

Identification of Continuous-time Hammerstein Systems by Simultaneous Perturbation Stochastic Approximation

Mohd Ashraf Ahmad^{a,b}, Shun-ichi Azuma^b, Toshiharu Sugie^b

^aFaculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, 26600, Pekan, Pahang, Malaysia

^bDepartment of Systems Science, Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Kyoto 606-8501, Japan

Abstract

This paper proposes an identification method for Hammerstein systems using simultaneous perturbation stochastic approximation (SPSA). Here, the structure of nonlinear subsystem is assumed to be unknown, while the structure of linear subsystem, such as the system order, is assumed to be available. The main advantage of the SPSA-based method is that it can be applied to identification of Hammerstein systems with less restrictive assumptions. In order to clarify this point, piecewise affine functions with a large number of parameters are adopted to approximate the unknown nonlinear subsystems. Furthermore, the linear subsystems are supposed to be described in continuous-time. Though this class of systems closely reflects the actual systems, there are few methods to identify such models. Hence, the SPSA-based method is utilized to identify the parameters in both linear and nonlinear subsystems simultaneously. The effectiveness of the proposed method is evaluated through several numerical examples. The results demonstrate that the proposed algorithm is useful to obtain accurate models, even for high-dimensional parameter identification.

Keywords: systems identification, stochastic approximation, continuous-time Hammerstein models.

1. Introduction

The modelling of real-world plants and processes, which are nonlinear in nature, remains a challenging problem. Both an expert and intelligent systems are therefore required to model accurately such plants and processes. One way to cope with this difficulty is to introduce identification of block oriented models. These models include a Hammerstein model (a static nonlinear subsystem followed by a linear subsystem), a Wiener model (a linear subsystem followed by a static nonlinear subsystem), or a Hammerstein-Wiener model (a linear subsystem sandwiched by two static nonlinear subsystems or vice-versa). In particular, an intelligent system, such as a system with a neural feed-forward controller, is modeled as a Hammerstein model. These models have been adopted by many researchers partly because they closely reflect actual nonlinear systems with relatively simple structures. As a result, these models have been successfully used to describe many practical plants, such as fuel cells (Li et al., 2008), valve actuators (Wang & Zhang, 2014), wind turbines (van der

Veen et al., 2013), spark ignition engine torques (Togun et al., 2012), and stirred tank reactors (Shi et al., 2011).

Among various types of nonlinear system models, the Hammerstein model is quite popular. In fact, the identification of Hammerstein systems has been widely reported in the literature (Pawlak, 1991; Ding & Chen, 2005; Zhao, 2006; Ding et al., 2006; Bai & Li, 2004; Liu & Bai, 2007; Greblicki, 2000; Ding et al., 2007b, 2011; Hasiewicz & Mzyk, 2004; Bai, 2004). Over the past two decades, various methods for identification of Hammerstein systems have been studied extensively. These can be roughly classified into several categories, such as the iterative method (Stoica, 1981; Narendra & Gallman, 1966; Rangan et al., 1995; Voros, 1997; Liu & Bai, 2007), the over-parameterization method (Chang & Luus, 1971; Hsia, 1976; Ding et al., 2007a), the blind approach (Bai & Fu, 2002), the subspace method (Verhaegen & Westwick, 1996), the least squares method (Ding & Chen, 2005; Goethals et al., 2005), the parametric instrumental variables method (Stoica & Soderstrom, 1981; Laurain et al., 2009), the stochastic method (Bilings & Fakhouri, 1978; Pawlak, 1991; Greblicki, 1996) and the non-parametric identification method (Greblicki & Pawlak, 1987; Krzyak, 1993; Bai, 2003;

Email addresses: mashraf@ump.edu.my (Mohd Ashraf Ahmad), sazuma@i.kyoto-u.ac.jp (Shun-ichi Azuma), sugie@i.kyoto-u.ac.jp (Toshiharu Sugie)

Krzyak, 1996).

Recently, a decomposition-based Newton iterative identification approach for a Hammerstein nonlinear FIR system with ARMA noise was presented by (Ding et al., 2014). Here, it was claimed that a fast convergence rates with more accurate parameter estimation can be achieved by using the Newton iterative method. In (Wang et al., 2015), a hierarchical least squares method has been proposed for the identification of Hammerstein state space model. A similar approach is applied to Hammerstein nonlinear controlled autoregressive models (Chen & Ding, 2015). Both results show that the hierarchical identification principle may improve the computational efficiency by decomposing one nonlinear system into several subsystems with smaller dimensions and fewer variables. In (Ma & Liu, 2015), a nonlinear recursive instrumental variables (RIV) identification method for Hammerstein ARMAX system is adopted. The effectiveness of the RIV method is shown in terms of identification accuracy and convergence speed, especially under colored noise. Meanwhile, a blind approach with new over-sampling strategy was adopted in (Yu et al., 2014) to produce a consistent parameter estimation in the presence of noise.

In spite of such abundant literature, several restrictions are inevitable in their studies from the theoretical view-point.

- (i) Most of the identification methods are restricted to the models in discrete-time, while it is natural to express a real system in a continuous-time domain.
- (ii) Many approaches assume that the static nonlinear system is given by a linear combination of several basis functions.
- (iii) In the over-parameterization identification method, the identification model contains the products of nonlinear and linear parameters, causing redundant parameter identification and a large computational load.

Though it looks that that we can handle general class of nonlinear subsystems by adopting so many basis functions (such as higher order polynomials and piecewise affine functions), this is not tractable in reality in the existing identification frame work.

On the other hand, there are different types of identification methods which utilize evolutionary computation, such as the cuckoo search algorithm (Gotmare et al., 2015), the stochastic gradient (Mao & Ding, 2015; Chen & Wang, 2015), and the PSO (particle swarm optimization) (Nanda et al., 2010; Ko, 2011; Wang et al., 2014;

Jingzhuo et al., 2014). These methods are quite flexible in nature, and do not suffer from (i) and (iii) mentioned above. In particular, the PSO is known to be effective in various systems control supplication (Maruta et al., 2009, 2013). However, they have a serious draw back.

- (iv) In swarm based optimization (including PSO), the computation times per iteration are proportional to the number of swarms. As a result, these methods require heavy computation time in the identification process, especially for static nonlinear systems with a large number of basis functions.

Hence, it is not tractable to handle static nonlinear subsystems consisting of large number of basis. Consequently, these evolutionary computation based methods cannot avoid (ii) as well.

To the best of the authors' knowledge, the simultaneous perturbation stochastic approximation (SPSA) method (Spall, 1992) could be only candidate to provide us with a promising tool for such system identification problems. This is because the SPSA method is well known to be effective for a variety of optimization problems, even for high-dimensional parameter tuning (Ahmad et al., 2014b). In comparison to the existing results, the SPSA method does not suffer from the afore-mentioned theoretical restrictions (i)-(iv). It may be expected to identify both linear and nonlinear subsystems simultaneously, even for large number of basis functions with less computational load. Meanwhile, one major drawback of SPSA may be to guarantee the local convergence only from the theoretical points of view.

Based on the above observations, this paper thus presents an identification method of Hammerstein systems in continuous-time using simultaneous perturbation stochastic approximation. We assume that the structure (i.e., the system order) of the linear subsystem is known in advance, while the structure of the nonlinear subsystem remains unknown. Here, a piecewise affine function is then used to approximate the unknown nonlinear function in the Hammerstein model. Next, based on the input and output data, the SPSA-based method is used to identify the coefficients of the linear time-invariant (LTI) model and the piecewise affine function simultaneously. In order to clarify the benefit of the SPSA-based method, a large number of parameters in the piecewise affine function are considered here. So far, there have been few papers discussing the identification of such Hammerstein models. Therefore, it is worth evaluating the effectiveness of the SPSA method.

The remainder of this paper is organized as follows. Section 2 formulates the identification problem

for Hammerstein models. In Section 3, the identification method using simultaneous perturbation stochastic approximation-based algorithm is presented. The nonlinear function identification based on piecewise affine function is also described in the same section. The method is validated through several numerical examples in Section 4. Finally, some concluding remarks are given in Section 5. This paper is based on **our preliminary** version (Ahmad et al., 2014a), published in a conference proceedings, and contains **the** full explanations and experiments omitted there.

Notation: The symbols $\mathbb{R}$ and $\mathbb{R}_+$ represent the set of real numbers and the set of positive real numbers, respectively. For the vector θ , we use $\|\theta\|_2$ to express the standard Euclidean norm. For $\delta \in \mathbb{R}_+$, $\text{sat}_\delta : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denotes the saturation function whose i -th element given as follows:

$$\text{The } i\text{-th element of } \text{sat}_\delta(\mathbf{x}) = \begin{cases} \delta & \text{if } \delta < x_i, \\ x_i & \text{if } -\delta \leq x_i \leq \delta, \\ -\delta & \text{if } x_i < -\delta \end{cases}$$

where $x_i \in \mathbb{R}$ is the i -th element of $\mathbf{x} \in \mathbb{R}^n$.

2. Problem Formulation

Consider the continuous-time single-input-single-output (SISO) Hammerstein model in Fig. 1, composed of a nonlinear function f and a linear dynamical system G described by the differential operator p ($:= \frac{d}{dt}$):

$$G(p) = \frac{B(p)}{A(p)} = \frac{b_m p^m + b_{m-1} p^{m-1} + \dots + b_0}{p^n + a_{n-1} p^{n-1} + \dots + a_0}. \quad (1)$$

Here, $u(t)$ is the input, $\bar{u}(t)$ is the unmeasurable output of the nonlinear function, namely $\bar{u}(t) = f(u(t))$, $\tilde{y}(t)$ is the measurement of $y(t)$ but is corrupted by the noise $v(t)$. The input-output relationship is described as follows:

$$\tilde{y}(t) = G(p)f(u(t)) + v(t). \quad (2)$$

In this paper, we address an identification problem of the Hammerstein model. Here, we assume that:

- m and n are known.
- a_i ($i = 0, 1, \dots, n-1$) and b_i ($i = 0, 1, \dots, m$) are positive real numbers.
- The function f is unknown, but $f(u(t))$ is a one-to-one map to $u(t)$. Moreover, $f(0) = 0$.
- $G(p)$ is stable and **minimum** phase.

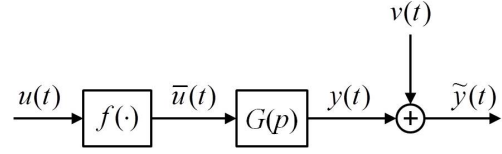


Figure 1: The continuous-time SISO Hammerstein model

- $b_m = 1$ so that $f(u(t))$ and $G(p)$ can be determined uniquely.

Remark 2.1. Note that any pair of $\beta f(u(t))$ and $G(p)/\beta$ for $\beta \neq 0$ would produce the identical input-output measurements, which make the parameterization of the Hammerstein model in Fig. 1 **non-unique** (Bai & Fu, 2002). For this reason, we normalize $G(p)$ by assuming $b_m = 1$.

In order to evaluate the identified model, we introduce the objective function

$$J(\hat{G}, \hat{f}) = \sum_{\kappa=0}^N (\tilde{y}(\kappa t_s) - \hat{y}(\kappa t_s))^2 \quad (3)$$

where t_s is the sampling **time** for the data set $(u(t), \tilde{y}(t))$ ($t = 0, t_s, 2t_s, \dots, Nt_s$), $\hat{G}$ and $\hat{f}$ are the identified models of G and f , and $\hat{y}(t) = \hat{G}(p)\hat{f}(u(t))$.

Then, the identification problem can be described as follows.

Problem 2.1. For the Hammerstein model in Fig. 1, suppose that the input-output data $(u(t), \tilde{y}(t))$ ($t = 0, t_s, 2t_s, \dots, Nt_s$) are given. Then, find $\hat{G}$ and $\hat{f}$ which minimizes $J(\hat{G}, \hat{f})$. $\square$

3. Identification Method Using Simultaneous Perturbation Stochastic Approximation

This section **proposes** a method to solve Problem 2.1. First, the standard SPSA algorithm (Spall, 1992) is briefly explained. **An** identification technique of the nonlinear function, based on piecewise affine function, is **then** described. Finally, the identification method is presented based on the SPSA-based algorithm.

3.1. Review of Simultaneous Perturbation Stochastic Approximation

Consider the optimization problem

$$\min_{\mathbf{x} \in \mathbb{R}^n} h(\mathbf{x}) \quad (4)$$

where $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function and $\mathbf{x} \in \mathbb{R}^n$ denotes the design variable.

The SPSA algorithm iteratively updates the design variable $\mathbf{x}$ to search a local optimal solution $\mathbf{x}^* \in \mathbb{R}^n$ of (4). The update law is given by

$$\mathbf{x}(k+1) = \mathbf{x}(k) - a(k)\mathbf{g}(\mathbf{x}(k), \Delta(k)) \quad (5)$$

for $k = 0, 1, \dots$, where $a(k) \in \mathbb{R}$ is the gain, $\Delta(k) \in \mathbb{R}^n$ is the perturbation vector, which is randomly generated, and $\mathbf{g}(\mathbf{x}(k), \Delta(k))$ is the update vector given by

$$\mathbf{g}(\mathbf{x}(k), \Delta(k)) = \begin{bmatrix} \frac{h(\mathbf{x}(k)+c(k)\Delta(k)) - h(\mathbf{x}(k)-c(k)\Delta(k))}{2c(k)\Delta_1(k)} \\ \frac{h(\mathbf{x}(k)+c(k)\Delta(k)) - h(\mathbf{x}(k)-c(k)\Delta(k))}{2c(k)\Delta_2(k)} \\ \vdots \\ \frac{h(\mathbf{x}(k)+c(k)\Delta(k)) - h(\mathbf{x}(k)-c(k)\Delta(k))}{2c(k)\Delta_n(k)} \end{bmatrix}. \quad (6)$$

In (6), $c(k) \in \mathbb{R}$ is another gain and $\Delta_i(k) \in \mathbb{R}$ is the i -th elements of $\Delta(k) \in \mathbb{R}^n$.

The idea of this algorithm is that the expectation of $\mathbf{g}(\mathbf{x}(k), \Delta(k))$ is nearly equal to the gradient of the objective function h , i.e., $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}(k))$, and thus (5) corresponds to a kind of stochastic steepest descent. The essential feature is that the gradient approximation is given by the only two measurements of the objective function, which are independent of the dimension of the design variable. Hence it does not require any explicit form of the objective function h , and can be a promising tool for solving large-scale optimization problems. Moreover, the measurement of the objective function h could include the noise term. A convergence condition of this algorithm (even if the measurement data is contaminated by noise) and a guidance to choose $a(k)$, $c(k)$ and $\Delta(k)$ are reported in (Spall, 1992).

3.2. Piecewise Affine Approximation of Nonlinear Functions

Since the nonlinear function $\hat{f}$ has to be parameterized in the identification process, we adopt a piecewise affine function to approximate the unknown nonlinear function f .

The piecewise affine function consists of a number of line segments that are connected to each other as shown in Fig. 2 and can be written as

$$\hat{f}(u) = \begin{cases} \eta_0 + M_1(u - w_0) & \text{if } w_0 \leq u < w_1, \\ \eta_1 + M_2(u - w_1) & \text{if } w_1 \leq u < w_2, \\ \vdots & \\ \eta_{\ell-1} + M_\ell(u - w_{\ell-1}) & \text{if } w_{\ell-1} \leq u < w_\ell, \end{cases} \quad (7)$$

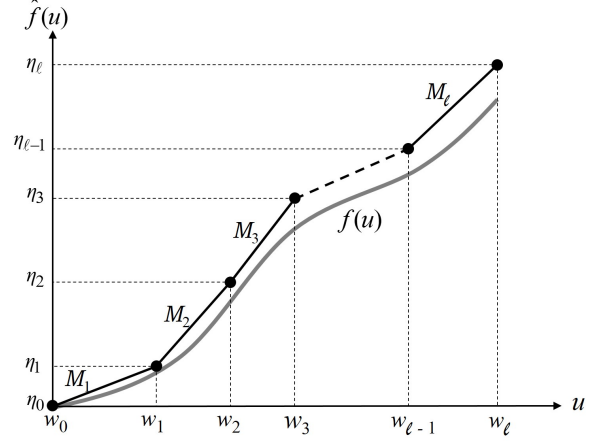


Figure 2: Piecewise affine approximation

where $M_i = (\eta_i - \eta_{i-1})/(w_i - w_{i-1})$ ($i = 1, 2, \dots, \ell$) are the segment slopes, w_i ($i = 0, 1, \dots, \ell$) $\in \mathbb{R}$ are the given input points satisfying $w_0 < w_1 < \dots < w_\ell$ and connecting the line segments, and η_i ($i = 0, 1, \dots, \ell$) $\in \mathbb{R}$ are the output points corresponding to each input point. For simplicity of the notation, let $\mathbf{w} := (w_0, w_1, \dots, w_\ell)^\top \in \mathbb{R}^{\ell+1}$. Furthermore, since $f(u)$ is a one-to-one map to u and $f(0) = 0$ by assumption, it is reasonable to consider

$$\eta_i = \begin{cases} \tilde{\eta}_i & \text{if } w_i > 0, \\ 0 & \text{if } w_i = 0, \\ -\tilde{\eta}_i & \text{if } w_i < 0, \end{cases} \quad (8)$$

for $i = 0, 1, \dots, \ell$, where $\tilde{\eta}_i \in \mathbb{R}_+$. By choosing ℓ large, the accuracy of the approximation increases. However, such choice results in a high-dimensional parameter identification problem, which is difficult to solve in practice.

3.3. Identification Method

This section presents how to apply the SPSA algorithm in Section 3.1 for the identification problem.

By using the piecewise affine approximation of $\hat{f}$, Problem 2.1 is reduced to the optimization problem with the objective function

$$h(\theta) = \sum_{\kappa=0}^N (\tilde{y}(\kappa t_s) - \hat{y}(\kappa t_s))^2 \quad (9)$$

for the design variable

$$\boldsymbol{\theta} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_{m-1} \\ \hat{a}_0 \\ \hat{a}_1 \\ \vdots \\ \hat{a}_{n-1} \\ \hat{\eta}_0 \\ \vdots \\ \hat{\eta}_\ell \end{bmatrix} \in \mathbb{R}^\rho \quad (10)$$

where $\rho = m + n + \ell + 1$. Solving the optimization problem in (9) by the SPSA algorithm in (5), we obtain a solution to Problem 2.1. On the other hand, in order to accelerate the design variable searching, we employ the logarithmic scale to the design variable $\boldsymbol{\theta}$ by setting $\theta_i = 10^{x_i}$ ($i = 1, 2, \dots, \rho$) with the objective function $h([10^{x_1} \ 10^{x_2} \ \dots \ 10^{x_\rho}]^\top)$. Then, our design procedure is summarized as follows:

Step 1: Determine the number k_{max} of the maximum iterations of the SPSA algorithm in (5). Determine the number ℓ and the location w_i ($i = 0, 1, \dots, \ell$). Let $x_i = \log \theta_i$ ($i = 1, 2, \dots, \rho$) and select the initial value $\mathbf{x}(0)$.

Step 2: Execute the SPSA algorithm in Section 3.1 for the objective function h in (9).

Step 3: After k_{max} iterations of the SPSA algorithm, we obtain $\mathbf{x}^* := \mathbf{x}(k_{max})$. Output $\boldsymbol{\theta}^* := [10^{x_1^*} \ 10^{x_2^*} \ \dots \ 10^{x_\rho^*}]^\top$ is a solution to Problem 2.1.

In Step 2, it is required to compute the value of $h(\boldsymbol{\theta})$ for a fixed $\boldsymbol{\theta}$. This is reduced to computing $\hat{y}(\kappa t_s)$ and it is performed as follows. First, assume that $u(t)$ is a continuous-time signal given by the zero-order hold for $u(t)$ ($t = 0, t_s, 2t_s, \dots, Nt_s$). Then, we can calculate

$$\hat{y}(t) = \frac{p^m + \hat{b}_{m-1}p^{m-1} + \dots + \hat{b}_0}{p^n + \hat{a}_{n-1}p^{n-1} + \dots + \hat{a}_0} \hat{f}(u(t)) \quad (11)$$

in the continuous-time environment. Then, $\hat{y}(t)$ is sampled to $\hat{y}(\kappa t_s)$ at a constant sampling time t_s for $\kappa = 0, 1, \dots, N$.

Remark 3.1. The standard SPSA algorithm in (5) does not always give a stable solution during the optimization process. This is due to a possibility that the updated design variables **grow** with k and **yield** an undesirable solution. In order to avoid this problem, we adopt a

modified SPSA algorithm, which has been proposed in (Baba et al., 2013). There, a saturation function $\text{sat}_\delta(\cdot)$ has been introduced in (5). That is,

$$\mathbf{x}(k+1) = \mathbf{x}(k) - \text{sat}_\delta(a(k)\mathbf{g}(\mathbf{x}(k), \Delta(k))). \quad (12)$$

4. Numerical Examples

In this section, the effectiveness of the proposed identification method is demonstrated with several numerical examples.

4.1. Example 1

Consider the system:

$$G(p) = \frac{B(p)}{A(p)}, \quad (13)$$

$$\begin{aligned} A(p) &= p^6 + 10.0000p^5 + 54.7700p^4 + 156.8000p^3 \\ &\quad + 87.0843p^2 + 25.2810p + 4.0197, B(p) = p^3, \\ f(u(t)) &= 125(u(t) + 0.5u^2(t) + 0.25u^3(t)). \end{aligned} \quad (14)$$

In this example, a Pseudo Random Binary Sequence (PRBS) signal with varying amplitude is considered as the input $u(t)$, while $v(t)$ is a white noise with zero mean and variance $\sigma_v^2 = 0.01$. Here, the amplitude of PRBS signal is varied in a range of $[-1, 1]$. The output data $\tilde{y}(t)$ is sampled at $t_s = 0.001$ over $N = 24000$.

Since the amplitude of PRBS signal is bounded in the range of $[-1, 1]$, we can divide the input points w_i of the piecewise affine function equally in the range of $[-1, 1]$ with $\ell = 10$. Then, the locations of the input points are given by $\mathbf{w} := (-1, -0.8, -0.6, -0.4, -0.2, 0.01, 0.2, 0.4, 0.6, 0.8, 1)^\top \in \mathbb{R}^{11}$. The design variables $\boldsymbol{\theta} \in \mathbb{R}^{17}$ and its corresponding linear and nonlinear coefficients are **presented** in Table 1. Next, we use the modified SPSA algorithm in (12) and set the parameters of the SPSA-based algorithm $a(k) = 15/(k+1)^{0.9}$, $c(k) = 0.01/(k+1)^{1/6}$, $\delta = 0.01$ and $k_{max} = 5000$. The initial values $\mathbf{x}(0)$ are tabulated in Table 1. In order to observe the randomization effect, we perform 25 independent trials.

Table 2 tabulates the best identified parameter values from the 25 trials and its parameter identification error ξ . In this table, $\bar{\boldsymbol{\theta}} \in \mathbb{R}^\rho$ is a vector, whose elements consists of true identification parameters, and ξ is given by

$$\xi = \left\| \left[\frac{\theta_1 - \bar{\theta}_1}{\bar{\theta}_1}, \dots, \frac{\theta_\rho - \bar{\theta}_\rho}{\bar{\theta}_\rho} \right]^\top \right\|_2 \quad (15)$$

where $\bar{\theta}_i$ is the i -th element of the vector $\bar{\boldsymbol{\theta}}$. It shows that the resultant identified parameters are close to the

Table 1: Identification parameters of Example 1

θ	Coefficients	$x(0)$	θ corresponding to $x(0)$
θ_1	$\hat{a}_0$	2.0000	100.0000
θ_2	$\hat{a}_1$	2.0000	100.0000
θ_3	$\hat{a}_2$	2.0000	100.0000
θ_4	$\hat{a}_3$	2.0000	100.0000
θ_5	$\hat{a}_4$	2.0000	100.0000
θ_6	$\hat{a}_5$	2.0000	100.0000
θ_7	$\tilde{\eta}_0$	2.0000	100.0000
θ_8	$\tilde{\eta}_1$	1.9031	80.0000
θ_9	$\tilde{\eta}_2$	1.7782	60.0000
θ_{10}	$\tilde{\eta}_3$	1.6021	40.0000
θ_{11}	$\tilde{\eta}_4$	1.3010	20.0000
θ_{12}	$\tilde{\eta}_5$	-2.0000	0.0100
θ_{13}	$\tilde{\eta}_6$	1.3010	20.0000
θ_{14}	$\tilde{\eta}_7$	1.6021	40.0000
θ_{15}	$\tilde{\eta}_8$	1.7782	60.0000
θ_{16}	$\tilde{\eta}_9$	1.9031	80.0000
θ_{17}	$\tilde{\eta}_{10}$	2.0000	100.0000

true identification parameters and parameter identification error is becoming smaller as k increases. Meanwhile, in order to **see the results more closely**, the graph of the nonlinear function f and the bode diagram of linear system G are shown in Figs. 3 and 4, respectively. In both Figs. 3 and 4, the thin **gray-dash lines represent** the identified responses from the 25 trials, while the thick solid-black line represents the true response. It shows that the SPSA-based method can, **with high probability, accurately** identify both linear and nonlinear subsystems in the Hammerstein model.

Furthermore, the statistical analysis of the objective function and the parameter identification error with different noise variances, which are 0.01, 0.25, and 1.0, are tabulated in Table 3. In particular, for each level of white noise variance, **their** mean, best, worst and standard deviation values are observed from the 25 trials. Note that a high noise level results in a higher value of the objective function, which can be clearly seen from the values of mean, best, worst and standard deviation. However, in terms of the parameter identification error, the increment of the mean, best, worst and standard deviation values are relatively small. Hence, it shows that the SPSA-based algorithm still can produce smaller parameter identification error even for high noise level.

4.2. Example 2

In Example 2, consider a fourth order linear subsystem with complex poles and a nonlinear subsystem with

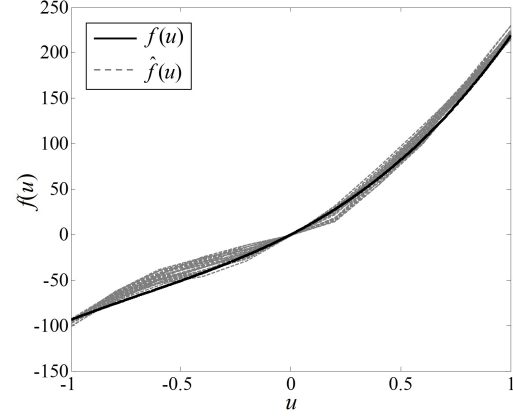


Figure 3: Identified nonlinear functions $\hat{f}(u)$ with the true nonlinear function $f(u)$ in Example 1

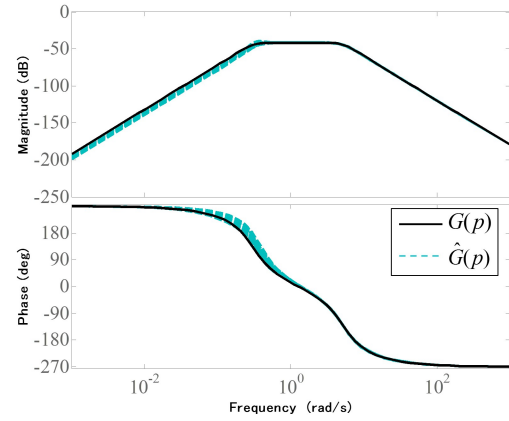


Figure 4: Identified linear systems $\hat{G}(p)$ with the true linear system $G(p)$ in Example 1

tangent hyperbolic function **given by**

$$G(p) = \frac{B(p)}{A(p)}, \quad (16)$$

$$A(p) = p^4 + 5.0000p^3 + 408.0000p^2 + 416.0000p + 1600.0000, B(p) = p + 0.2500,$$

$$f(u(t)) = 250 \tanh(1.5u(t)). \quad (17)$$

Here, we consider the same $u(t)$, $v(t)$, t_s and N as in Example 1.

Moreover, w for the piecewise affine function is set to be similar as in Example 1. The design variables $\theta \in \mathbb{R}^{16}$ and its corresponding linear and nonlinear coefficients are tabulated in Table 4. Next, we utilize the

Table 2: The best identified parameters of Example 1

k	100	500	1000	1500	2000	3000	4000	5000	$\bar{\theta}$
θ_1	61.6595	11.0398	6.3880	5.1552	4.6136	4.5745	4.4521	4.4662	4.0197
θ_2	64.5654	46.6753	46.1803	40.5091	38.0603	34.3867	32.3358	30.4906	25.2810
θ_3	147.9108	191.6781	123.9439	107.0407	103.9615	96.9761	93.9654	91.3663	87.0843
θ_4	162.1810	182.4933	162.2306	156.4483	158.4630	156.3911	155.0739	155.9906	156.8000
θ_5	117.4898	66.7712	56.5164	55.3787	55.6544	55.0488	54.7945	54.4622	54.7700
θ_6	85.1138	11.4642	10.1681	9.7844	9.9173	9.8485	9.7886	9.8931	10.0000
θ_7	85.1138	140.1687	103.8806	98.9751	97.4951	95.0511	94.0789	94.3057	93.7500
θ_8	98.4215	60.4572	56.4219	59.7494	61.4385	66.7463	68.0568	70.8570	76.0000
θ_9	61.3976	45.9962	42.8683	45.3468	49.4876	49.7910	51.1490	52.6335	59.2500
θ_{10}	35.6500	53.2509	42.3143	41.5301	41.6026	40.5197	40.8274	41.1615	42.0000
θ_{11}	21.4304	30.6678	42.9538	36.8894	31.3125	28.0662	27.0867	26.0172	22.7500
θ_{12}	0.0170	0.0204	0.0145	0.0140	0.0138	0.0133	0.0143	0.0147	1.2563
θ_{13}	16.2566	23.8840	18.1743	20.8733	21.0038	22.6007	23.7102	24.3970	27.7500
θ_{14}	34.0455	23.4115	32.7060	42.3298	48.1361	53.5851	55.7866	56.6143	62.0000
θ_{15}	58.6342	69.5208	91.8034	94.9460	100.7211	100.7297	101.6655	100.9917	104.2500
θ_{16}	78.1790	180.0901	166.5565	160.8019	164.4206	162.3344	159.2057	158.6079	156.0000
θ_{17}	107.1519	280.8116	231.9416	223.7404	225.7308	219.6245	217.6851	217.7133	218.7500
ξ	16.7809	2.7240	1.8658	1.4611	1.2704	1.1372	1.0830	1.0462	-

Table 3: The statistical analysis of the objective function and parameter identification error with different noise variances in Example 1

Noise variance, σ_v^2		0.01	0.25	1.0
Objective function, $J(\hat{G}, \hat{f})$	Mean	0.1601	3.9916	15.9640
	Best	0.1590	3.9666	15.8638
	Worst	0.1608	4.0071	16.0264
	Std.	0.0006	0.0114	0.0445
Parameter identification error, ξ	Mean	1.2511	1.3368	1.5126
	Best	1.0462	1.0843	1.1640
	Worst	1.5190	1.8672	1.8715
	Std.	0.1507	0.2009	0.1976

modified SPSA algorithm in (12) and set the parameters of the SPSA-based algorithm $a(k) = 800/(k + 11)^{0.9}$, $c(k) = 0.01/(k + 1)^{1/6}$, $\delta = 0.01$ and $k_{max} = 5000$. The initial values $\mathbf{x}(0)$ are tabulated in Table 4. Then, 25 independent trials are performed to observe the randomization effect.

The best identified parameter values and its parameter identification error ξ are tabulated in Table 5. It shows that the SPSA-based method provides a close approximation to the true parameters, even if the linear subsystem has both slow and fast frequency modes. This is clearly seen from the graph of the nonlinear function f and the bode diagram of linear system G in Figs. 5 and 6, respectively. Thus, we can confirm the superiority of the SPSA-based algorithm in identifying the Hammerstein model, particularly for high-dimensional param-

Table 4: Identification parameters of Example 2

θ	Coefficients	$\mathbf{x}(0)$	θ corresponding to $\mathbf{x}(0)$
θ_1	$\hat{b}_0$	1.0000	10.0000
θ_2	$\hat{a}_0$	1.0000	10.0000
θ_3	$\hat{a}_1$	2.0000	100.0000
θ_4	$\hat{a}_2$	2.0000	100.0000
θ_5	$\hat{a}_3$	2.0000	100.0000
θ_6	$\tilde{\eta}_0$	2.3010	200.0000
θ_7	$\tilde{\eta}_1$	2.2041	160.0000
θ_8	$\tilde{\eta}_2$	2.0792	120.0000
θ_9	$\tilde{\eta}_3$	1.9031	80.0000
θ_{10}	$\tilde{\eta}_4$	1.6021	40.0000
θ_{11}	$\tilde{\eta}_5$	-1.6990	0.0200
θ_{12}	$\tilde{\eta}_6$	1.6021	40.0000
θ_{13}	$\tilde{\eta}_7$	1.9031	80.0000
θ_{14}	$\tilde{\eta}_8$	2.0792	120.0000
θ_{15}	$\tilde{\eta}_9$	2.2041	160.0000
θ_{16}	$\tilde{\eta}_{10}$	2.3010	200.0000

ter identifications.

Moreover, Table 6 shows a similar pattern of statistical data as that of Table 3 in the previous example, in terms of the mean, best, worst and standard deviation values of the objective function and the parameter identification error, respectively, with different noise variances. This illustrates that the SPSA-based algorithm may effectively handle a high noise level by producing relatively smaller parameter identification error.

Table 5: The best identified parameters of Example 2

k	100	500	1000	2000	3000	4000	5000	$\bar{\theta}$
θ_1	6.1660	0.8844	0.5472	0.2605	0.1064	0.1328	0.2442	0.2500
θ_2	107.1519	336.4346	1753.7017	1462.7104	1567.0862	1471.5679	1669.2042	1600.0000
θ_3	70.7946	230.7463	494.7244	428.9361	395.1450	494.9514	417.6439	416.0000
θ_4	154.8817	291.5834	421.2388	414.8130	407.0802	408.5309	404.4467	408.0000
θ_5	12.3027	5.5955	4.5178	4.3391	5.3482	5.5801	5.2192	5.0000
θ_6	123.3190	81.1749	199.6981	226.4612	224.1082	250.6631	239.6720	226.2871
θ_7	130.0529	79.2151	176.3993	178.6007	212.6533	219.5686	229.4416	208.4137
θ_8	84.9535	61.3159	135.0453	150.9329	184.8741	230.2703	185.4966	179.0745
θ_9	81.8634	59.7296	125.3047	141.2728	124.8094	158.8265	134.5685	134.2624
θ_{10}	29.6524	35.7899	47.4979	60.8397	91.3870	101.7111	80.0356	72.8282
θ_{11}	0.0155	0.0217	0.0325	0.0401	0.0780	0.0554	0.0266	3.7497
θ_{12}	42.8608	54.5442	55.1930	62.5358	67.8905	43.4431	76.3260	72.8282
θ_{13}	118.3287	92.5177	133.3952	154.7946	117.0218	105.3911	119.4875	134.2624
θ_{14}	117.2685	92.0525	153.4238	183.2179	189.8304	183.1949	163.2413	179.0745
θ_{15}	179.5230	109.8968	232.6424	226.4605	191.6792	169.6387	213.6546	208.4137
θ_{16}	214.3039	131.0503	234.2946	255.4309	212.8191	236.1105	231.3931	226.2871
ξ	24.3699	3.3692	1.6478	1.0799	1.1861	1.3282	1.0205	-

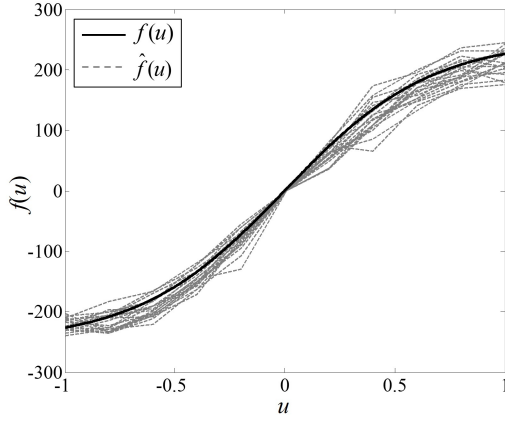
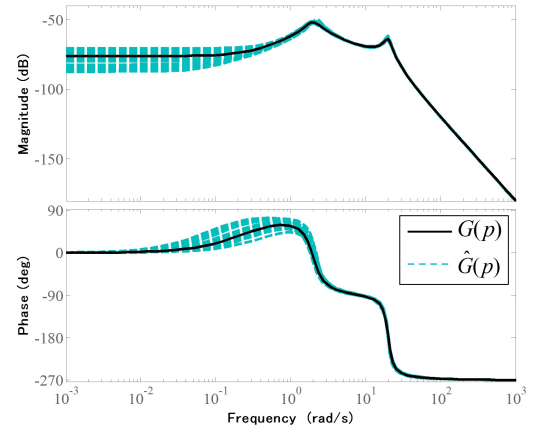
Figure 5: Identified nonlinear functions $\hat{f}(u)$ with the true nonlinear function $f(u)$ in Example 2Figure 6: Identified linear systems $\hat{G}(p)$ with the true linear system $G(p)$ in Example 2

Table 6: The statistical analysis of the objective function and parameter identification error with different noise variances in Example 2

Noise variance, σ_v^2		0.01	0.25	1.0
Objective function, $J(\hat{G}, \hat{f})$	Mean	0.1601	3.9949	15.9749
	Best	0.1594	3.9749	15.8868
	Worst	0.1607	4.0083	16.0216
	Std.	0.0003	0.0093	0.0377
Parameter identification error, ξ	Mean	1.1579	1.4794	1.8459
	Best	1.0205	1.0583	1.4237
	Worst	1.3569	1.7621	2.3504
	Std.	0.0984	0.1698	0.2285

5. Conclusion

In this paper, an identification method of continuous-time Hammerstein models has been presented, which is based on Simultaneous Perturbation Stochastic Approximation (SPSA). The simulation results demonstrate that the proposed method has a good potential to identify continuous-time Hammerstein models, even for high-dimensional parameter identifications. In particular, the method is shown to be effective in approximating both linear and nonlinear subsystems in the Hammerstein models from the view-points of quadratic output prediction error and parameter identification error. Al-

though the results of this work are based only on Hammerstein models, the basic idea can be extended to identify other types of nonlinear models including Wiener ones shown in Fig. 7, where q is another static nonlinear function.

We believe that our proposed identification scheme could be the only practical solution at the moment to solve this identification problem as compared to the other existing methods, such as cuckoo search (Gotmare et al., 2015) and particle swarm optimization (Wang et al., 2014; Jingzhuo et al., 2014). Because the SPSA method only uses two measurements of performance index per iteration, while for cuckoo search and particle swarm optimization methods, the number of measurements per iteration depends on the number of agents, which tends to be very large in order to deal with high-dimensional identification problem. Hence, their computation burden would be too heavy to use.

In the future, it is worth improving the SPSA algorithm, such that it converges much faster when the number identification parameters become larger. Furthermore, it is interesting to validate the applicability of the SPSA-based method in identifying Hammerstein-Wiener models as shown in Fig. 8 and to do the same for multi-input-multi-output (MIMO) nonlinear systems under open-loop, closed-loop or time-delay conditions.

ACKNOWLEDGMENT

This work was partly supported by Grant-in-Aid for Scientific Research (A) 25249058 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

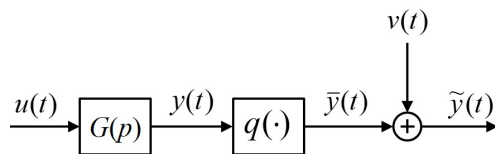


Figure 7: The continuous-time SISO Wiener model

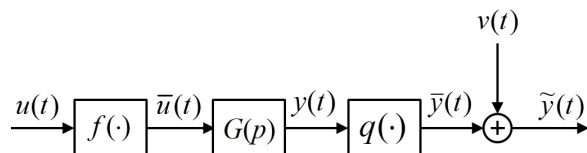


Figure 8: The continuous-time SISO Hammerstein-Wiener model

References

- Ahmad, M., Azuma, S., & Sugie, T. (2014a). Identification of continuous-time Hammerstein models using simultaneous perturbation stochastic approximation. In *Proceedings of the 14th International Conference on Control, Automation and Systems* (pp. 1107–1111).
- Ahmad, M., Azuma, S., & Sugie, T. (2014b). Performance analysis of model-free PID tuning of MIMO systems based on simultaneous perturbation stochastic approximation. *Expert Systems with Applications*, 41, 6361–6370.
- Baba, I., Azuma, S., & Sugie, T. (2013). Controller design for optimizing fuel consumption of hybrid electric vehicles: A model-free approach based on simultaneous perturbation stochastic approximation. *Transactions of the Society of Instrument and Control Engineers*, 49, 887–894.
- Bai, E. W. (2003). Frequency domain identification of Hammerstein models. *IEEE Transactions on Automatic Control*, 48, 530–542.
- Bai, E.-W. (2004). Decoupling the linear and nonlinear parts in hammerstein model identification. *Automatica*, 40, 671–676.
- Bai, E. W., & Fu, M. (2002). A blind approach to Hammerstein model identification. *IEEE Transactions on Signal Processing*, 50, 1610–1619.
- Bai, E.-W., & Li, D. (2004). Convergence of the iterative Hammerstein system identification algorithm. *IEEE Transactions on Automatic Control*, 49, 1929–1940.
- Bilings, S. A., & Fakhouri, S. Y. (1978). Identification of a class of nonlinear systems using correlation analysis. *Proceedings of the Institution of Electrical Engineers*, 7, 691–697.
- Chang, F., & Luus, R. (1971). A noniterative method for identification using Hammerstein model. *IEEE Transactions on Automatic Control*, AC-16, 464–468.
- Chen, H., & Ding, F. (2015). Hierarchical least squares identification for Hammerstein nonlinear controlled autoregressive systems. *Circuits, Systems, and Signal Processing*, 34, 61–75.
- Chen, J., & Wang, X. (2015). Identification of Hammerstein systems with continuous nonlinearity. *Information Processing Letters*, 115, 822–827.
- Ding, F., & Chen, T. (2005). Identification of Hammerstein nonlinear ARMAX systems. *Automatica*, 41, 1479–1489.
- Ding, F., Chen, T., & Iwai, Z. (2007a). Adaptive digital control of Hammerstein nonlinear systems with limited output sampling. *SIAM Journal Control Optimization*, 45, 2257–2276.
- Ding, F., Deng, K., & Liu, X. (2014). Decomposition based newton iterative identification method for a Hammerstein nonlinear FIR system with ARMA noise. *Circuits, Systems, and Signal Processing*, 9, 2881–2893.
- Ding, F., Liu, X., & Liu, G. (2011). Identification methods for Hammerstein nonlinear systems. *Digital Signal Processing*, 21, 215–238.
- Ding, F., Shi, Y., & Chen, T. (2006). Gradient-based identification methods for Hammerstein nonlinear ARMAX models. *Nonlinear Dynamics*, 45, 31–43.
- Ding, F., Shi, Y., & Chen, T. (2007b). Auxiliary model-based least-squares identification methods for Hammerstein output-error systems. *Systems and Control Letters*, 56, 373–380.
- Goethals, I., Pelckmans, K., Suykens, J., & Moor, B. D. (2005). Identification of MIMO Hammerstein models using least squares support vector machines. *Automatica*, 41, 1263–1272.
- Gotmare, A., Patidar, R., & George, N. (2015). Nonlinear system identification using a cuckoo search optimized adaptive Hammerstein model. *Expert Systems with Applications*, 42, 2538–2546.
- Greblicki, W. (1996). Nonlinearity estimation in Hammerstein sys-

- tems based on ordered observations. *IEEE Transactions on Signal Processing*, 44, 1224–1233.
- Greblicki, W. (2000). Continuous-time Hammerstein system identification. *IEEE Transactions on Automatic Control*, 45, 1232–1236.
- Greblicki, W., & Pawlak, M. (1987). Hammerstein system identification by non-parametric regression estimation. *International Journal of Control*, 45, 343–354.
- Hasiewicz, Z., & Mzyk, G. (2004). Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 49, 1370–1375.
- Hsia, T. (1976). A multi-stage least squares method for identifying Hammerstein model nonlinear systems. In *Proceedings of IEEE Conference on Decision and Control* (pp. 934–938).
- Jingzhuo, S., Juanping, Z., Jingtao, H., Meiyu, X., Juwei, Z., & Lei, Z. (2014). Identification of ultrasonic motors nonlinear Hammerstein model. *Journal of Control, Automation and Electrical Systems*, 25, 537–546.
- Ko, C.-N. (2011). Integration of support vector regression and annealing dynamical learning algorithm for MIMO system identification. *Expert Systems with Applications*, 38, 15224–15233.
- Krzyak, A. (1993). Identification of nonlinear block-oriented systems by the recursive kernel estimate. *International Journal of Franklin Institute*, 330, 605–627.
- Krzyak, A. (1996). On nonparametric estimation of nonlinear dynamic systems by the fourier series estimate. *Signal Processing*, 52, 299–321.
- Laurain, V., Gilson, M., & Garnier, H. (2009). Refined instrumental variable methods for identifying Hammerstein models operating in closed loop. In *Proceedings of IEEE Conference on Decision and Control and Chinese Control Conference* (pp. 314–320).
- Li, C.-H., Zhu, X.-J., Cao, G.-Y., Sui, S., & Hu, M.-R. (2008). Identification of the Hammerstein model of a PEMFC stack based on least squares support vector machines. *Journal of Power Sources*, 175, 303–316.
- Liu, Y., & Bai, E.-W. (2007). Iterative identification of Hammerstein systems. *Automatica*, 43, 346–354.
- Ma, L., & Liu, X. (2015). A nonlinear recursive instrumental variables identification method of Hammerstein ARMAX system. *Nonlinear Dynamics*, 79, 1601–1613.
- Mao, Y., & Ding, F. (2015). Multi-innovation stochastic gradient identification for Hammerstein controlled autoregressive autoregressive systems based on the filtering technique. *Nonlinear Dynamics*, 79, 1745–1755.
- Maruta, I., Kim, T.-H., Song, D., & Sugie, T. (2013). Synthesis of fixed-structure robust controllers using a constrained particle swarm optimizer with cyclic neighborhood topology. *Expert Systems with Applications*, 40, 3595–3605.
- Maruta, I., Kim, T.-H., & Sugie, T. (2009). Fixed-structure H-infinity controller synthesis: A meta-heuristic approach using simple constrained particle swarm optimization. *Automatica*, 45, 553–559.
- Nanda, S., Panda, G., & Majhi, B. (2010). Improved identification of Hammerstein plants using new CPSO and IPSO algorithms. *Expert Systems with Applications*, 37, 6818–6831.
- Narendra, K., & Gallman, P. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, AC-11, 546–550.
- Pawlak, M. (1991). On the series expansion approach to the identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 36, 763–767.
- Rangan, S., Wolodkin, G., & Polla, K. (1995). Identification methods for Hammerstein systems. In *Proceedings of IEEE Conference on Decision and Control* (pp. 697–702).
- Shi, J., Xu, X., & Dai, Y. (2011). Identification of Hammerstein LSSVM-ARMA systems and its application in continuous stirred tank reactor. *Energy Procedia*, 13, 5395–5365.
- Spall, J. (1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37, 332–341.
- Stoica, P. (1981). On the convergence of an iterative algorithm used for Hammerstein system identification. *IEEE Transactions on Automatic Control*, 26, 967–969.
- Stoica, P., & Soderstrom, T. (1981). Instrumental-variable methods for identification of Hammerstein systems. *International Journal of Control*, 35, 459–476.
- Togun, N., Baysec, S., & Kara, T. (2012). Nonlinear modeling and identification of a spark ignition engine torque. *Mechanical Systems and Signal Processing*, 26, 294–304.
- van der Veen, G., van Wingerde, J.-W., & Verhaegen, M. (2013). Global identification of wind turbines using a Hammerstein identification method. *IEEE Transactions on Control System Technology*, 21, 1471–1478.
- Verhaegen, M., & Westwick, D. (1996). Identifying MIMO Hammerstein systems in the context of subspace model identification methods. *International Journal of Control*, 63, 331–349.
- Voros, J. (1997). Parameter identification of discontinuous Hammerstein systems. *Automatica*, 33, 1141–1146.
- Wang, D., Ding, F., & Ximei, L. (2015). Least squares algorithm for an input nonlinear system with a dynamic subspace state space model. *Nonlinear Dynamics*, 75, 49–61.
- Wang, D., Ren, Y., Liu, C., & Han, P. (2014). Identification of thermal process using Hammerstein model based on particle swarm optimization algorithm. *Lecture Notes in Electrical Engineering*, (pp. 1961–1968).
- Wang, J., & Zhang, Q. (2014). Detection of asymmetric control valve stiction from oscillatory data using an extended Hammerstein system identification method. *Journal of Process Control*, 24, 1–12.
- Yu, C., Zhang, C., & Xie, L. (2014). A new deterministic identification approach to Hammerstein systems. *IEEE Transactions on Signal Processing*, 62, 131–140.
- Zhao, W. (2006). Recursive identification for Hammerstein system with ARX subsystem. *IEEE Transactions on Automatic Control*, 51, 1966–1974.