,0}$ and θ_0 respectively.

In terms of the input nonlinearity, the oscillatory signal $u(t)$ covers the ascent and descent paths; thus, the samples of $u(t)$ and the inner signal $x(t)$ contain enough information to distinguish different input nonlinearity models. If the true input nonlinearity model is inside the model set described by $f(u(t), \theta_f)$, then the informative data set ensures the parameter vector θ_f to reach its true value

$\theta_{f,0}$. Therefore, the identifiability problem becomes: given the true nonlinearity model, is the oscillatory data set informative enough to give a unique value of θ equal to θ_0, or equivalently, $G(q,\theta)$ equal to $G(q,\theta_0)$?

Given the true nonlinearity model and the inner signal $x(t)$, the estimation of the parameter θ is the same as the one obtained by applying the prediction error method to (9), i.e.,

$$\min_{\theta} V(\theta) = \min_{\theta} \frac{1}{N}\sum_{t=1}^{N}\frac{1}{2}\varepsilon^2(t) \tag{10}$$

$$\text{s.t. } \varepsilon(t) = H^{-1}(q,\theta)[y(t) - G(q,\theta)x(t)].$$

To simplify notations, we rewrite the prediction error $\varepsilon(t)$ as

$$\begin{aligned}\varepsilon(t) &= H_\theta^{-1}[y(t) - G_\theta x(t)] = H_\theta^{-1}[(G_0 - G_\theta)x(t) + H_0 e(t)]\\ &= H_\theta^{-1}[(G_0 - G_\theta)x(t) + (H_0 - H_\theta)e(t)] + e(t),\end{aligned} \tag{11}$$

where G_0 and H_0 stand for the true process model $G(q,\theta_0)$ and disturbance model $H(q,\theta_0)$, respectively. Under a standard assumption to avoid algebraic loops [30]: either the controller (owing to the zero-order holder in Fig. 1) or both G_0 and G_θ contain at least one sample time delay, the last term $e(t)$ in the right-hand side of (11) is uncorrelated to the other two terms that may be only correlated to $e(t-i)$ for $i \geq 1$. Then, the identifiability analysis requires the statistical relationship between $(G_0 - G_\theta)x(t)$ and $(H_0 - H_\theta)e(t)$. However, $x(t)$ and $e(t)$ in the feedback control loop depicted in Fig. 1 are connected in a complicated manner due to the input nonlinearity. Thus, the description of this statistical relationship hinders the appearance of identifiability analysis in the literature.

The statistical relationship can be analyzed by exploiting cyclo-stationarity of the signals in the feedback control loop. In Assumption A1 (Section 2), $y(t)$ is assumed to be CWSS. Because the noise source $e(t)$ is WSS and H_θ is LTI, the cyclo-stationarity of $y(t)$ comes solely from $x(t)$; as G_θ is LTI, this is possible if and only if $x(t)$ is CWSS [15]. It is a known fact that a CWSS signal with cyclo-period p can be regarded as the output of a linear-periodic time varying (LPTV) system with period p driven by a zero-mean stationary white noise [1]. Using this fact, $x(t)$ can be represented as

$$x(t) = \sum_{n=-\infty}^{\infty} k(t,n)a(n). \tag{12}$$

Here $k(t,n)$ is the Green's function of an LPTV system with $a(t)$ as input and $x(t)$ as output, and stands for the system response at time t to an impulse at time n (t and n are integers), and $a(t)$ is a zero-mean stationary white noise.

The statistical relationship between $x(t)$ and $a(t)$ can be described in the frequency domain via bispectra of cyclo-stationarity signals. The bispectrum of a CWSS signal $s(t)$ is defined as the two-dimensional Fourier transform of the correlation function $R_s(t_1,t_2)$ in (4) [27],

$$\Phi_s(\omega',\omega) = \frac{1}{2\pi}\sum_{t_1=-\infty}^{\infty}\sum_{t_2=-\infty}^{\infty} R_s(t_1,t_2)\,e^{-j\omega' t_1}e^{j\omega t_2}. \tag{13}$$

In the frequency domain, the LPTV system in (12) is fully specified by the bifrequency map $K(\omega',\omega)$ [1,26],

$$\begin{aligned}K(\omega',\omega) &= \frac{1}{2\pi}\sum_{t=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} k(t,n)e^{-j\omega' t}e^{j\omega n}\\ &= \sum_{l=-\infty}^{\infty} F_l(\omega)\delta\left(\omega - \omega' + \frac{2\pi l}{p}\right),\end{aligned} \tag{14}$$

where $\delta(\cdot)$ is the Dirac-delta function, and $F_l(\omega)$ is a finite function on the line $\omega - \omega' + \frac{2\pi l}{p} = 0$. Using a fact describing the transformed bispectra through LPTV systems (Section A in [1]; Fact 4 in [36]), $F_l(\omega)$ in (14) can be determined from the bispectrum of $x(t)$ as

$$\Phi_x(\omega',\omega) = \sum_{l=0}^{p-1} P_l(\omega)\delta\left(\omega - \omega' + \frac{2\pi l}{p}\right), \tag{15}$$

where

$$\begin{aligned}P_l(\omega) &= \sum_{k=0}^{p-1} F_{k+l}\left(\omega + \frac{2\pi l}{p}\right) F_k^*(\omega)\,\Phi_a\left(\omega - \frac{2\pi k}{p}\right)\\ &= \sigma_a^2\sum_{k=0}^{p-1} F_{k+l}\left(\omega + \frac{2\pi l}{p}\right) F_n^*(\omega).\end{aligned}$$

Here σ_a^2 is the variance of $a(t)$. As $a(t)$ is WSS white noise, the cross bispectra of $x(t)$ and $a(t)$ are nonzero only at the diagonal line $\omega = \omega'$, i.e.,

$$\Phi_{xa}(\omega',\omega) = \sigma_a^2 F_0(\omega), \tag{16}$$

$$\Phi_{ax}(\omega',\omega) = \sigma_a^2 F_0^*(\omega). \tag{17}$$

Denote the contribution of $e(t)$ to $a(t)$ as $\lambda\sigma_a^2 = \sigma_e^2$ for a positive real number λ less than 1, where σ_e^2 is the variance of $e(t)$. From (16) and (17), we have

$$\Phi_{xe}(\omega',\omega) = \lambda\sigma_a^2 F_0(\omega), \tag{18}$$

$$\Phi_{ex}(\omega',\omega) = \lambda\sigma_a^2 F_0^*(\omega). \tag{19}$$

With these results, the identifiability can be shown to hold by adopting the approach in [36]. To save the space, we would like to present the main results only and omit some related derivation and proofs that are available in Sections 6 and 7 of [36].

First, $\varepsilon(t)$ in (11) inherits the cyclo-stationarity from $x(t)$, because G_0, H_0, G_θ and H_θ are LTI, and $e(t)$ is WSS. The loss function $V(\theta)$ in (10) as the data length N goes to infinity can be represented in the frequency domain as

$$\overline{V}(\theta) = \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\frac{1}{4\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\Phi_\varepsilon(\omega',\omega)\mathrm{d}\omega\mathrm{d}\omega', \text{ with } \omega = \omega'. \tag{20}$$

Here $\Phi_\varepsilon(\omega',\omega)$ is the bispectrum of $\varepsilon(t)$. Eq. (20) says that $\overline{V}(\theta)$ depends only on the diagonal part of $\Phi_\varepsilon(\omega',\omega)$ for $\omega = \omega'$. Second, based on (11), the bispectrum $\Phi_\varepsilon(\omega',\omega)$ can be written as

$$\begin{aligned}\Phi_\varepsilon(e^{j\omega'},e^{j\omega}) &= \frac{G_0(\omega') - G_\theta(\omega') + B_\theta(\omega')}{H_\theta(\omega')}\,\Phi_x(\omega',\omega)\\ &\times\frac{G_0(\omega) - G_\theta(\omega) + B_\theta(\omega)}{H_\theta(\omega)} + \frac{H_0(\omega') - H_\theta(\omega')}{H_\theta(\omega')}\\ &\cdot\left(\Phi_e(\omega',\omega) - \frac{\Phi_{xe}(\omega',\omega)\Phi_{ex}(\omega',\omega)}{\Phi_x(\omega',\omega)}\right)\frac{H_0(\omega) - H_\theta(\omega)}{H_\theta(\omega)}\\ &+ \Phi_e(\omega',\omega),\end{aligned} \tag{21}$$

where the bias term is

$$B_\theta(\omega) = \frac{(H_0(\omega) - H_\theta(\omega))\,\Phi_{ex}(\omega',\omega)}{\Phi_x(\omega',\omega)}. \tag{22}$$

Substituting (21) into (20) gives a frequency-domain description of asymptotic estimated parameters. This description can be regarded as a generalized version of the expression (8.77) in [25] for identification of LTI systems with CWSS inputs. Finally, since only the diagonal part of $\Phi_\varepsilon(\omega',\omega)$ plays a role in (20), we can focus our attention on the diagonal parts of $\Phi_x(\omega',\omega)$, $\Phi_{xe}(\omega',\omega)$ and $\Phi_{ex}(\omega',$

ω). With (15), (18) and (19), the first and second weighting terms for $\omega = \omega'$ are

$$\Phi_x(\omega', \omega) = \sigma_a^2 \sum_{l=0}^{p-1} |F_l(\omega)|^2,$$

and

$$\Phi_e(\omega', \omega) - \frac{\Phi_{xe}(\omega', \omega)\Phi_{ex}(\omega', \omega)}{\Phi_x(\omega', \omega)} = \sigma_e^2 - \frac{\lambda^2 \sigma_a^4 |F_0(\omega)|^2}{\sigma_a^2 \sum_{l=0}^{p-1} |F_l(\omega)|^2}$$

$$= \frac{\lambda \sigma_a^2}{\sum_{l=0}^{p-1} |F_l(\omega)|} \left((1-\lambda)|F_0(\omega)|^2 + \sum_{l=1}^{p-1} |F_l(\omega)|^2 \right).$$

If $\lambda = 0$, i.e., $e(t)$ does not contribute to $a(t)$, then the bias term $B_\theta(\omega)$ in (22) is absent, and the first weighting term $\Phi_x(\omega', \omega)$ is nonzero so that G_0 is a unique solution of G_θ. Otherwise, if $0 < \lambda < 1$, the two weighting terms are nonzero. Thus, the minimal value of the loss function in (20) is achieved when the bispectrum $\Phi_\varepsilon(\omega', \omega)$ in (21) is equal to $\Phi_e(\omega', \omega)$ with $\omega' = \omega$. Under the assumption that G_0 and H_0 are covered by the model set of G_θ and H_θ, as a result, H_0 is a unique solution of H_θ, leading to the vanish of the bias term $B_\theta(\omega)$ in (22) so that G_0 is a unique solution of G_θ.

5. Examples

We presents three examples to illustrate the effectiveness of the proposed identification method. Because the ARX model in (1) and the point-slope-based input nonlinearity model in (2) are used, the two models might involve quite a few unknown parameters so that it would be cumbersome to present these parameters, and be more convenient to present the time or frequency response of the ARX model and the input-vs-output plot of the nonlinearity model. As a complement, all the data sets and the identified models for the three examples are provided via Matlab data and program files available online for the purpose of academic research. [2]

Example 5.1. Consider the feedback control loop depicted in Fig. 1. The process model is

$$G(s) = \frac{1.25e^{-2s}}{80s + 1}.$$

The controller is a discretized counterpart of a continuous-time PI controller

$$C(s) = 1.303 \left(1 + \frac{1}{10s}\right),$$

with the sampling period 0.5 s. The process noise v is generated as the output of a digital filter $1/(1 + 0.5q^{-1})$ driven by a Gaussian white noise with zero mean and variance equal to 0.001. The sticky valve model is the data-driven model proposed in [20] with two parameters $f_s = 5.0$ and $f_d = 0$, whose character is given in the left subplot of Fig. 8. The presence of valve stiction leads to oscillations in the loop, even though the setpoint is constant.

The identification method is applied to the collected data $\{u(t), y(t)\}_{t=1}^{N=2800}$ shown in Fig. 6. The step responses of actual and identified process models are given in Fig. 7. The actual and identified input nonlinearities are compared in Fig. 8. The model quality is quantitatively measured by the fitness between $\hat{y}_s(t)$ and $y(t)$,

$$\text{Fitness} = 100 \left(1 - \frac{\|\hat{y}_s(t) - y(t)\|_2}{\|y(t) - E\{y(t)\}\|_2}\right). \quad (23)$$

[2] www.mech.pku.edu.cn/robot/teacher/wangjiandong/research.htm.

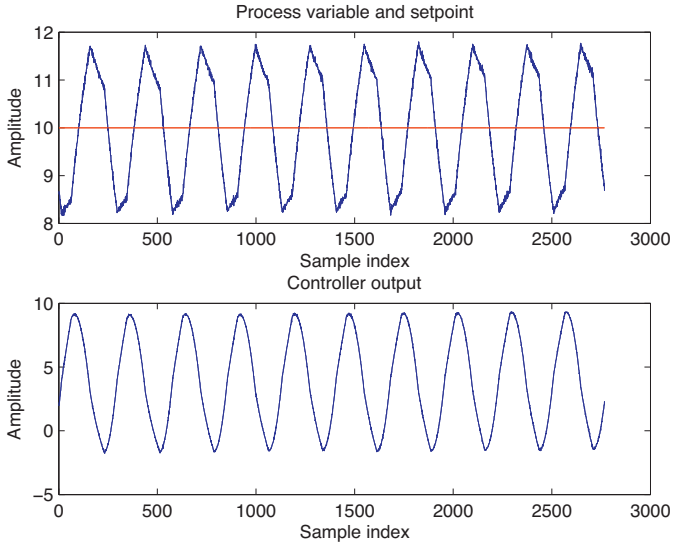

Fig. 6. The time trends of $u(t)$ (bottom) and $y(t)$ (top) in Example 1.

The fitness in this example is 92.7828%, with $y(t)$ and $\hat{y}_s(t)$ shown in Fig. 9. The identified input nonlinearity can reveal the basic shape of the actual valve stiction. The identified process model has the dynamics close to the actual one. Note that due to the gain ambiguity, the step response of the identified process model has a different steady-state gain with the actual one.

To support the identifiability analysis in Section 4, the signals $y(t)$, $u(t)$ and $x(t)$ are found to be CWSS with the cyclo-period $p = 280$ samples, estimated via the variability method in [37]. The cyclic spectrum $\Phi_x(\omega', \omega)$ for $\omega = \omega'$ shown in Fig. 10 is obtained via the bispectrum estimator in [16]; it covers a broad frequency range up to high frequencies. Thus, as expected from (21), the estimated process model matches well the actual one except at high frequencies, as shown in the Bode plots in Fig. 11. Note that the steady-state gains of the two models are adjusted for comparison in Fig. 11.

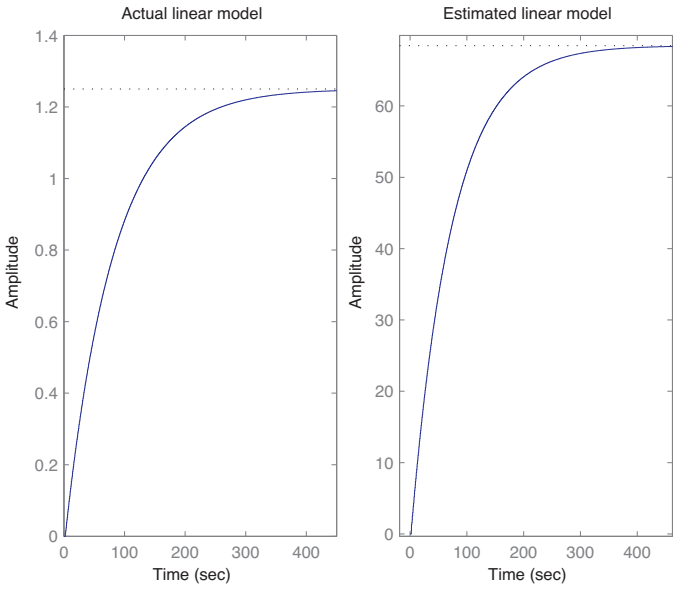

Fig. 7. The step responses of actual (left) and identified (right) process models in Example 1.

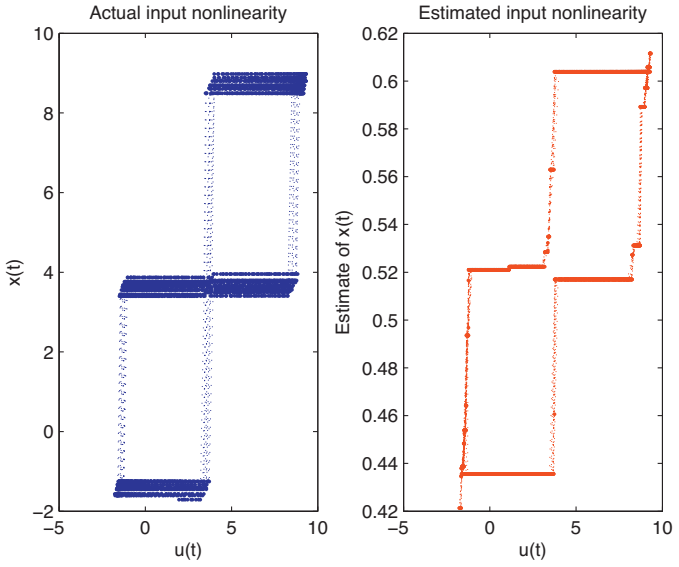

Fig. 8. The actual (left) and identified (right) input nonlinearities in Example 1.

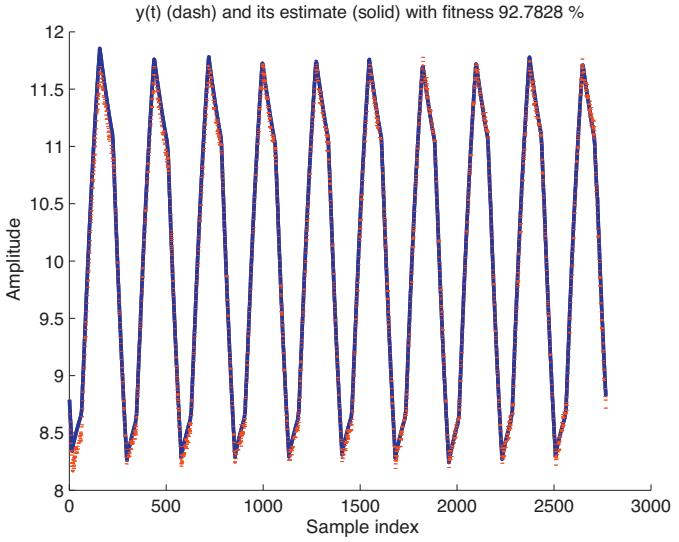

Fig. 9. The actual output $y(t)$ (dash) and simulated one $\hat{y}_s(t)$ (solid) in Example 1.

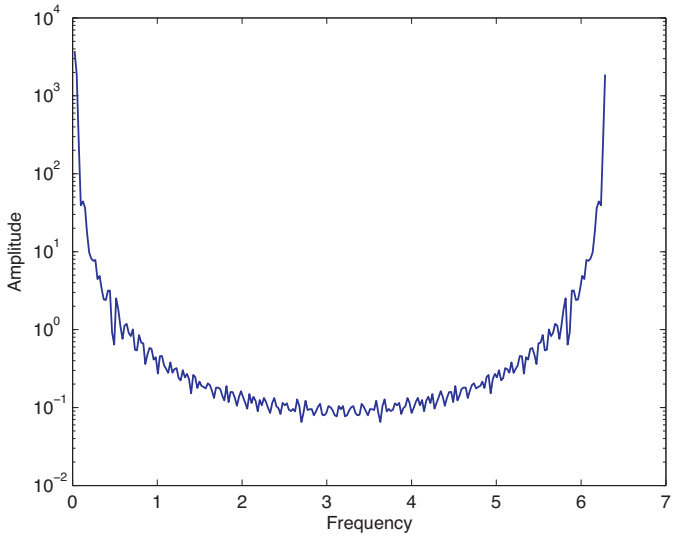

Fig. 10. The cyclic spectrum $\Phi_x(\omega', \omega)$ for $\omega = \omega'$.

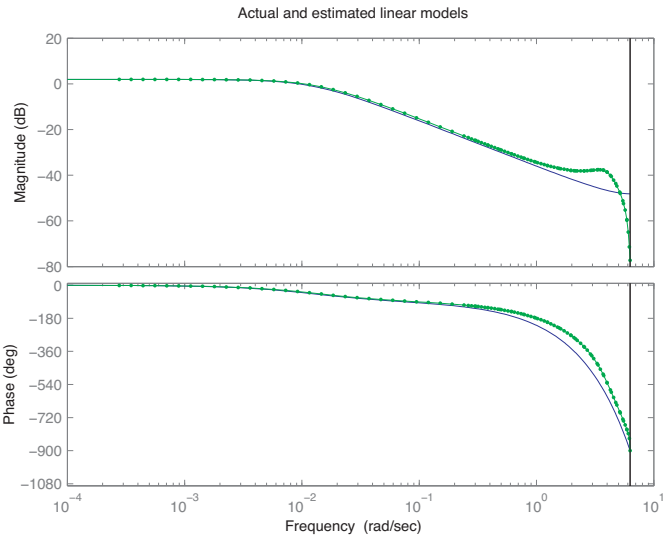

Fig. 11. The Bode plots of the actual (solid) and identified (dot-dash) process models in Example 1.

Example 5.2. The proposed identification method is applied to the experimental data in Fig. 2. The simulated output $\hat{y}_s(t)$ fits well with the actual output $y(t)$ with the fitness in (23) equal to 81.1891% as shown in Fig. 12. The estimated input nonlinearity is presented in Fig. 13 (right subfigure), and is close to the actual characteristic of the control valve in Fig. 3 (reproduced as the left subfigure of Fig. 13).

Example 5.3. Real-time measurements of a flow control loop are collected with the sampling period of 1 second from August 2011 to December 2011 in a large-scale petro-chemical plant in China. The variation range of the process output usually is about 1 unit; however, severe oscillations occur with the variation range about 15 units, as shown in Fig. 14. The setpoint takes a constant value 110 and the actual valve position is not available. There is a discretized counterpart of a continuous-time PI controller $C(s) = 0.24\left(1 + (1/85.3s)\right)$ in the control loop.

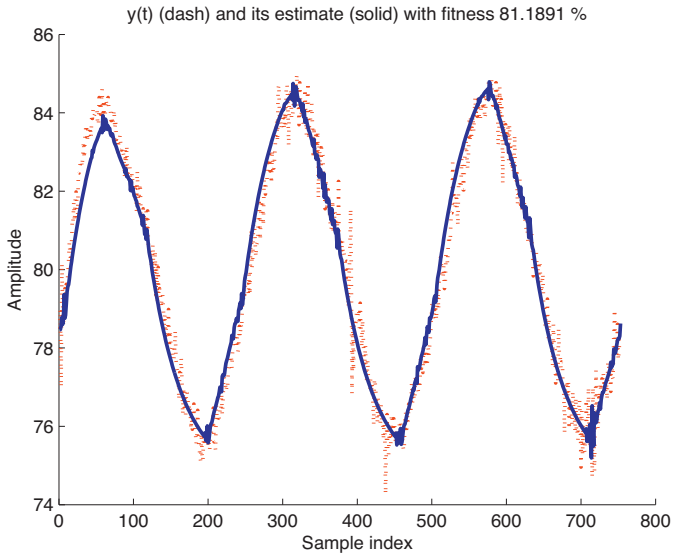

Fig. 12. The actual output $y(t)$ (dash) and simulated one $\hat{y}_s(t)$ (solid) in Example 2.

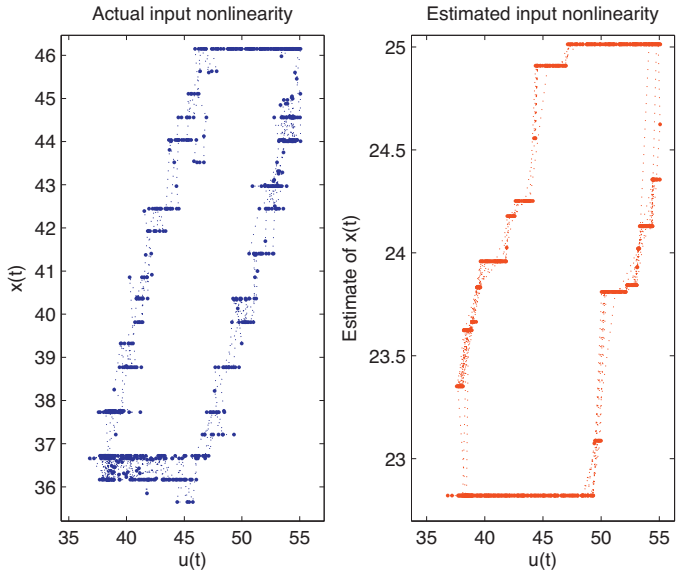

Fig. 13. The actual (left) and identified (right) input nonlinearities in Example 2.

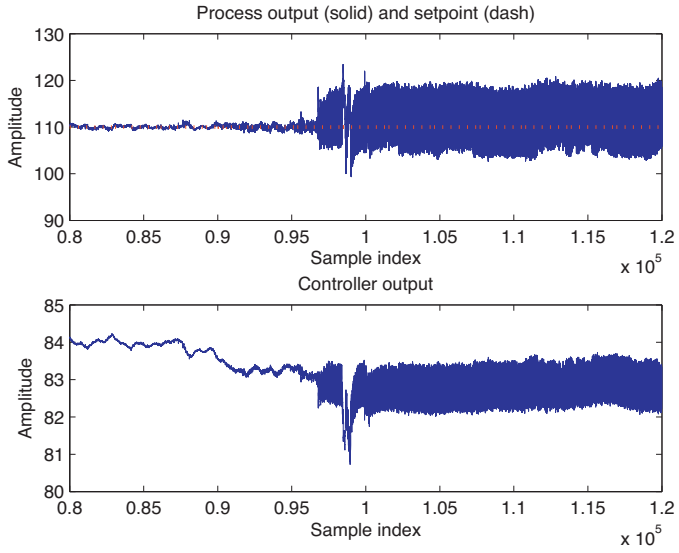

Fig. 14. The time trends of $u(t)$, $y(t)$ and $r(t)$ in Example 3.

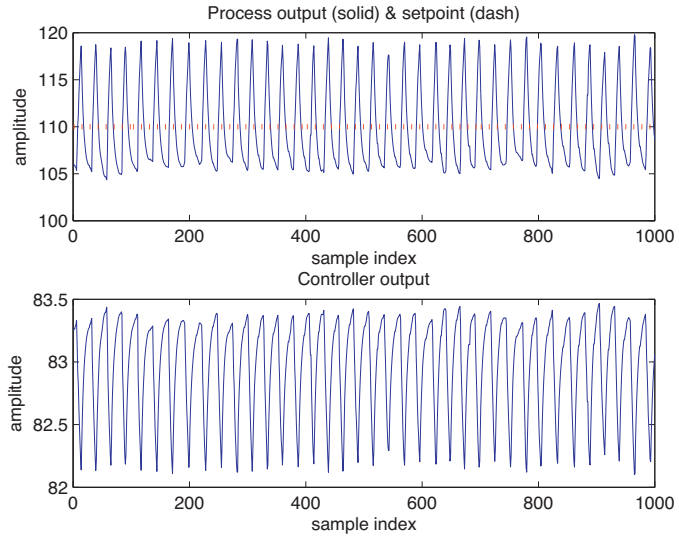

Fig. 15. The identification data $u(t)$, $y(t)$ and $r(t)$ in Example 3.

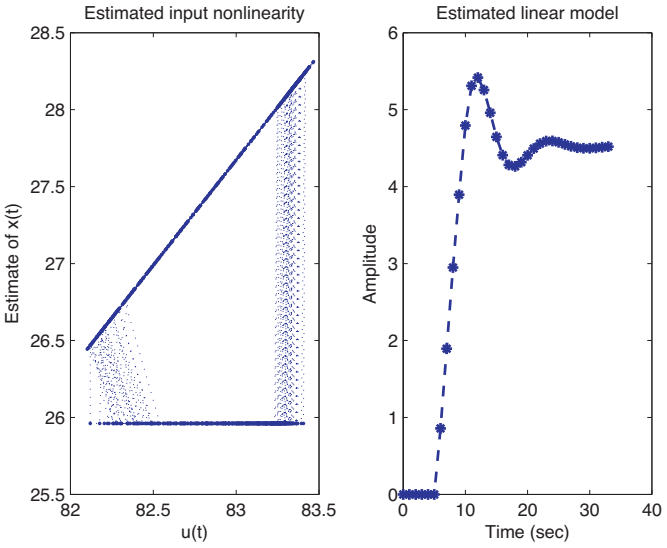

Fig. 16. The estimated input nonlinearity and process model in Example 3 from the proposed method.

The proposed method is applied to the data $\{u(t), y(t)\}_{t=1}^{N=1000}$ given in Fig. 15, which are selected from some parts of the oscillatory data in Fig. 14 as the identification data. The identified process model is

$$\hat{G}(q) = \frac{0.858q^{-1} + 0.094q^{-2} + 0.312q^{-3} + 0.175q^{-4} + 0.357q^{-5}}{1 - 1.097q^{-1} + 0.454q^{-2} - 0.096q^{-3} + 0.135q^{-4} + 0.001q^{-5}} q^{-5},$$

whose step response is presented in Fig. 16. The identified input nonlinearity is also presented in Fig. 16, where the knots for the ascend and descend pathes are the same pair, $\hat{c}_1^A = \hat{c}_1^D = 82.1020$, $\hat{c}_2^A = \hat{c}_2^D = 82.4670$, but the slope parameters are very different, $\hat{g}_1^A = 0$ and $\hat{g}_1^D = 1.36981$. The estimated values of $x(t)$ corresponding to $\hat{c}_1^A$ and $\hat{c}_1^D$ are $\hat{k}_A = 25.9610$ and $\hat{k}_D = 26.4441$, respectively. The fitness in (23) between $\hat{y}_s(t)$ and $y(t)$ is 68.5815% as shown in Fig. 17.

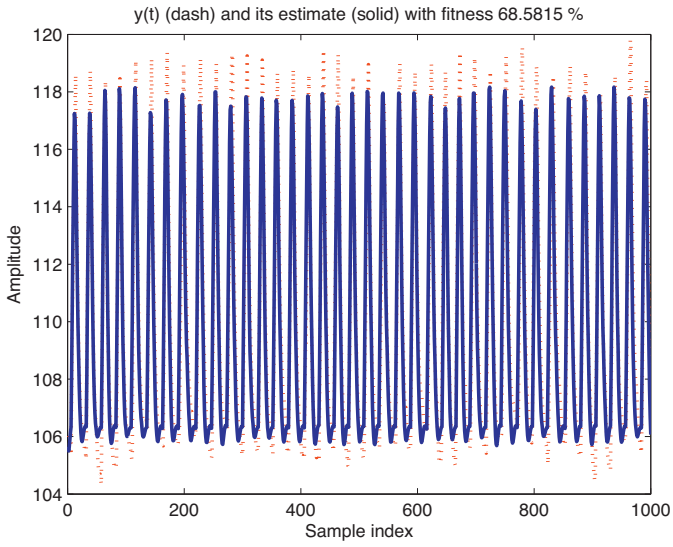

Fig. 17. The actual output $y(t)$ (dash) and simulated one $\hat{y}_s(t)$ (solid) in Example 3 from the proposed method.

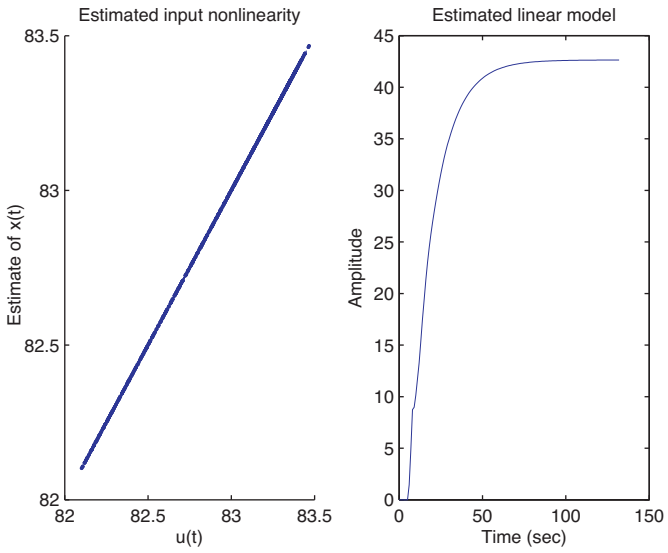

Fig. 18. The estimated input nonlinearity and process model in Example 3 from the two-stage method.

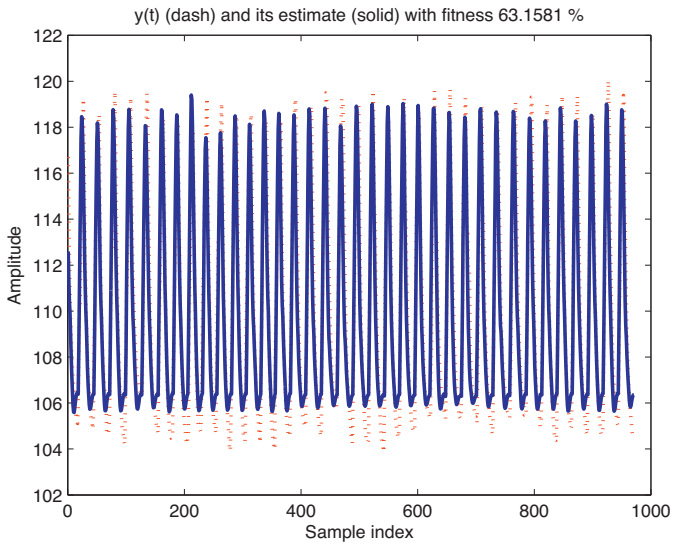

Fig. 20. The actual output $y(t)$ (dash) and simulated one $\hat{y}_s(t)$ (solid) for the validation data in Example 3.

As a comparison, the two-stage identification method in [21] is applied to the identification data in Fig. 15. The data-driven stiction model proposed by He et al. [20] is used to represent the control valve. As the optimization problem to be solved in the two-stage identification method is indeed a separable least-squares problem with two unknown parameters (Eq. (14) in [21]), a grid search method over 2-dimensional search space is exploited to obtain the optimal estimates, instead of the pattern search or genetic algorithms in [21]. A linear control valve is obtained as shown in Fig. 18, i.e., $\hat{f}_s = \hat{f}_d = 0$, and the fitness between $\hat{y}_s(t)$ and $y(t)$ is equal to 53.7318% as shown in Fig. 19. Note that the two-stage method is able to yield satisfactory estimates for Example 1 (the detailed results are omitted here).

The proposed method and the two-stage method obtain different results here. In particular, no stiction is detected by using the two-stage identification method taking He's data-driven stiction model; by contrast, the proposed method detects an asymmetric control valve stiction, namely, the ascent path is sticky while the descent path is not. As a cross validation, the estimated model from the proposed method is applied to another data set shown in Fig. 20. The fitness between $\hat{y}_s(t)$ and $y(t)$ for the validation data is equal to 63.1581%, saying that the estimated model from the proposed method is quite reliable.

The detected control valve stiction is confirmed by a further investigation. This flow control loop has other two parallel flow

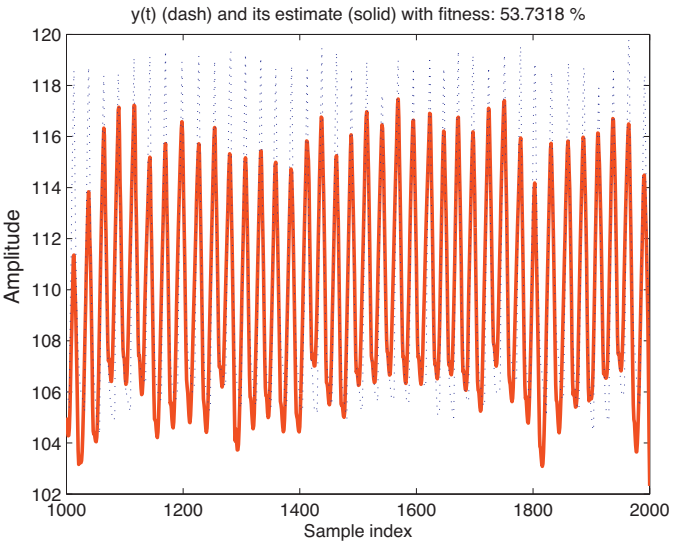

Fig. 19. The actual output $y(t)$ (dash) and simulated one $\hat{y}_s(t)$ (solid) in Example 3 from the two-stage method.

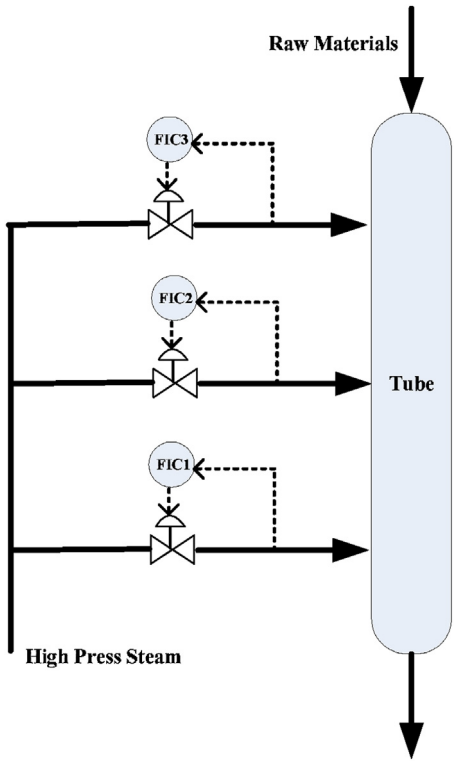

Fig. 21. Diagram of the three flow control loops.

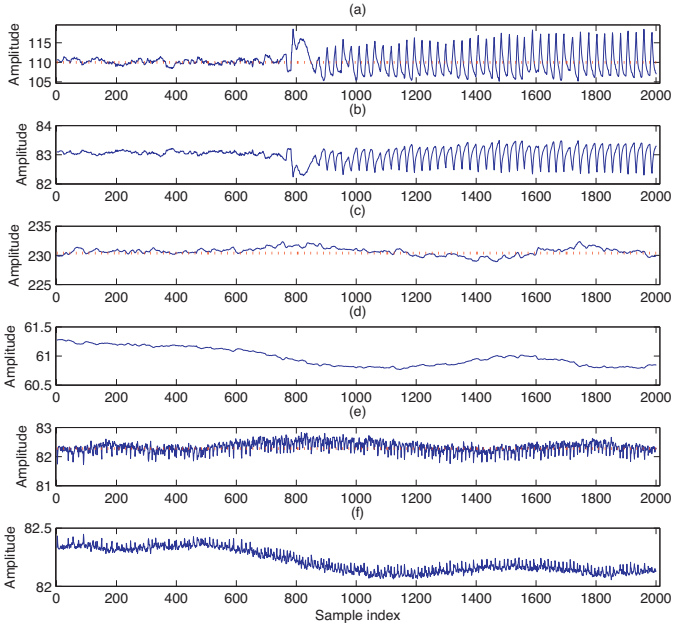

Fig. 22. Signals in three flow control loops in Example 3: Loop FIC1: (a) the output (solid) and setpoint (dash) and (b) the controller output, Loop FIC2: (c) the output (solid) and setpoint (dash) and (d) the controller output, Loop FIC3: (e) the output (solid) and setpoint (dash) and (f) the controller output.

control loops, all of them injecting high pressure steams into raw materials to maintain a certain level of fluidity, as depicted by the process diagram in Fig. 21. The upstream of the three loops is in a stable condition. The current flow control loop denoted by the controller FIC1 in Fig. 21 runs into oscillation, while the other two flow control loops denoted by FIC2 and FIC3 do not have oscillatory behaviors, as shown in Fig. 22. The failure of the two-stage method is believed to be caused by the structure limitation of He's data-driven stiction model that can not describe such an asymmetric control valve stiction.

6. Conclusion

Motivated by the problem of detecting control valve stiction, we formulated an identification problem for a class of extended Hammerstein systems with hysteretic input nonlinearity based on oscillatory input and output data, and proposed an iterative identification method. The point-slope-based input hysteretic model is composed by two ascent and descent paths, and is flexible enough to describe asymmetric control valve stiction. The structure of the formulated extended Hammerstein system was shown to be identifiable. The oscillatory signals in feedback control loops were proven to be informative enough to achieve identifiability. Numerical, experimental and industrial examples have demonstrated the effectiveness of the proposed identification method.

Several interesting issues are left open. First, the convergence of the proposed iterative method has not been proved, although the divergent has not appeared in our experience. Second, in some cases, the oscillatory controller output $u(t)$ caused by control valve stiction does not satisfy Assumption A1; as a result, the input nonlinearity contains other paths along with the two ascent and descent paths, and the model in (2) is no longer valid. Therefore, one important research work is to characterize the control valve stiction for the general types of input and output data that do not satisfy Assumption A1.

Acknowledgment

The authors would like to thank Sinopec Yangzi Petro-chemical Co., Nanjing, China, for providing industrial data used in this study.

References

[1] S. Akkarakaran, P.P. Vaidyanathan, Bifrequency and bispectrum maps: a new look at multirate systems with stochastic inputs, IEEE Trans. Signal Process. 48 (2000) 723–736.
[2] E.W. Bai, Identification of linear systems with hard input nonlinearities of known structure, Automatica 38 (2002) 853–860.
[3] E.W. Bai, Discussion on: subspace-based identification algorithms for Hammerstein and Wiener models, Eur. J. Control 11 (2005) 137–138.
[4] E.W. Bai, K. Li, Convergence of the iterative algorithm for a general Hammerstein system identification, Automatica 46 (2010) 1891–1896.
[5] W.L. Bialkowski, Dreams versus reality: a view from both sides of the gap, Pulp Paper Can. 94 (1993) 19–27.
[6] V. Cerone, D. Regruto, Bounding the parameters of linear systems with input backlash, IEEE Trans. Autom. Control 52 (2007) 531–536.
[7] Y. Cheng, C. Yu, Relay feedback identification for actuators with hysteresis, Ind. Eng. Chem. Res. 39 (2000) 4239–4249.
[8] M.A.A.S. Choudhury, N.F. Thornhill, S.L. Shah, Modeling valve stiction, Control Eng. Pract. 13 (2005) 641–658.
[9] M.A.A.S. Choudhury, S.L. Shah, N.F. Thornhill, Diagnosis of Process Nonlinearities and Valve Stiction, Springer, London, 2006.
[10] L. Deng, Y. Tan, Modeling hysteresis in piezoelectric actuators using NARMAX models, Sens. Actuat. A: Phys. 149 (2009) 106–112.
[11] L. Desborough, P. Nordh, R. Miller, Control system reliability: process out of control, Ind. Comput. 8 (2001) 52–55.
[12] G. Dolanc, S. Strmcnik, Identification of nonlinear systems using a piecewise-linear Hammerstein model, Syst. Control Lett. 54 (2005) 145–158.
[13] R. Dong, Y. Tan, H. Chen, Recursive identification for dynamic systems with backlash, Asian J. Control 12 (2010) 26–38.
[14] M. Farenzena, J.O. Trierweiler, Valve stiction estimation using global optimisation, Control Eng. Pract. 20 (2012) 379–385.
[15] W.A. Gardner (Ed.), Cyclostationarity in Communications and Signal Processing, IEEE Press, Piscataway, N.J, 1994.
[16] G.B. Giannakis, Cyclostationary signal analysis, in: V.K. Madisetti, D.B. Williams (Eds.), Digital Signal Processing Handbook, CRC Press LLC, Boca Raton, 1999.
[17] F. Giri, Y. Rochdi, A. Brouri, F.Z. Chaoui, Parameter identification of Hammerstein systems containing backlash operators with arbitrary-shape parametric borders, Automatica 47 (2011) 1827–1833.
[18] F. Giri, Y. Rochdi, F.Z. Chaoui, A. Brouri, Identification of Hammerstein systems in presence of hysteresis-backlash and hysteresis-relay nonlinearities, Automatica 44 (2008) 767–775.
[19] F. Giri, E.W. Bai (Eds.), Block-oriented Nonlinear System Identification, Springer Verlag, London, 2010.
[20] Q.P. He, J. Wang, M. Pottmann, S.J. Qin, A curve fitting method for detecting valve stiction in oscillating control loops, Ind. Eng. Chem. Res. 46 (2007) 4549–4560.
[21] M. Jelali, Estimation of valve stiction in control loops using separable least squares and global search algorithms, J. Process Control 18 (2008) 632–642.
[22] M. Jelali, B. Huang (Eds.), Detection and Diagnosis of Stiction in Control Loops: State of the Art and Advanced Methods, Springer Verlag, London, 2010.
[23] S. Karra, M.N. Karim, Comprehensive methodology for detection and diagnosis of oscillatory control loops, Control Eng. Pract. 17 (2009) 939–956.
[24] H.K. Khalil, Nonlinear Systems, third ed., Prentice Hall, Englewood Cliffs, N.J, 2002.
[25] L. Ljung, System Identification: Theory for the User, second ed., Prentice Hall, Englewood Cliffs, N.J., 1999.
[26] C.M. Loeffler, C.S. Burrus, Optimal design of periodically time-varying and multirate digital filters, IEEE Trans. Acoust. Speech Signal Process. 32 (1984) 991–997.
[27] A. Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, New York, 1965.
[28] M.A. Paulonis, J.W. Cox, A practical approach for large-scale controller performance assessment, diagnosis, and improvement, J. Process Control 13 (2003) 155–168.
[29] U. Nallasivam, S. Babji, R. Rengaswamy, Stiction identification in nonlinear process control loops, Comput. Chem. Eng. 34 (2010) 1890–1898.
[30] T. Söderström, P. Stoica, System Identification, Prentice Hall, London, 1989.
[31] R. Srinivasan, R. Rengaswamy, S. Narasimhan, R. Miller, Control loop performance assessment. 2. Hammerstein model approach for stiction diagnosis, Ind. Eng. Chem. Res. 44 (2005) 6719–6728.
[32] N.F. Thornhill, A. Horch, Advances and new directions in plant-wide disturbance detection and diagnosis, Control Eng. Pract. 15 (2007) 1196–1206.
[33] F. Tjarnstrom, L. Ljung, Variance properties of a two-step ARX estimation procedure, Eur. J. Control 9 (2003) 422–430.
[34] T.H. Van Pelt, D.S. Bernstein, Nonlinear system identification using Hammerstein and nonlinear feedback models with piecewise linear static maps, Int. J. Control 74 (2001) 1807–1823.
[35] J. Voros, Modeling and identification of systems with backlash, Automatica 46 (2010) 369–374.

[36] J. Wang, T. Chen, B. Huang, Closed-loop identification via output fast sampling, J. Process Control 14 (2004) 555–570.
[37] J. Wang, T. Chen, B. Huang, Cyclo-period estimation for discrete-time cyclostationarity signals, IEEE Trans. Signal Process. 54 (2006) 83–94.
[38] A. Sano, T. Chen, B. Huang, Identification of Hammerstein systems without explicit parameterisation of non-linearity, Int. J. Control 82 (2009) 937–952.
[39] J. Wang, Q. Zhang, Identification of an extended Hammerstein system with input hysteresis nonlinearity for control valve stiction characterization, in: The 16th IFAC Symp. System Identification, Brussels, Belgium, July 11–13, 2012, 2012, pp. 268–273.
[40] Y. Zhu, Multivariable System Identification for Process Control, Elsevier Science, Oxford, 2001.