Nonlinear Dyn (2015) 79:1257–1273
DOI 10.1007/s11071-014-1740-3

ORIGINAL PAPER

Identification of extended Hammerstein systems with hysteresis-type input nonlinearities described by Preisach model

Lei Fang · Jiandong Wang · Qinghua Zhang

Received: 21 May 2014 / Accepted: 1 October 2014 / Published online: 14 October 2014

Abstract This paper introduces Preisach model to describe hysteretic input nonlinearities of extended Hammerstein systems. Preisach model is proved to be an effective model to cover most of hysteretic input nonlinearities studied in the existing literatures on identification of extended Hammerstein systems. That is, Preisach model encloses the backlash-type nonlinearities and the hysteresis-relay nonlinearities as special cases, but excludes the switch-type nonlinearities. Hence, the usage of Preisach model is able to resolve an issue that types of hysteretic input nonlinearities are often inaccessible in practice, but most of existing identification methods have to know the type of input hysteretic nonlinearity a priori. For identification of extended Hammerstein systems with input nonlinearity described by Presiach model, the persistently exciting (PE) condition of input is established and the consistency of estimated model parameters is proved. In particular, the requirement on inputs to meet with the PE condition covers a wide class of signals, which resolves another issue in the existing studies that specially designed inputs are usually required. Numerical and industrial examples are provided to illustrate the obtained results.

Keywords System identification · Hammerstein system · Backlash · Relay · Switch · Preisach model · Persistent excitation

L. Fang · J. Wang (✉)
College of Engineering, Peking University, Beijing 100871, China
e-mail: jiandong@pku.edu.cn

L. Fang
e-mail: fanglei@pku.edu.cn

Q. Zhang
INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France
e-mail: qinghua.zhang@irisa.fr

1 Introduction

Hammerstein system, consisting of a static input nonlinearity followed by a linear time-invariant (LTI) subsystem, is a class of block-oriented nonlinear systems, widely used as a mathematical model to describe many physical and engineering systems [12,15]. If the static input nonlinearity is extended with some particular memory effects, then the resulting system has a structure similar to the conventional Hammerstein system, but for the purpose of distinction, it is referred to as *extended* Hammerstein system in this context[1].

Identification of extended Hammerstein systems has been receiving an increasing attention. Cheng and Yu [11] estimated the width of backlash hysteresis (the same as Fig. 1a having parallel borders with $m_l = m_r = m$) as well as the ultimate gain and frequency

[1] In Giri et al. [14] and Rochidi et al. [22], the extended Hammerstein system is respectively referred to as "Hammerstein system in presence of memory hard nonlinearity" and "Hammerstein-like system" to distinguish from conventional Hammerstein system.

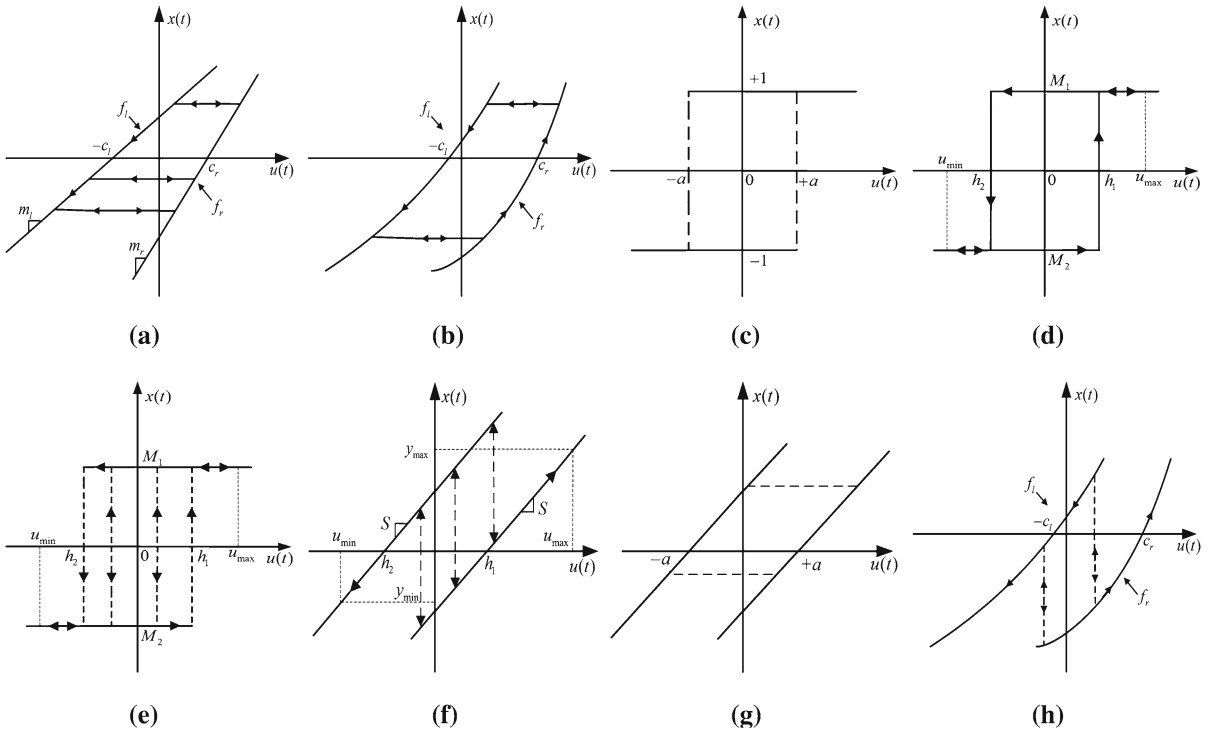

Fig. 1 Hysteresis nonlinearities with input $u(t)$ and output $x(t)$: **a** backlash with linear boundaries in Cerone and Regruto [7], Miyashita and Yamakita [20], Vörös [29], Cheng and Yu [11] and Giri et al. [14]; **b** general backlash in Vörös [30], Rochdi et al. [22], and Giri et al. [13]; **c** hysteresis relay in Bai [2]; **d** hysteresis relay of type 1 in Giri et al. [14]; **e** hysteresis relay of type 2 in Giri et al. [14]; **f** hysteresis backlash of type 2 in Giri et al. [14]; **g** hysteresis switch in Bai [2]; **h** general switch in Rochdi et al. [22]

of linear subsystems based on relay feedback tests. Bai [2] identified two hysteresis input nonlinearities parameterized by one single parameter, namely hysteresis relay (Fig. 1c) and hysteresis switch (Fig. 1g), based on the idea of separable least squares. Cerone and Regruto [7] derived parameter bounds of linear subsystems with input backlash nonlinearity (Fig. 1a), when the output measurement errors are bounded. Miyashita and Yamakita [20] employed binary sequences as input to decouple the identification of dynamic block from that of piecewise hysteresis nonlinearities (Fig. 1a). Giri et al. [14] designed specific inputs to identify linear subsystems and two types of hysteresis backlash (Fig. 1a with $m_l = m_r = m$ and Fig. 1f) and two types of hysteresis-relay nonlinearities (Fig. 1d, e) in a separated manner. Wang et al. [31] exploited input's piecewise constant property to estimate the unmeasurable inner signals so that the proposed identification approach is applicable to Hammerstein systems with nonlinearities such as backlash that can preserve the input's piecewise constant property. Vörös [29,30] applied the key-term separation principle to identification of cascaded systems consisting of linear subsystems following input backlash with linear borders (Fig. 1a) and with polynomial borders (Fig. 1b), respectively. Rochdi et al. [22] and Giri et al. [13] designed two experiments to decompose the identification of Hammerstein-like systems containing the general backlash (Fig. 1b) or the general switch (Fig. 1h) into two simpler problems involving static nonlinearities.

There are two main issues in the above-cited articles. First, specific type of hysteresis in terms of the shape of hysteresis loop is assumed to be valid, and some particular identification method is proposed to estimate unknown parameters by taking advantages of special characters of this specific geometrical shape of the hysteresis loop. In practice, however, types of hysteretic

input nonlinearities are often inaccessible, so that it is difficult to make a valid choice from the existing identification methods; in addition, the identification method proposed for one specific input nonlinearity is hardly applicable to different types of nonlinearities. Hence, it is desirable to identify a general model that can enclose all (or most of) the hysteretic input nonlinearities in Fig. 1 as special cases. Second, special inputs, each specific to a type of hysteresis, are often required in the above-cited articles, e.g., the rectangular waves [11], the binary sequences [20], the piecewise constant inputs [31], and the specially designed inputs [13,14,22]. As a result, the identification methods proposed in these articles cannot be applied to the case where these special inputs are not allowed or inputs are not subject to users' design. These two issues motivate us to address the following questions in this study:

Q1 Does there exist a hysteresis model that can cover all or most of the hysteretic input nonlinearities presented in Fig. 1?

Q2 If the answer to Q1 is positive, can the hysteresis model be identified under a general type of input, releasing the requirement on special inputs in the existing articles?

The contributions of this paper are twofold. First, we prove that Preisach model covers the backlash-type nonlinearities in Fig. 1a, b, and the hysteresis-relay nonlinearities in Fig. 1c, d, but excludes the switch-type nonlinearities in Fig. 1e–h. Second, we establish the persistently exciting (PE) condition of inputs for the identification of extended Hammerstein systems involving the Preisach model, and prove the consistency of estimated model parameters. These contributions provide partial, but significant, answers to the questions Q1 and Q2: a single model covers half[2] of the cases presented in Fig. 1, and the same type of input excitation allows consistent model estimation in all the covered cases.

Even though Preisach model is a well-known mathematical model in the field of magnetic, ferromagnetic, and smart materials [6,19,28], the relations between Preisach model and the nonlinearities in Fig. 1 investigated in the field of system identification have not been studied to the best of our knowledge. Besides enclosing the nonlinearities in Fig. 1a–d, Preisach model also has the merit of simplicity in terms of requirements on exciting inputs for identification of extended Hammerstein systems. The requirement on the associated PE condition (to be established in Proposition 3) is not restrictive and covers a wide class of signals; thus, identification can be made under quite general experimental conditions, relieving the different specific requirements in most of related works on special inputs.

It may be necessary to clarify the connection between this study on the identification of extended Hammerstein systems with hysteresis-type input nonlinearity and the existing works on the identification of standalone hysteretic systems. The modeling and control of hysteretic systems have been active research topics [6,19,28]. A fundamental difference between these two topics is that the output of a standalone hysteretic system is directly measurable, while the output of the hysteresis block in an extended Hammerstein system is unavailable for system identification. It is a well-known fact that the unavailability of inner signals between blocks is the main challenge for identification of block-oriented nonlinear systems [12].

The rest of the paper is organized as follows. Section 2 formulates the identification problem. Section 3 investigates the ability of Preisach model in describing the hysteresis nonlinearities depicted in Fig. 1. Section 4 is on the identification method for extended Hammerstein systems. Section 5 establishes the PE conditions of input and the consistency of estimated model parameters. Section 6 presents numerical and industrial examples to support the obtained results. Finally, a conclusion is given in Sect. 7.

2 Problem formulation

Consider an extended Hammerstein system, depicted in Fig. 2,

$$x(t) = f[u, \xi](t), \tag{1}$$

$$A(q^{-1})y(t) = B(q^{-1})x(t) + e(t), \tag{2}$$

where $A(q^{-1}) = 1+a_1q^{-1}+\cdots+a_{n_a}q^{-n_a}$, $B(q^{-1}) = b_1q^{-n_k-1} + \cdots + b_{n_b}a^{-n_k-n_b}$, $u(t)$, $y(t)$, $x(t)$, and $e(t)$ are the discrete time input, output, inner signal and noise, respectively. The inner signal $x(t)$ and the noise $e(t)$ are not measurable. Symbol q^{-1} stands for a one-sample delay operator, i.e., $q^{-1}x(t) = x(t-1)$.

[2] In fact, most of the existing literatures on identification of extended Hammerstein systems are focused on the backlash-type and hysteresis-relay nonlinearities, not on the switch-type nonlinearities.

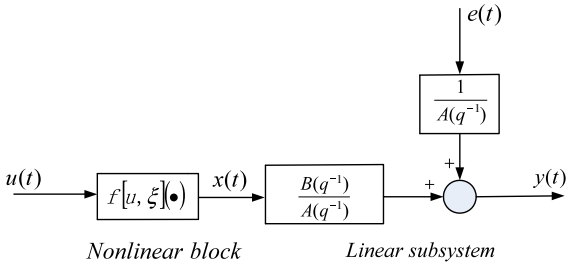

Fig. 2 The diagram of an extended Hammerstein system

With a temporary abuse of t as a continuous variable, the nonlinear block $f[u, \xi](\cdot)$ is described by Preisach model [18,19],

$$f[u, \xi](t) = \iint_{(\beta,\alpha)\in P} \mu(\beta, \alpha)\gamma_{\beta\alpha}[u, \xi](t)\, \mathrm{d}\beta\mathrm{d}\alpha. \tag{3}$$

Here, P stands for the so-called effective Preisach plane,

$$P \triangleq \{(\beta, \alpha) \mid \beta \geq u_{\min},\ \alpha \leq u_{\max},\ \beta \leq \alpha\}, \tag{4}$$

where the range of $u(t)$ is denoted as $[u_{\min}, u_{\max}]$. The notation $\mu(\beta, \alpha)$ is a weighting function associated with the point $(\beta, \alpha) \in P$. The symbol $\gamma_{\beta\alpha}[\cdot, \cdot]$ is the relay operator

$$\gamma_{\beta\alpha}[u, \zeta](t) = \begin{cases} +1, & \text{if } u(t) > \alpha, \\ -1, & \text{if } u(t) < \beta, \\ \gamma_{\beta\alpha}[u, \zeta](t^-) \triangleq \lim_{t'<t, t'\to t} \gamma_{\beta\alpha}[u, \zeta](t'), & \text{if } \beta \leq u(t) \leq \alpha, \end{cases}$$

where ζ stands for the initial value $\gamma_{\beta\alpha}[u, \zeta](0^-)$ of the operator. Thus, ξ in (3) represents the initial states of all the relay operators on the entire Preisach plane P. Clearly, the inner signal $x(t)$ in (1) depends on the current and previous samples of $u(t)$. For the simplicity of notation, $f[u, \xi](\cdot)$ is represented by $f(\cdot)$ in the sequel.

The following assumptions are made throughout the paper:

- A1 The noise $e(t)$ is a zero-mean white noise with finite variance and is independent of the input $u(t')$ for all values of $t' \leq t$.
- A2 The linear subsystem is asymptotically stable, i.e., the polynomial $A(q^{-1})$ has all zeros inside the unit circle. The polynomials $A(q^{-1})$ and $B(q^{-1})$ are coprime, i.e., there is no common factor between $A(q^{-1})$ and $B(q^{-1})$.

Under Assumption A1, the linear subsystem in (2) is an auto-regression with extra input (ARX) model, possibly operating in a feedback control loop having at least one-sample delay in the loop in order to avoid algebraic loops [24]. A high-order ARX model can approximate any LTI dynamic systems arbitrarily well [17] (p. 336); in addition, even if the noise additive to $y(t)$ is heavily colored, a consistent model of the linear subsystem can be obtained via a high-order ARX model estimation and a subsequent model reduction (see e.g., Chapter 6 of [33] and [27]). For these reasons, the ARX model is chosen here to describe the linear subsystem. Assumption A2 is standard in the identification of linear systems for the establishment of PE condition and parameter consistency [17].

With the measurements $\{u(t), y(t)\}_{t=1}^{N}$, the identification objective is to estimate the unknown parameters of Preisach model $f(\cdot)$, and those of linear subsystem, $a \triangleq [a_1, \ldots, a_{n_a}]^T$ and $b \triangleq [b_1, \ldots, b_{n_b}]^T$, associated with the structure parameters (n_a, n_b, n_k).

3 Relation between Presiach model and hysteresis-type nonlinearities in Fig. 1

This section investigates the ability of Preisach model in describing the hysteresis-type nonlinearities depicted in Fig. 1. It will be proved that Preisach model includes the backlash-type nonlinearities in Fig. 1a, b and the hysteresis-relay nonlinearities in Fig. 1c, d as special cases, but excludes the switch-type nonlinearities in Fig. 1d–g. In order to make the paper self-sustained, some results about Preisach model [6,19,28] are given in Appendix "1" and "2".

3.1 Backlash-type nonlinearities

The backlash-type nonlinearities include the backlash with linear boundaries in Fig. 1a and the general backlash with polynomial or arbitrary-shape borders in Fig. 1b. They have one common property: when the input changes its direction, the output first holds and then continuously varies along some borders. Among them, the backlash nonlinearity with linear boundaries in Fig. 1a is chosen here as the representative for investigation, because the result to be established can be

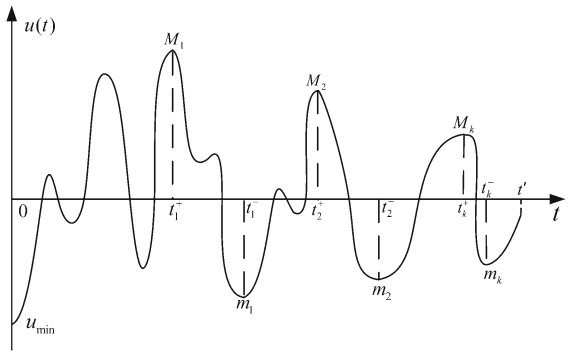

Fig. 3 An illustrative input $u(t)$ in the proof of Proposition 1

extended analogously to the general backlash in Fig. 1b. The mathematic model of the backlash nonlinearity in Fig. 1b is [7],

$$x(t) = \begin{cases} f_l(u(t)) \triangleq m_l\,(u(t) + c_l)\,, & \text{if } u(t) \le z_l(t), \\ f_r(u(t)) \triangleq m_r\,(u(t) - c_r)\,, & \text{if } u(t) \ge z_r(t), \\ x(t-1), & \text{if } z_l(t) < u(t) < z_r(t), \end{cases} \tag{5}$$

where $m_l > 0, m_r > 0, c_l > 0, c_r > 0$ are constant parameters and

$$z_l(t) = \frac{x(t-1)}{m_l} - c_l, \quad z_r(t) = \frac{x(t-1)}{m_r} + c_r.$$

Proposition 1 *The backlash nonlinearity (5) with linear boundaries in Fig. 1a can be described by Preisach model. The resulting weighting function $\mu(\beta, \alpha)$ is a one-dimensional impulse function on a specific line $\beta = f_l^{-1}(f_r(\alpha))$.*

Proof Owing to Representation Theorem in [18] (Theorem 1 in Appendix "1"), it is sufficient to prove that the backlash nonlinearity (5) satisfies the three required properties, namely the wiping-out property, congruency property, and rate-independence property.

First, let us prove the wiping-out property, which says that the series of dominant input extrema wipe out all other input extrema, by inductive reasoning as follows. For a general input $u(t)$ with $t \in [0, t']$ in the range $[u_{\min}, u_{\max}]$, the dominant input maximum M_k and the dominant input minimum m_k for a positive integer k are defined, respectively, as,

$$M_k = \max_{t \in [t_{k-1}^-, t']} u(t),\ u(t_k^+) = M_k, \tag{6}$$

$$m_k = \min_{t \in [t_k^+, t']} u(t),\ u(t_k^-) = m_k. \tag{7}$$

with $t_0^- \triangleq 0$. Figure 3 presents an illustrative example of $u(t)$, where M_k and m_k denote the inverses of limiting descending branch $f_l(u)$ and limiting ascending branch of the backlash nonlinearity as $f_l^{-1}(x)$ and $f_r^{-1}(x)$, respectively. The minimum and maximum of $x(t)$ on the input interval $[u_{\min}, u_{\max}]$ are $x_{\min} \triangleq f_l(u_{\min})$ and $x_{\max} \triangleq f_r(u_{\max})$, respectively. From the definition of backlash in (5), only two cases may occur for the dominant input maximum M_1:

- In the case of $M_1 \in [u_{\min}, f_r^{-1}(x_{\min})]$, $u(t)$ for $t \in [0, t']$ always gives a constant output equal to $x_{\min}$. Thus, all the previous local input extrema for $t < t_1^+$ are wiped out by M_1.
- In the case of $M_1 \in (f_r^{-1}(x_{\min}), u_{\max}]$, $x(t)$ is increased along the right border f_r to $f_r(M_1)$. No matter what the input path for $t \in [0, t_1^+]$ is, the definition of backlash in (5) says that all the previous input extrema for $t < t_1^+$ do not affect $x(t)$ after the time instant $t = t_1^+$. That is, all the previous input extrema are wiped out by M_1.

Similarly, two cases need to be discussed for the dominant input minimum m_1:

- In the case of $m_1 \in [f_l^{-1}(f_r(M_1)), M_1]$ with $M_1 \in (f_r^{-1}(x_{\min}), u_{\max}]$, $x(t)$ moves along the inner horizontal line connecting the two borders f_l and f_r. The input $u(t)$ for $t \in (t_1^+, t']$ always gives a constant output value equal to $f_r(M_1)$. Thus, the wiping-out property is satisfied.
- In the case of $m_1 \in [u_{\min}, f_l^{-1}(f_r(M_1)))$, $x(t)$ is decreased along the left border f_l to $f_l(m_1)$. The definition of backlash in (5) says that the previous local input extrema for $t \in [t_1^+, t_1^-]$ do not affect $x(t)$ after the time instant $t = t_1^-$. That is, the previous local input extrema in $(t_1^+, t_1^-]$ are wiped out by m_1.

Analogously to M_1 and m_1, the other dominant input extreme series $\{M_2, M_3, \ldots\}$ and $\{m_2, m_3, \ldots\}$ wipe out the effects of all the intermediate local input extreme values appeared before the time instants t_k^+ and t_k^-, respectively. Therefore, the wiping-out property holds for the backlash (5).

Second, the congruency property is proved, namely all hysteresis loops corresponding to the same extreme values of input are congruent. When $u(t)$ oscillates between two extreme values u_1 and u_2 for $u_1 < u_2$, there exists two cases forming hysteresis loops as shown in Fig. 4:

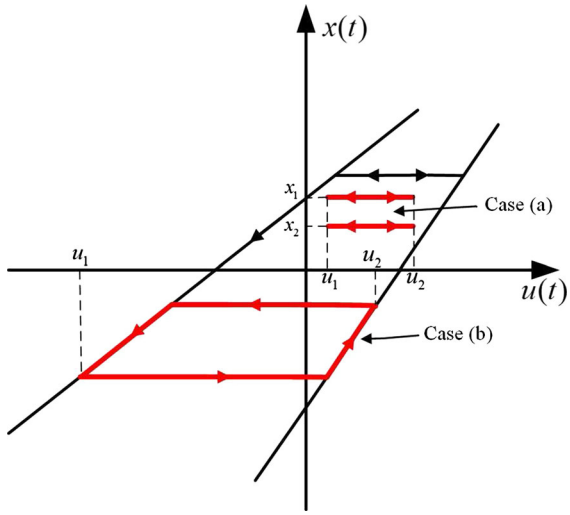

Fig. 4 Hysteresis loops for the proof of congruency property in Proposition 1

(a) In the case that the input oscillates between the left and right borders, no matter what the initial outputs are, the resulting hysteresis loops are degenerated to horizontal lines (shown as Case (a) in Fig. 4), which are congruent in the geometric sense.
(b) In the case that the input interval $[u_1, u_2]$ is large enough, the output increases along the right border to the local maximum $f_r(u_2)$ and decreases along the left border to the local minimum $f_l(u_1)$. Thus, only one hysteresis loop is formed (Case (b) in Fig. 4), and the congruency property certainly holds.

Next, the backlash nonlinearity in (5) is clearly rate-independent, i.e., $x(t)$ is independent of the variation velocity of $u(t)$. Therefore, the backlash nonlinearity satisfies the wiping-out property, congruency property, and rate-independence property, and can be described by Preisach model.

Finally, we derive the weighting function based on the first-order reversal functions defined in [19] (recalled at Appendix "2"). The first-order descending reversal function $f_{\alpha'\beta'}$ is the path of $x(t)$ for a monotonic increment of $u(t)$ starting from the minimum value $u_{\min}$ to the value α' (denoted as $f_{\alpha'}$) followed by a subsequent monotonic decrement to β'. Thus, the weighting function $\mu(\beta, \alpha)$ for the backlash (5) is obtained as follows.

- In the case of $\alpha' \in [u_{\min}, f_r^{-1}(x_{\min})]$, Eq. (5) yields $f_{\alpha'} = f_{\alpha'\beta'} = x_{\min}$ so that (43) and (44) (appeared later in Appendix "2") lead to

$$\mu(\beta, \alpha) = 0, \quad \forall \alpha \in [u_{\min}, f_r^{-1}(x_{\min})], \quad \forall \beta \in [u_{\min}, \alpha].$$

- In the case of $\alpha' \in (f_r^{-1}(x_{\min}), u_{\max}]$ and $\beta' \in (f_l^{-1}(f_r(\alpha')), \alpha']$, Eq. (5) yields $f_{\alpha'\beta'} = f_{\alpha'} = f_r(\alpha') = m_r(\alpha' - c_r)$ so that

$$\mu(\beta, \alpha) = 0, \quad \forall \alpha \in (f_r^{-1}(x_{\min}), u_{\max}], \quad \forall \beta \in (f_l^{-1}(f_r(\alpha)), \alpha].$$

- In the case of $\alpha' \in (f_r^{-1}(x_{\min}), u_{\max}]$ and $\beta' \in [u_{\min}, f_l^{-1}(f_r(\alpha'))]$, Eq. (5) gives $f_{\alpha'} = f_r(\alpha') = m_r(\alpha' - c_r)$ and $f_{\alpha'\beta'} = f_l(\beta') = m_l(\beta' + c_l)$. As a result, eqs. (43) and (44) give

$$\mu(\beta, \alpha) = 0, \quad \forall \alpha \in (f_r^{-1}(x_{\min}), u_{\max}], \quad \forall \beta \in [u_{\min}, f_l^{-1}(f_r(\alpha)).$$

- In the case of $\alpha' \in (f_r^{-1}(x_{\min}), u_{\max}]$ and $\beta' = f_l^{-1}(f_r(\alpha'))$, it is ready to obtain from the geometric interpretation of the first-order descending reversal curve (similar to Fig. 13 in Appendix "2"),

$$\frac{1}{2}\left(f_l(\beta') - f_l(\beta'_-)\right) = \iint_{T(\beta'_-, \alpha') - T(\beta', \alpha')} \mu(\beta, \alpha) \mathrm{d}\beta \mathrm{d}\alpha \tag{8}$$

where β'_- is the limiting point approaching β' from the left-hand side. The set $T(\beta'_-, \alpha') - T(\beta', \alpha')$ forms a trapezoid in the Presiach plane with four vertices (β'_-, β'_-), (β', β'), (β', α'), and (β'_-, α'). A counterpart of (8) can be similarly obtained for the first-order ascending reversal function, which is the path of $x(t)$ for a monotonic decrement of $u(t)$ starting from the maximum value $u_{\max}$ to the value β'' followed by a subsequent monotonic increment to α''. That is, in the case of $\beta'' \in [u_{\min}, f_l^{-1}(x_{\max}))$ and $\alpha'' = f_r^{-1}(f_l(\beta''))$, we have

$$\frac{1}{2}\left(f_r(\alpha'') - f_r(\alpha''_-)\right) = \iint_{T(\beta'', \alpha'') - T(\beta'', \alpha''_-)} \mu(\beta, \alpha) \mathrm{d}\beta \mathrm{d}\alpha, \tag{9}$$

where α''_- is the limiting point approaching α'' from the left-hand side. The set $T(\beta'', \alpha'') - T(\beta'', \alpha''_-)$ forms a trapezoid in the $\alpha - \beta$ plane with four vertices (α'', α''), (α''_-, α''_-), (β'', α'') and (β'', α''_-). If $\beta'' = \beta'_-$, then the left-hand sides of (8) and (9) are

the same, i.e.,

$$f_l(\beta') - f_l(\beta'_-) = f_l\left(f_l^{-1}\left(f_r(\alpha'')\right)\right) - f_l\left(f_l^{-1}\left(f_r(\alpha''_-)\right)\right) = f_r(\alpha'') - f_r(\alpha''_-).$$

The intersection of the set $T(\beta'_-, \alpha') - T(\beta', \alpha')$ and $T(\beta'', \alpha'') - T(\beta'', \alpha''_-)$ is a rectangular with the vertices (β'_-, α'), (β', α'_-), (β', α') and (β'_-, α'_-), and the weighting function $\mu(\beta, \alpha)$ is nonzero only in this rectangular, i.e.,

$$\int_{\beta'_-}^{\beta'}\int_{\alpha'_-}^{\alpha'} \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha = \frac{1}{2}\left(f_l(\beta') - f_l(\beta'_-)\right) = \frac{1}{2}\left(f_r(\alpha') - f_r(\alpha'_-)\right),$$

which says that $\mu(\beta, \alpha)$ is a one-dimensional impulse lying on the line $\beta = f_l^{-1}(f_r(\alpha))$. □

The proof of Proposition 1 can be generalized to the general backlash with polynomial or arbitrary-shape boundaries in Fig. 1b.

Corollary 1 *The general backlash nonlinearity with two polynomial or arbitrary-shape boundaries $f_l(u)$ and $f_r(u)$ depicted in Fig. 1b can be described by Preisach model. The weighting function $\mu(\beta, \alpha)$ is a one-dimensional impulse function on a specific curve $\beta = f_l^{-1}(f_r(\alpha))$.*

Remark 1 Theorem 2.7 in Sect. 4 of Visintin [28] also provides a conclusion that Presiach model can describe the backlash nonlinearity. However, the proof therein was based on the measure theory and a connection between a generalized play operator and Preisach model, while the proof here is very different and more constructive: we strictly follow the Representation Theorem (Theorem 1 in Appendix "1") of Preisach model. In addition, Proposition 1 establishes the weighting function $\mu(\beta, \alpha)$ for the backlash nonlinearity; this result will be verified later in Example 1 at Sect. 6.

3.2 Switch-type nonlinearities

The switch-type nonlinearities include hysteresis relay of type 2 in Fig. 1e, hysteresis backlash of type 2 in Fig. 1f, hysteresis switch in Fig. 1g and general switch in Fig. 1h. Among them, the hysteresis backlash of type 2 in Fig. 1f is a representative of the above switch-type nonlinearities, because the proof of Proposition 2 given next can be extended to other types of switch nonlinearities. The mathematical model of the nonlinearity in Fig. 1f is [14]

$$x(t) = \begin{cases} S\left(u(t) - h_1\right), & \text{if } u(t) > u(t-1), \\ S\left(u(t) - h_2\right), & \text{if } u(t) < u(t-1), \\ y(t-1), & \text{if } u(t) = u(t-1), \end{cases} \tag{10}$$

where S, h_1 and h_2 are constant positive parameters, and $h_1 > h_2$.

Proposition 2 *The nonlinearity in (10) cannot be described by Preisach model.*

Proof This is proved by contradiction. We first assume that this nonlinearity can be described by Preisach model. When $u(t)$ changes its direction at the value $u(t) = \alpha'$ from increment to decrement, even if the decrement is infinitesimal, i.e., $\beta' = \alpha'_-$, where α'_- denotes the limiting point approaching α from the left-hand side, the nonlinearity output $x(t)$ in (10) immediately jumps from $S\left(\alpha' - h_1\right)$ to $S\left(\alpha'_- - h_2\right)$. Thus, the geometric representation of Presiach model (Appendix "2") says that

$$\iint_{T(\alpha'_-,\alpha')} \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha = \frac{S}{2}(\alpha' - \alpha'_- + h_2 - h_1) \neq 0. \tag{11}$$

Here, $T(\beta', \alpha')$ is a triangle area in the Preisach plane with three vertices (β', β'), (α', α') and (β', α'). Equation (11) implies that $\mu(\beta, \alpha)$ must be a two-dimensional impulse function on the line $\beta' = \alpha'$, which causes a contradiction shown as follows. From (42) (in Appendix "2") and the definition of Preisach plane partition $P = S^+(t) + S^-(t)$, we have

$$x(t) = -\iint_P \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha + 2\iint_{S^+(t)} \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha = x_{\min} + 2\iint_{S^+(t)} \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha. \tag{12}$$

where $x_{\min}$ is the value of $x(t)$ associated with the minimum value $u_{\min}$ of $u(t)$. If $u(t)$ is monotonically increased from $u_{\min}$ to an arbitrary value $\alpha' > u_{\min}$,

then (11) and (12) yield

$$\begin{aligned} x(t) &= x_{\min} + 2 \iint_{T(u_{\min},\alpha')} \mu(\beta,\alpha)\mathrm{d}\beta\mathrm{d}\alpha \\ &= x_{\min} + S(h_2 - h_1) \times \infty \\ &= -\infty. \end{aligned} \tag{13}$$

Here, the symbol ∞ on the right hand of second equality stands for the infinite points on the line $\beta = \alpha$ inside the triangle $T(u_{\min}, \alpha')$, where the weighting function $\mu(\beta, \alpha)$ is nonzero. Equation (13) clearly contradicts a fact that the output of the nonlinearity in (10) always takes a finite value. Hence, the nonlinearity in (10) cannot be represented by Preisach model. □

The proof of Proposition 2 for the nonlinearity in Fig. 1f can be extended to other types of switch-type nonlinearities in Fig. 1e, g, h.

Corollary 2 *The switch-type nonlinearities in Fig. 1e, g, h cannot be described by Preisach model.*

Proposition 2 and Corollary 2 say that Preisach model cannot describe the switch-type nonlinearities in Fig. 1e–h. These switch-type nonlinearities share a common character, namely the output jumps from one border to another one, as long as the input changes its direction possibly in one limited or entire input range. As shown in the proof of Proposition 2, such a discontinuous behavior leads to two-dimensional impulse-type weighting functions associated with infinite points on the Presiach plane, which conflicts with a fact that the nonlinearity output always takes a finite value. For the backlash nonlinearities in Proposition 1 and Corollary 1, the nonlinearity output continuously varies (without jumpy behaviors); this property leads to a one-dimensional impulse-type weighting function, which ensures that the output is always finite.

3.3 Hysteresis-relay nonlinearities

The proofs of Propositions 1 and 2 imply that if the nonlinearity output only jumps at specific points, instead of an interval with infinite points, then only some specific points on the Preisach plane are associated with two-dimensional impulse weighting functions so that the output is finite. Therefore, the hysteresis relay studied in [2] (Fig. 1c) and hysteresis relay of type 1 in [14] (Fig. 1d) can be described by Presiach model, because the output jumps only at two distinct places.

Corollary 3 *The hysteresis-relay nonlinearities in Fig. 1c, d can be described by Presiach model. The weighting functions are a two-dimensional impulse function on one specific point $(-a, a)$ for the nonlinearity in Fig. 1c and (h_2, h_1) for that in Fig. 1d, respectively.*

4 Identification method

This section presents the identification method to estimate unknown parameters of extended Hammerstein system involving Presiach model.

Preisach model has to be discretized for system identification purpose. The discretized Preisach model is [5,16,23],

$$x(t) = \sum_{i=1}^{L} \sum_{j=i}^{L} \mu_{ij} \gamma_{ij}(t) \tag{14}$$

where μ_{ij} is the weighting parameter and $\gamma_{ij}(t)$ is the discretized relay operator,

$$\gamma_{ij}(t) = \begin{cases} +1, & \text{if } u(t) > \frac{u_j + u_{j+1}}{2}, \\ -1, & \text{if } u(t) < \frac{u_i + u_{i+1}}{2}, \\ \gamma_{ij}(t-1), & \text{otherwise.} \end{cases} \tag{15}$$

Here, u_i is the discretized threshold

$$u_i = u_{\min} + (i-1)\delta, \quad i = 1, 2, \ldots, L+1, \tag{16}$$

where L is the discretization level and δ is the discretization step size,

$$\delta = \frac{u_{\max} - u_{\min}}{L}. \tag{17}$$

The connection between discretized Preisach model and its continuous counterpart was established by Theorem 4.1 in [5]. It states that, as the discretization level L approaches infinity, the output of the discretized Preisach model in (14) converges to its continuous counterpart in (3).

To obtain a compact representation of (14), stack all the weighting parameters μ_{ij}'s into a vector μ,

$$\begin{aligned} \mu &\triangleq [\mu_1, \ldots, \mu_K]^T \\ &\triangleq [\mu_{11}, \ldots, \mu_{1L}, \mu_{22}, \ldots, \mu_{2L}, \ldots, \mu_{LL}]^T, \\ &\text{for } K = \frac{L(L+1)}{2}, \end{aligned} \tag{18}$$

and define a state vector $\gamma(t)$ as

$$\begin{aligned}\gamma(t) &\triangleq \left[\gamma_1(t), \ldots, \gamma_K(t)\right]^T \\ &\triangleq \left[\gamma_{11}(t), \ldots, \gamma_{1L}(t), \gamma_{22}(t), \ldots, \right. \\ &\quad \left. \gamma_{2L}(t), \ldots, \gamma_{LL}(t)\right]^T. \end{aligned} \tag{19}$$

With μ in (18) and $\gamma(t)$ in (19), the discretized Preisach model (14) can be rewritten as

$$x(t) = \gamma^T(t)\mu. \tag{20}$$

Substituting (20) into (2) yields

$$y(t) = -\phi^T(t)a + b^T\Gamma(t)\mu + e(t), \tag{21}$$

where

$$\phi(t) = [y(t-1), \ldots, y(t-n_a)]^T,$$

$$\Gamma(t) = \begin{bmatrix} \gamma_1(t-n_k-1) & \cdots & \gamma_K(t-n_k-1) \\ \vdots & \ddots & \vdots \\ \gamma_1(t-n_k-n_b) & \cdots & \gamma_K(t-n_k-n_b) \end{bmatrix}.$$

The gain ambiguity between the parameter vectors b and μ is removed here by setting b to have a unit Euclidean norm $\|b\| = 1$ and the first nonzero element of b to be positive [1]. If the structure parameters (n_a, n_b, n_k) and the discretization level L are known a priori, the estimation of the unknown parameter vectors a, b and μ, based on the measurements $\{u(t), y(t)\}_{t=1}^N$, is a bilinearly parameterized estimation problem [1,32]. Many well-developed identification methods for conventional Hammerstein systems, such as the over-parameterized method [1,9] and the iterative method [4], are directly applicable.

Next, we present the over-parameterized method with regularization technique to estimate the parameter vectors a, b and μ. With the definitions

$$\begin{aligned}\theta &\triangleq [b_1\mu_1, \ldots, b_1\mu_K, b_2\mu_1, \ldots, b_2\mu_K, \ldots, \\ &\quad b_{n_b}\mu_1, \ldots, b_{n_b}\mu_K, a_1, \ldots, a_{n_a}]^T \end{aligned} \tag{22}$$

and

$$\begin{aligned}\psi(t) &\triangleq \gamma_1(t-n_k-1), \ldots, \gamma_K(t-n_k-1), \ldots, \\ &\quad \gamma_1(t-n_k-n_b), \ldots, \gamma_K(t-n_k-n_b), \\ &\quad -y(t-1), \ldots, -y(t-n_a)]^T, \end{aligned}$$

the bilinear equation (21) becomes

$$y(t) = \psi^T(t)\theta + e(t). \tag{23}$$

For the measurements $\{u(t), y(t)\}_{t=1}^N$, the over-parameterized method consists of the following steps:

Step 1. Compute the least square estimate

$$\hat{\theta} = [\hat{\theta}_1, \ldots, \hat{\theta}_{n_bK+n_a}]^T = \left(\Psi_N^T\Psi_N\right)^{-1}\Psi_N^T Y_N, \tag{24}$$

where

$$\begin{aligned}\Psi_N &\triangleq [\psi(1), \ldots, \psi(N)]^T, \\ Y_N &\triangleq [y(1), \ldots, y(N)]^T \end{aligned} \tag{25}$$

Step 2. Construct a $\mathbb{R}^{n_b\times K}$ matrix $\hat{\Theta}_{b\mu}$ from $\hat{\theta}$ as

$$\hat{\Theta}_{b\mu} = \begin{bmatrix} \hat{\theta}_1 & \hat{\theta}_2 & \cdots & \hat{\theta}_K \\ \hat{\theta}_{K+1} & \hat{\theta}_{K+2} & \cdots & \hat{\theta}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\theta}_{(n_b-1)K+1} & \hat{\theta}_{(n_b-1)K+2} & \cdots & \hat{\theta}_{n_bK} \end{bmatrix}$$

and perform the singular value decomposition,

$$\hat{\Theta}_{b\mu} = USV^T,$$

where $U \in \mathbb{R}^{n_b\times n_b}$, $S \in \mathbb{R}^{n_b\times K}$ and $V \in \mathbb{R}^{K\times K}$. The singular values of $\hat{\Theta}_{b\mu}$ are listed in the diagonal elements of S in a descending order, i.e., $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{n_b,K\}} \geq 0$.

Step 3. Let U_1 and V_1 be the first columns of U and V, respectively. Denote s_1 as the sign of the first nonzero element of U_1. The estimates of a, b and μ are obtained as

$$\hat{a} = \hat{\theta}(n_bK+1 : n_bK+n_a),\ \hat{b} = s_1U_1,\ \hat{\mu} = s_1\sigma_1V_1. \tag{26}$$

The total number K of the weighting parameters μ_{ij} is quadratically increased with the discretization level L defined as (17), namely $K = L(L+1)/2$. As a result, the increment of K may lead to the dimensional problem of the over-parameterized method that deteriorates the estimation accuracy [3]. To resolve this issue, the regularization technique for linear system identification (see e.g., [10,21]) can be introduced to improve the estimation accuracy. The regularized estimation may lead to smaller mean square error via some bias sacrifice but significant variance improvement.

The regularized optimization problem is formulated for (23) by following the counterpart in linear system identification (see, e.g., that in [10]) as

$$\hat{\theta} = \arg\min_{\theta}\left(y(t) - \psi^T(t)\theta\right) + \theta^T D\theta. \tag{27}$$

where the regularization matrix D is

$$D = \text{diag}\{d(1), d(2), \ldots, d(n_bK + n_a)\}$$

with the weighting factor $d(i)$ monotonically increasing with i for the first n_bK terms, e.g., $d(i) = 0.2\left(e^{\frac{5(i-1)}{n_bK}} - 1\right)$ for $i = 1, 2, \ldots, n_bK$, and for the rest n_a terms, e.g., $d(i) = 0.2\left(e^{\frac{5(i-n_bK-1)}{n_a}} - 1\right)$ for $i = n_bK + 1, \ldots, n_bK + n_a$. The solution of this regularized optimization problem is

$$\begin{aligned}\hat{\theta}_{\text{reg}} &= [\hat{\theta}_{1\text{reg}}, \ldots, \hat{\theta}_{(n_bK+n_a)\text{reg}}]^T \\ &= \left(\Psi_N^T\Psi_N + D\right)^{-1}\Psi_N^T Y_N, \end{aligned} \tag{28}$$

where Ψ_N^T and Y_N are defined in (25). The final estimates of a, b and μ are obtained by using $\hat{\theta}_{\text{reg}}$ to replace $\hat{\theta}$ in (24) at Step 1 and following the rest Steps 2 and 3 of the over-parameterized method. Example 4 in Sect. 6 illustrates the effectiveness of this method.

5 Persistently exciting condition and consistency of estimation

In this section, we establish the persistently exciting (PE) condition of input for identification of extended Hammerstein systems with discretized Preisach model as the input nonlinearity and prove the consistency of estimated model parameters. Here, the structure parameters n_a, n_b, n_k and L are assumed to be known a priori; in addition, n_k is set to zero by properly shifting $u(t)$ for the ease of notation.

First, analogously to the concept of a strongly persistently exciting (SPE) signal for identification of conventional Hammerstein systems defined in [8,25], we define a similar concept for extended Hammerstein systems in this context:

Definition 1 The input $u(t)$ of extended Hammerstein system described by (2) and (20) is SPE of orders (n, K), denoted as SPE(n, K), where n is a positive integer and $K = \frac{L(L+1)}{2}$, if there exists a positive integer M and two positive constants c_1 and c_2, such that

$$c_1 I \le \sum_{t=t_0}^{t_0+M} \varphi(t)\varphi^T(t) \le c_2 I, \quad \forall t_0 \tag{29}$$

where the symbol I is the identity matrix and

$$\varphi(t) = [\gamma^T(t-1), \ldots, \gamma^T(t-n)]^T \tag{30}$$

with the K-dimensional vector $\gamma(t)$ defined in (19).

Second, we recall the sufficient PE condition established in Tan and Baras [26] for identification of the discretized Preisach model (20). To ease the understanding of this condition, we introduce $\widetilde{u}(t)$, referred to as the virtual input, obtained by rounding $u(t)$ to the nearest discretized threshold u_i in (16), i.e.,

$$\widetilde{u}(t) = \begin{cases} u_{L+1}, & \text{if } u(t) > \frac{1}{2}(u_L + u_{L+1}), \\ u_i, & \text{if } \frac{1}{2}(u_{i-1} + u_i) < u(t) \le \frac{1}{2}(u_i + u_{i+1}), \quad \text{for } i = 2, \ldots, L-1, \\ u_1, & \text{if } u(t) \le \frac{1}{2}(u_1 + u_2). \end{cases} \tag{31}$$

For two different inputs, if they share the same virtual input $\widetilde{u}(t)$, then the outputs of the discretized Preisach model are the same at all time instants.

Lemma 1 *(Theorem 3.2 in [26]) For the discretized Preisach model (20), the difference sequence of $\gamma(t)$ in (19),*

$$\gamma_d(t) \triangleq \gamma(t) - \gamma(t-1), \tag{32}$$

is PE, if there exists a positive integer M, such that for any t_0 one can find a sequence $\{u'(t)\}_{t=t_0}^{t_0+M-1}$ belonging to the PE equivalence class[3] $\{\underline{u(t)}\}_{t=t_0}^{t_0+M-1}$ of $u(t)$, where $\{u'(t)\}_{t=t_0}^{t_0+M-1}$ satisfies the following conditions: (a) The discretized thresholds u_i's for $i = 1, 2, \ldots, L+1$ (defined in (16)) are covered by local maxima and local minima of the virtual signal $\widetilde{u}'(t)$; (b) Two local maxima (minima) of $\widetilde{u}'(t)$ adjacent in time index t differ from each other by no more than one discretization step size δ in (17); (c) The series of local maxima of $\widetilde{u}'(t)$ is nonincreasing (nondecreasing), and the series of local minima is nondecreasing (nonincreasing).

Next, we generalize this important result in Lemma 1 to identification of extended Hammerstein systems described by (2) and (20).

Proposition 3 *Let a sequence $\{\bar{u}(k)\}_{k=t_0}^{t_0+M-1}$ be the one satisfying the PE conditions in Lemma 1. If the sample $\bar{u}(k)$ and n consecutive samples $\{u(nk-n+1), u(nk-n+2), \ldots, u(nk)\}$ of an input sequence $\{u(t)\}_{t=nt_0-n+1}^{n(t_0+M-1)}$ have the same virtual input value, then $u(t)$ is SPE(n, K).*

[3] See Section III-B and Fig. 5 in [26] for the definition of PE equivalence class and an illustrative example.

Proof In order to show that (29) holds here so that $u(t)$ is SPE(n, K), it is sufficient to prove that there is a positive integer N such that the inequality

$$\sum_{t=nt_0-n+1}^{nt_0-n+N} \varphi(t)\varphi^T(t) > 0 \tag{33}$$

is uniform for all t_0. Since the matrix on the left side of (33) is at least positive semi-definite for all t_0, it is equivalent to prove (33) by showing that the equality

$$z^T \left[\sum_{t=nt_0-n+1}^{nt_0-n+N} \varphi(t)\varphi^T(t) \right] z = 0 \tag{34}$$

holds if and only if the vector $z \triangleq [z_1^T, \ldots, z_n^T]^T$ is equal to zero, where $z_i \in \mathbb{R}^K$ for $i = 1, \ldots, n$. For a nonzero vector z, (34) is rewritten as

$$\sum_{t=nt_0-n+1}^{nt_0-n+N} z^T\varphi(t)\varphi(t)^T z = 0. \tag{35}$$

By selecting $N = nM$ and noticing that $z^T\varphi(t)$ and $\varphi(t)^T z$ are two equivalent scalars, (35) yields

$$\begin{aligned} z^T\varphi(t) = 0,\ t \\ = nt_0 - n+1, nt_0 - n+2, \ldots, n\,(t_0+M-1)\,. \end{aligned} \tag{36}$$

That is,

$$\underbrace{\begin{bmatrix} \gamma(nt_0-n) & \gamma(nt_0-n+1) & \cdots & \gamma\,(n\,(t_0+M-1)-1) \\ \gamma(nt_0-n-1) & \gamma(nt_0-n) & \cdots & \gamma\,(n\,(t_0+M-1)-2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(nt_0-2n+1) & \gamma(nt_0-2n+2) & \cdots & \gamma\,(n\,(t_0+M-1)-n) \end{bmatrix}}_{\Gamma} \begin{bmatrix} z_1^T \\ z_2^T \\ \cdots \\ z_n^T \end{bmatrix} = 0. \tag{37}$$

Because every n consecutive samples of the input sequence $\{u(t)\}_{t=nt_0-n+1}^{n(t_0+M-1)}$ share the same virtual input value, $\gamma_{ij}(t)$ in (15) and $\gamma(t)$ in (19) lead to the following equalities,

$$\gamma(nk-n+1) = \gamma(nk-n+2) = \cdots = \gamma(nk), \tag{38}$$

for $k = t_0, \ldots, t_0 + M - 1$. Owing to (38), subtracting the equality $z^T\varphi(t) = 0$ in (36) for adjacent values of t (i.e., subtracting the equality associated with the j-th column of Γ in (37) from the counterpart with the $(j+1)$-th column) gives

$$\begin{bmatrix} z_1^T & z_2^T & \cdots & z_n^T \end{bmatrix} \begin{bmatrix} \overline{\gamma}_d(t_0) & 0 & \cdots & 0 & \cdots & \overline{\gamma}_d(t_0+M-1) & 0 & \cdots & 0 \\ 0 & \overline{\gamma}_d(t_0) & \cdots & 0 & \cdots & 0 & \overline{\gamma}_d(t_0+M-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \overline{\gamma}_d(t_0) & \cdots & 0 & 0 & \cdots & \overline{\gamma}_d(t_0+M-1) \end{bmatrix} = 0, \tag{39}$$

where

$$\overline{\gamma}_d(j) \triangleq \gamma(n(j-1)+1) - \gamma(n(j-1)).$$

It follows from (39) that

$$\begin{aligned} z_i^T[\overline{\gamma}_d(t_0), \ldots, \overline{\gamma}_d(t_0+M-1)] = [0, \ldots, 0], \\ \text{for } i = 1, 2, \ldots, n. \end{aligned} \tag{40}$$

Because the sample $\bar{u}(k)$ and n consecutive samples $\{u(nk-n+1), u(nk-n+2), \ldots, u(nk)\}$ have the same virtual input value, $\{\overline{\gamma}_d(t)\}_{t=t_0}^{t_0+M-1}$ in (40) are the same as the counterparts of $\gamma_d(t)$ in (32) associated with $\{\bar{u}(k)\}_{k=t_0}^{t_0+M-1}$. It is known from the proof of Lemma 1 in [26] that $\{\gamma_d(t)\}_{t=1}^{M}$ spans $\mathbb{R}^K$, so that (40) says $z_i = 0, \quad \forall i$. □

Finally, the consistency of estimated model parameters for extended Hammerstein system from the over-parameterized method is established based on Proposition 3.

Proposition 4 *Assume that the true system can be described by the extended Hammerstein model described by (2) and (20) with known structure parameters n_a, n_b, n_k and L. If $u(t)$ is SPE(n_b, K), then the estimated parameter vectors $\hat{a}$, $\hat{b}$ and $\hat{\mu}$ in (26) from the over-parameterized method are consistent.*

Proof Since $\gamma(t)$ in (19) is solely related to $u(t)$, it is clear from Assumption A1 ($u(t)$ and $e(t)$ are independent) that $E\{\gamma(t)e(s)\} = 0$ for all t and s with $t \neq s$. Thus, $E\left\{\psi(t)\psi^T(t)\right\}$ with $\psi(t)$ defined in (23) becomes

$$
\begin{aligned}
R_\psi &\triangleq E\left\{\psi(t)\psi^T(t)\right\} \\
&= E\left\{\begin{bmatrix} \gamma(t-1) \\ \vdots \\ \gamma(t-n_b) \\ \hdashline \widetilde{\tilde{y}}(t-1) \\ \vdots \\ \widetilde{\tilde{y}}(t-n_a) \end{bmatrix} \begin{bmatrix} \gamma(t-1) \\ \vdots \\ \gamma(t-n_b) \\ \hdashline \widetilde{\tilde{y}}(t-1) \\ \vdots \\ \widetilde{\tilde{y}}(t-n_a) \end{bmatrix}^T\right\} \\
&\quad + E\left\{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hdashline \widetilde{\tilde{e}}(t-1) \\ \vdots \\ \widetilde{\tilde{e}}(t-n_a) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hdashline \widetilde{\tilde{e}}(t-1) \\ \vdots \\ \widetilde{\tilde{e}}(t-n_a) \end{bmatrix}^T\right\} \\
&\triangleq \left[\begin{array}{c:c} R_\varphi & R_{\varphi\widetilde{y}} \\ \hdashline R_{\varphi\widetilde{y}}^T & R_{\widetilde{y}} \end{array}\right] + \left[\begin{array}{c:c} \mathrm{O} & \mathrm{O} \\ \hdashline \mathrm{O} & R_{\widetilde{e}} \end{array}\right].
\end{aligned} \tag{41}
$$

Here, $\varphi(t)$ is defined in (30), and

$$
\widetilde{e}(t) \triangleq -\frac{1}{A(q^{-1})}e(t), \quad \widetilde{y}(t) \triangleq \frac{B(q^{-1})}{A(q^{-1})}\gamma^T(t)\mu.
$$

According to Assumptions A1 and A2, namely $e(t)$ is a white noise with zero mean, and the filter $\frac{1}{A(q^{-1})}$ is asymptotically stable, it follows from Lemma 3.1 in [25] that $R_{\widetilde{e}} > 0$. In addition, Proposition 3 says that the covariance matrix of $\varphi(t)$ in (30) is positive definite, i.e., $R_\varphi > 0$. Thus, it follows from Complement C5.1 and Lemma A.3 in [24] that the covariance matrix R_ψ in (41) is positive definite. Therefore, the consistency of $\hat{a}$, $\hat{b}$ and $\hat{\mu}$ in (26) can be established by following the same argument in the proofs of Lemma A1 and Theorem 2.1 in [1]. □

6 Examples

This section presents three numerical and one industrial examples to support the above-obtained results.

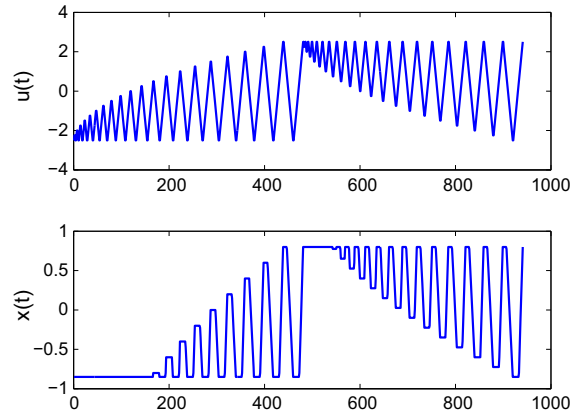

Fig. 5 The input $u(t)$ (*top*) and output $x(t)$ (*bottom*) in Example 1

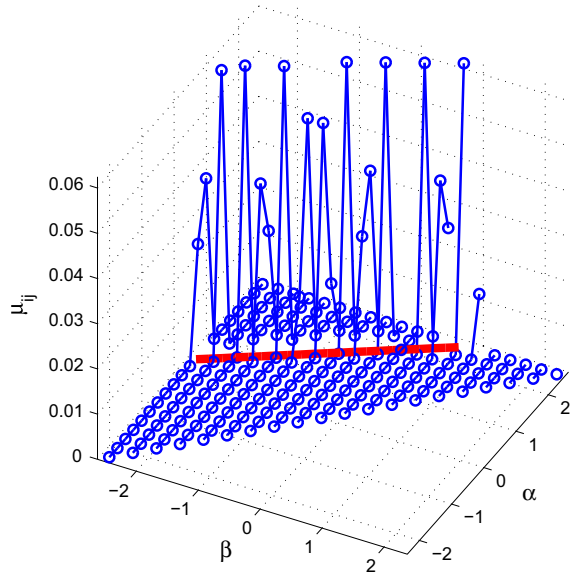

Fig. 6 Estimated weighting parameters μ_{ij} of the discretized Preisach model in Example 1

Example 1 This example is to support Proposition 1 that Presiach model can describe the backlash nonlinearity (5), and the weighting parameters are nonzero only at a specific line. Parameters of the actual backlash nonlinearity (5) are $m_l = 0.5$, $c_l = 0.8$, $m_r = 0.8$ and $c_r = 1.5$. To simplify the illustration, the structure parameter $L = 20$ is assumed to be known a priori; moreover, the linear subsystem is set to be unit, so that the inner signal $x(t)$ is the same as the output $y(t)$ and becomes measurable. The input $u(t)$, in the range $[-2.5, 2.5]$, is designed by following Lemma 1 with

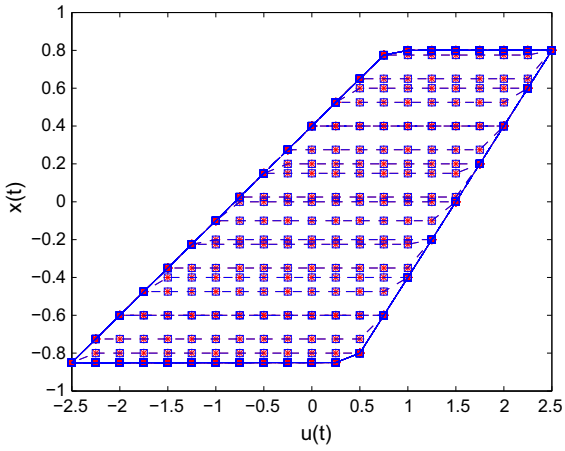

Fig. 7 The estimated backlash nonlinearity (*square*) and the actual one (*star*) in Example 1

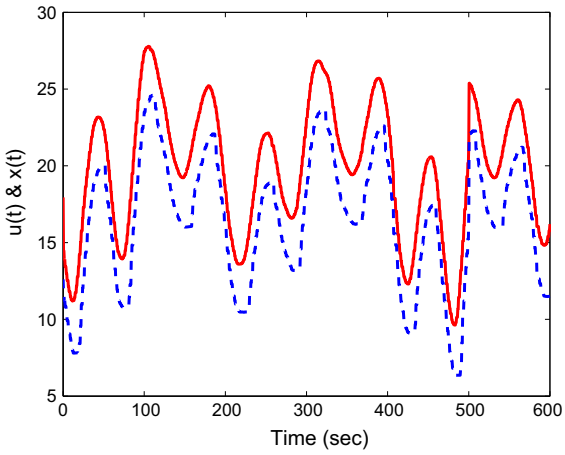

Fig. 8 The measured input $u(t)$ (*solid*) and output $x(t)$ (*dash*) of an industrial control valve in Example 2

$L = 20$. The samples of $u(t)$ and $x(t)$ are presented in Fig. 5. With the unit linear subsystem, the over-parameterized method reduces to the standard least squares method and is used to estimate the parameters of the discretized Preisach model. The estimated weighting parameters μ_{ij}'s in (14) are nonzero at the line $f_r(\alpha) = f_l(\beta)$ on the Preisach plane, shown as the thick solid line in Fig. 6. As shown in Fig. 7, the estimated Preisach model is exactly the same as the actual backlash nonlinearity. Hence, these results in Figs. 6 and 7 clearly support Proposition 1. □

Example 2 This example is to demonstrate the applicability of Preisach model in modeling hysteresis nonlinearities of industrial control valves. The input $u(t)$ and

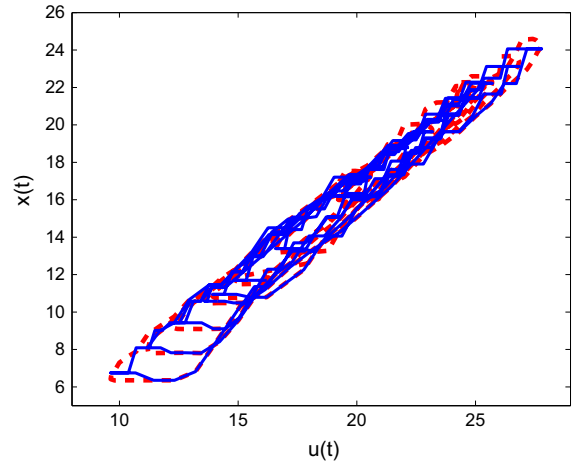

Fig. 9 The measured nonlinearity (*dash*) and the estimated Preisach model (*solid*) in Example 2

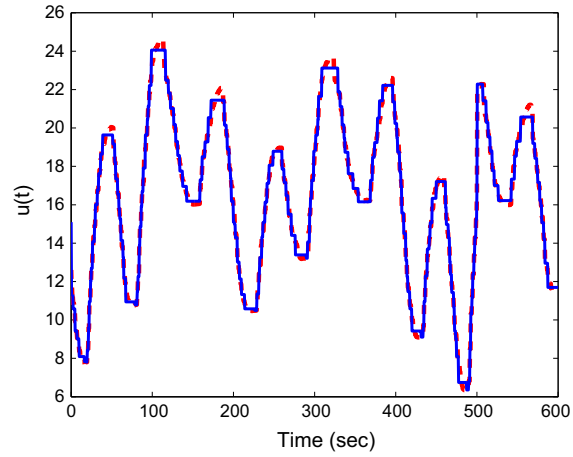

Fig. 10 The measured output $x(t)$ (*dash*) and the estimated output $\hat{x}(t)$ (*solid*) in Example 2

the output $x(t)$ of a control valve from a large-scale thermal power plant at Shandong Province in China are measured with sampling period 0.5 s and presented in Fig. 8. The actual nonlinearity is clearly revealed by plotting $x(t)$ against $u(t)$, shown in Fig. 9 (dash line). Preisach model is used to approximate such a hysteresis nonlinearity that is close to the general backlash in Fig. 1(b). The least squares method is used to estimate the discretized Preisach model, while, similar to Example 1, the linear subsystem is set to be unit. The estimated nonlinearity and its output are presented in Figs. 9 and 10, respectively. The results clearly sup-

Table 1 Actual and estimated parameters in Example 3

	Actual	Mean	SD
μ_{11}	2	2.0000	0.0407
μ_{12}	3	3.0034	0.0670
μ_{13}	4	4.0011	0.0590
μ_{14}	5	4.9968	0.0408
μ_{22}	2.5	2.4966	0.0494
μ_{23}	3.5	3.4971	0.0923
μ_{24}	4.5	4.5100	0.0617
μ_{33}	6	6.0025	0.0528
μ_{34}	6.5	6.4961	0.0883
μ_{44}	7	7.0004	0.0494
a_1	−0.8	−0.7999	0.0026
a_2	0.5	0.5000	0.0017
b_1	0.6	0.5999	0.0019
b_2	0.8	0.8001	0.0014

port that Preisach model is able to well approximate the hysteresis nonlinearity of this industrial control valve. □

Example 3 This example is to illustrate the PE condition in Proposition 3 and the consistency of estimated model parameters in Proposition 4. The actual input hysteresis nonlinearity is formulated by the discretized Preisach model in (14) as a superposition of 10 relay operators with the weighting parameters listed in Table 1. The discretized level L is equal to 4, and the discretized thresholds are $u_1 = -1$, $u_2 = 0$, $u_3 = 1$, $u_4 = 2$ and $u_5 = 3$. The actual input nonlinearity is shown in Fig. 11. The actual parameters of the linear subsystem are $a = [-0.8,\ 0.5]^T$ and $b = [0.6,\ 0.8]^T$. The structure parameters $n_a = 2$, $n_b = 2$, $n_k = 0$ and $L = 4$ are assumed to be known a priori.

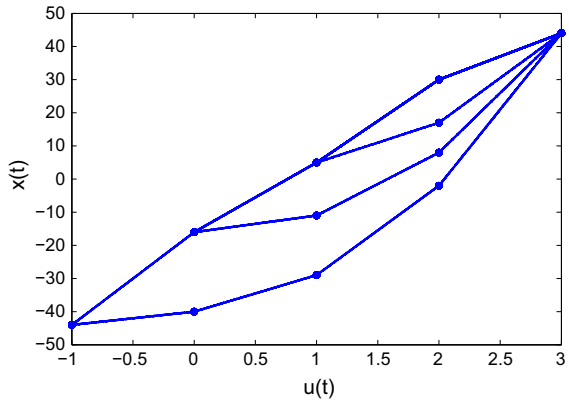

Fig. 11 The actual input nonlinearity in Example 3

The input $u(t)$ is designed to follow the conditions in Proposition 3. One hundred Monte Carlo experiments

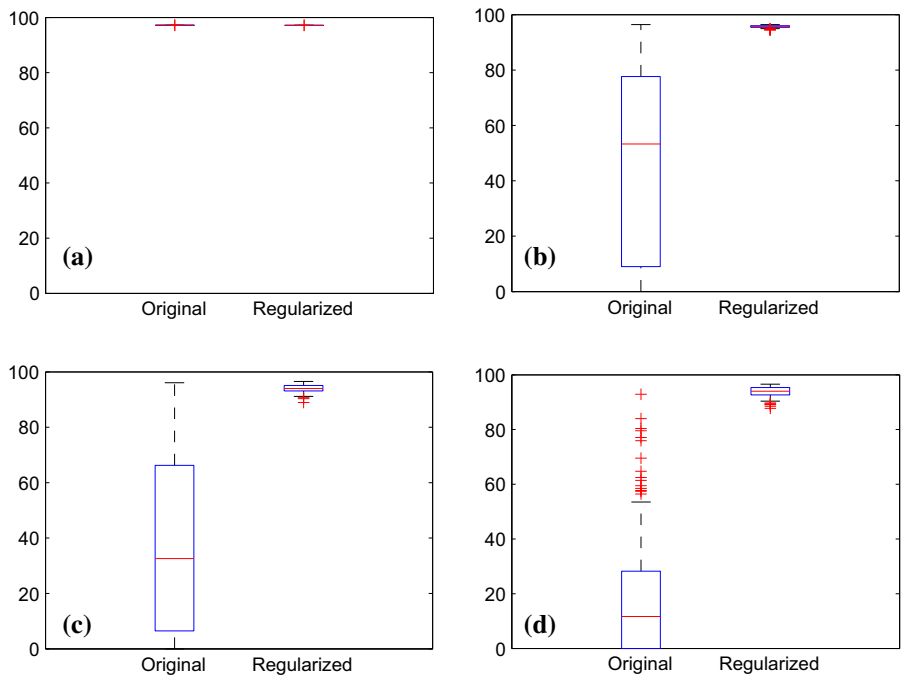

Fig. 12 Boxplots of fitnesses between $\hat{y}_s(t)$ and $y(t)$ for 100 Monte Carlo experiments using the over-parameterized method without (*left*) and with (*right*) regularization using different structure parameters $\hat{n}_a = kn_a$, $\hat{n}_b = kn_b$, $\hat{n}_k = 0$ and $\hat{L} = kL$: **a** $k = 1$, **b** $k = 2$, **c** $k = 3$, **d** $k = 4$

are performed with the same input $u(t)$ and different realizations of the noise $e(t)$, which is a Gaussian white noise with zero mean and unit variance. In each experiment, 2000 samples of $u(t)$ and $y(t)$ are collected to yield the estimated parameters $\hat{a}$, $\hat{b}$ and $\hat{\mu}$ in (26) from the over-parameterized method. The sample means and standard deviations of estimated parameters are listed in Table 1. The obtained results confirm that the estimated model parameters are consistent. □

Example 4 This example is to present the effectiveness of incorporating the regularization technique in the over-parameterized method. The configuration is the same as Example 3. The model quality is quantitatively measured by the fitness between $\hat{y}_s(t)$ and $y(t)$,

$$F_y = \max\left\{100\left(1 - \frac{\left\|\hat{y}_s(t) - y(t)\right\|_2}{\|y(t) - E\{y(t)\}\|_2}\right), 0\right\},$$

where $\hat{y}_s(t)$ is the simulated output,

$$\hat{y}_s(t) = \frac{\sum_{j=1}^{n_b} \hat{b}_j q^{-j}}{1 + \sum_{i=1}^{n_a} \hat{a}_i q^{-i}} \gamma^T(t - n_k)\,\hat{\mu}.$$

If the structure parameters n_a, n_b, n_k and L are known a prior, then the boxplots of the fitness between $\hat{y}_s(t)$ and $y(t)$ for 100 Monte Carlo experiments using the over-parameterized method and the one with regularization are illustrated in Fig. 12a. It can be seen that the output fitness for the over-parameterized method with regularization is close to that from the over-parameterized method, with a very small decline (about 0.02). The reason is that the over-parameterized method provides consistent estimated parameters but the regularized method introduces some bias with the usage of regularization matrix D.

If higher-order structure parameters are applied, e.g., $\hat{n}_a = kn_a$, $\hat{n}_b = kn_b = 4$, $\hat{n}_k = 0$ and $\hat{L} = kL$ for $k = 2, 3, 4$, then the boxplots of the resulting output fitnesses are shown in Fig. 12b–d. As expected, the performance of the over-parameterized method without regularization deteriorates severely with increasing orders, while the one with regularization achieves satisfactory output fitnesses. This comparison strongly supports the usage of regularization technique, particularly in the case with high-order structure parameters.

7 Conclusion

In this paper, Preisach model was used to represent the hysteresis-type input nonlinearities of extended Hammerstein systems. Proposition 1, Corollaries 1 and 3 proved that Preisach model is an effective model covering the backlash-type nonlinearities in Fig. 1a, b and the hysteresis-relay nonlinearities in Fig. 1c, d. The PE condition of input and the consistency of estimated model parameters were, respectively, established in Propositions 3 and 4. Owing to the adoption of Preisach model and the PE conditions established, the two motivational questions Q1 and Q2 proposed in Sect. 1 are positively answered. Several numerical and industrial examples supported the obtained results.

Acknowledgments 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.

Appendix 1: Representation theorem

This appendix recalls the Representation Theorem established by Mayergoyz [18] that gives a sufficient and necessary condition for the hysteresis-type nonlinearities that can be represented by Preisach model:

Theorem 1 *(Representation Theorem) [18] The wiping-out property and the congruency property constitute the necessary and sufficient condition for a rate-independent hysteresis nonlinearity to be represented by Preisach model on the set of piecewise continuous inputs.*

The three properties in Theorem 1 are defined as follows:

Property 1 *(Wiping-out property) [19] (pp. 15–17 therein) Only the series of dominant input extrema are stored by Preisach model, and all other input local extrema are wiped out.*

Property 2 *(Congruency property) [19] (pp. 18–19 therein) All hysteresis loops corresponding to the same extreme values of input are congruent in the geometric sense (i.e., the loops are identical up to translation).*

Property 3 *(Rate-independence property) [28] (p. 13 therein) The hysteresis nonlinearity output is invariant to any increasing time homeomorphism (i.e., the output is independent to the variation velocity of the input).*

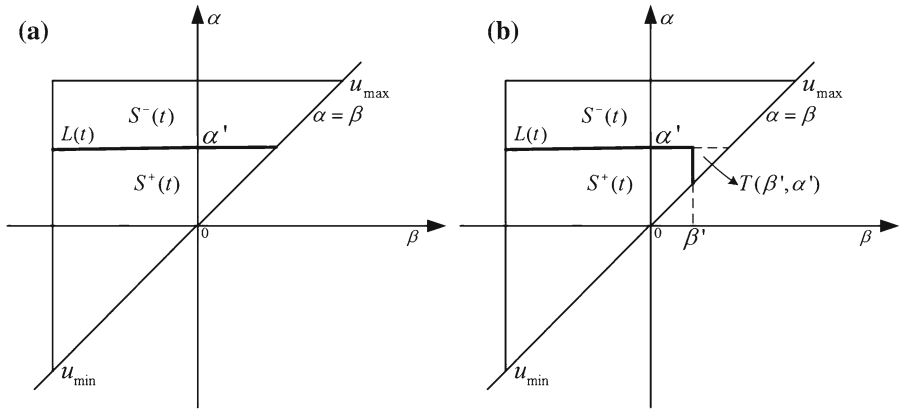

Fig. 13 The memory curve $L(t)$ and the sets $S^+(t)$ and $S^-(t)$: **a** the input $u(t)$ is first increased from $u_{\min}$ to α', and **b** u is decreased to β' afterward

Appendix 2: First-order reversal function

This appendix serves to recall the first-order reversal function defined in [19].

First, the geometric interpretation of Preisach model is a useful tool [18,19]. Define two sets $S^+(t)$, $S^-(t)$ and one memory curve $L(t)$ as follows:

- $S^+(t)$ consists of all the points (β, α) for which the relay operator $\gamma_{\beta\alpha}(t) = 1$ at the time instant t;
- $S^-(t)$ consists of all the points (β, α) for which $\gamma_{\beta\alpha}(t) = -1$ at the time instant t;
- $L(t)$ is the boundary line between $S^+(t)$ and $S^-(t)$.

With the sets $S^+(t)$ and $S^-(t)$, Preisach model in (3) can be equivalently represented as

$$f(t) = \iint_{S^+(t)} \mu(\beta, \alpha)\mathrm{d}\beta\mathrm{d}\alpha - \iint_{S^-(t)} \mu(\beta, \alpha)\mathrm{d}\beta\mathrm{d}\alpha. \quad (42)$$

Second, the first-order descending reversal function is formulated as the path of $f(t)$ for a monotonic increment of $u(t)$ starting from the minimum value $u_{\min}$ to the value α' followed by a subsequent monotonic decrement to β'; the corresponding $\alpha - \beta$ diagram is shown in Fig. 13. Denote $f_{\alpha'}$ and $f_{\alpha',\beta'}$ as the output values of $f(t)$ in (42) for $u(t) = \alpha'$ and $u(t) = \beta'$, respectively. The so-called Everett function is defined as [18],

$$F(\beta', \alpha') = \frac{1}{2}\left(f_{\alpha'} - f_{\alpha'\beta'}\right). \quad (43)$$

It is revealed from the $\alpha - \beta$ diagrams in Fig. 13 that the triangle $T(\beta', \alpha')$ with three vertices (β', β'), (α', α') and (β', α') is added to the negative set S^- and subtracted from the positive set S^+, when $u(t)$ is monotonically decreased from α' to β'. From (42) and (43), we have

$$F(\beta', \alpha') = \iint_{T(\beta',\alpha')} \mu(\beta, \alpha)\mathrm{d}\beta\mathrm{d}\alpha.$$

Thus, the weighting function $\mu(\beta', \alpha')$ can be determined by differentiating $F(\beta', \alpha')$ with respect to β' and α', i.e.,

$$\mu(\beta', \alpha') = \frac{\partial^2 F(\beta', \alpha')}{\partial\beta'\partial\alpha'}. \quad (44)$$

References

1. Bai, E.W.: An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems. Automatica **34**(3), 333–338 (1998)
2. Bai, E.W.: Identification of linear systems with hard input nonlinearities of known structure. Automatica **38**(5), 853–860 (2002)
3. Bai, E.W.: Discussion on: "Subspace-based identification algorithms for Hammerstein and Wiener models". Eur. J. Control **11**(2), 137–138 (2005)
4. Bai, E.W., Li, K.: Convergence of the iterative algorithm for a general Hammerstein system identification. Automatica **46**(11), 1891–1896 (2010)
5. Banks, H.T., Kurdila, A.J., Webb, G.: Identification of hysteretic control influence operators representing smart actuator part I: formulation. Math. Probl. Eng. **3**(4), 287–328 (1997)
6. Brokate, M., Sprekels, J.: Hysteresis and Phase Transitions. Springer, New York (1996)
7. Cerone, V., Regruto, D.: Bounding the parameters of linear systems with input backlash. IEEE Trans. Autom. Control **52**(3), 531–536 (2007)
8. Chan, K.H., Bao, J., Whiten, W.J.: Identification of MIMO Hammerstein systems using cardinal spline functions. J. Process Control **16**(7), 659–670 (2006)
9. Chang, F., Luus, R.: A noniterative method for identification using Hammerstein model. IEEE Trans. Autom. Control **16**(5), 464–468 (1971)

10. Chen, T., Ohlsson, H., Ljung, L.: On the estimation of transfer functions, regularizations and Gaussian processes—revisited. Automatica **48**(8), 1525–1535 (2012)
11. Cheng, Y.C., Yu, C.C.: Relay feedback identification for actuators with hysteresis. Ind. Eng. Chem. Res. **39**(11), 4239–4249 (2000)
12. Giri, F., Bai, E.W. (eds.): Block-Oriented Nonlinear System Identification. Springer, Berlin (2010)
13. Giri, F., Rochdi, Y., Brouri, A., Chaoui, F.Z.: Parameter identification of Hammerstein systems containing backlash operators with arbitrary-shape parametric borders. Automatica **47**(8), 1827–1833 (2011)
14. Giri, F., Rochdi, Y., Chaoui, F.Z., Brouri, A.: Identification of Hammerstein systems in presence of hysteresis-backlash and hysteresis-relay nonlinearities. Automatica **44**(3), 767–775 (2008)
15. Haber, R., Keviczky, L.: Nonlinear System Identification: Input–Output Modeling Approach. Kluwer Academic Publishers, Dordrecht (1999)
16. Iyer, R.V., Shirley, M.E.: Hysteresis parameter identification with limited experimental data. IEEE Trans. Magn. **40**(5), 3227–3239 (2004)
17. Ljung, L.: System Identification: Theory for the User, 2nd edn. Prentice Hall, Englewood Cliffs (1999)
18. Mayergoyz, I.D.: Mathematical models of hysteresis. Phys. Rev. Lett. **56**(15), 1518–1521 (1986)
19. Mayergoyz, I.D.: Mathematical Models of Hysteresis and Their Applications, 2nd edn. Academic Press, London (2003)
20. Miyashita, N., Yamakita, M.: Identification of Hammerstein systems with piecewise nonlinearities with memory. In: Proceedings of 46th IEEE Conference on Decision and Control, pp. 5749–5754. New Orleans, LA, USA, Dec 12–14 (2007)
21. Pillonetto, G., Dinuzzo, F., Chen, T., Nicolao, G.D., Ljung, L.: Kernel methods in system identification, machine learning and function estimation: a survey. Automatica **50**(3), 657–682 (2014)
22. Rochdi, Y., Giri, F., Gning, J.B., Chaoui, F.Z.: Identification of block-oriented systems in the presence of nonparametric input nonlinearities of switch and backlash types. Automatica **46**(5), 864–877 (2010)
23. Shirley, M.E., Venkataraman, R.: On the identification of Preisach measures. Smart Struct. Mater. **5049**, 326–336 (2003)
24. Söderström, T., Stoica, P.: System Identification. Prentice Hall, Englewood Cliffs (1989)
25. Stoica, P., Söderström, T.: Instrumental-variable methods for identification of Hammerstein systems. Int. J. Control **35**(3), 459–476 (1982)
26. Tan, X., Baras, J.S.: Adaptive identification and control of hysteresis in smart materials. IEEE Trans. Autom. Control **50**(6), 827–839 (2005)
27. Tjärnström, F., Ljung, L.: Variance properties of a two-step ARX estimation procedure. Eur. J. Control **9**(4), 422–430 (2003)
28. Visintin, A.: Differential Models of Hysteresis. Springer, New York (1994)
29. Vörös, J.: Modeling and identification of systems with backlash. Automatica **46**(2), 369–374 (2010)
30. Vörös, J.: Recursive identification of nonlinear cascade systems with time-varying general input backlash. J. Dyn. Syst. Meas. Control 135, 014,504-1 (2013)
31. Wang, J., Sano, A., Chen, T., Huang, B.: Identification of Hammerstein systems without explicit parameterisation of non-linearity. Int. J. Control **82**(5), 937–952 (2009)
32. Wang, J., Zhang, Q., Ljung, L.: Revisiting Hammerstein system identification through the two-stage algorithm for bilinear parameter estimation. Automatica **45**(11), 2627–2633 (2009)
33. Zhu, Y.: Multivariable System Identification for Process Control. Elsevier Science, Oxford (2001)