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

Mohd Ashraf Ahmad\textsuperscript{a,b}, Shun-ichi Azuma\textsuperscript{b}, Toshiharu Sugi\textsuperscript{b}

\textsuperscript{a}Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, 26600, Pekan, Pahang, Malaysia
\textsuperscript{b}Department of Systems Science, Graduate School of Informatics, Kyoto University, Yoshida-Hommachi, 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;
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 framework.

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 supplantation (Maruta et al., 2009, 2013). However, they have a serious drawback.

(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 evolitional 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 \( \theta \), we use \( \| \theta \|_2 \) to express the standard Euclidean norm. For \( \delta \in \mathbb{R}_+ \), \( \text{sat}_\delta : \mathbb{R}^n \to \mathbb{R}^n \) denotes the saturation function whose \( i \)-th element given as follows:

\[
\text{sat}_\delta(x) = \begin{cases} 
\delta & \text{if } \delta < x_i, \\
-x_i & \text{if } -\delta \leq x_i \leq \delta, \\
-x_i & \text{if } x_i < -\delta.
\end{cases}
\]

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

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 \( p \) (\( \approx \frac{d}{dt} \)):

\[
G(p) = \frac{B(p)}{A(p)} = \frac{b_m p^m + b_{m-1} p^{m-1} + \cdots + b_0}{p^n + a_{n-1} p^{n-1} + \cdots + a_0}, \tag{1}
\]

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

\[
\tilde{y}(t) = G(p)f(u(t)) + v(t). \tag{2}
\]

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

- \( m \) and \( n \) are known.
- \( a_i \) (\( i = 0, 1, ..., n - 1 \)) and \( b_i \) (\( i = 0, 1, ..., m \)) 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.

\[
\begin{array}{c}
u(t) \\
\downarrow \\
\tilde{u}(t) \\
\downarrow \\
\tilde{y}(t) \\
\downarrow \\
y(t) \\
\end{array}
\]

\[
\begin{array}{c}
u(t) \\
\downarrow \\
\tilde{u}(t) \\
\downarrow \\
\tilde{y}(t) \\
\downarrow \\
y(t) \\
\end{array}
\]

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

\[
J(\hat{G}, \hat{f}) = \sum_{t=0}^{N} (\hat{y}(t) - \tilde{y}(t))^2 \tag{3}
\]

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

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 \((u(t), \tilde{y}(t)) \) (\( t = 0, t_s, 2t_s, ..., Nt_s \)) are given. Then, find \( \hat{G} \) and \( \hat{f} \) which minimizes \( J(\hat{G}, \hat{f}) \).

3. Identification Method Using Simultaneous Perturbation Stochastic Approximation

This section proposes a method to solve Problem 2.1. First, the standard SPSPA 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 SPSPA-based algorithm.

3.1. Review of Simultaneous Perturbation Stochastic Approximation

Consider the optimization problem

\[
\min_{x \in \mathbb{R}^n} h(x) \tag{4}
\]
where \( h : \mathbb{R}^n \rightarrow \mathbb{R} \) is the objective function and \( x \in \mathbb{R}^n \) denotes the design variable.

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

\[
x(k + 1) = x(k) - a(k)g(x(k), \triangle(k))
\]

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

\[
g(x(k), \triangle(k)) = \begin{bmatrix}
\frac{h(x(k) + c(k)_0 \triangle(k)) - h(x(k))}{2c(k)_0} \\
\frac{h(x(k) + c(k)_1 \triangle(k)) - h(x(k) - c(k)_1 \triangle(k))}{2c(k)_1} \\
\vdots \\
\frac{h(x(k) + c(k)_n \triangle(k)) - h(x(k) - c(k)_n \triangle(k))}{2c(k)_n}
\end{bmatrix}.
\]

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

The idea of this algorithm is that the expectation of \( g(x(k), \triangle(k)) \) is nearly equal to the gradient of the objective function \( h \), i.e., \( \frac{d}{dx} h(x(k)) \), 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 \( \triangle(k) \) 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

\[
\hat{f}(u) = \begin{cases} 
\eta_0 + M_1(u - w_0) & \text{if } w_0 \leq u < w_1, \\
\eta_1 + M_2(u - w_1) & \text{if } w_1 \leq u < w_2, \\
\vdots & \\
\eta_{i-1} + M_i(u - w_{i-1}) & \text{if } w_{i-1} \leq u < w_i,
\end{cases}
\]

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

\[
\eta_i = \begin{cases} 
\bar{\eta}_i & \text{if } w_i > 0, \\
0 & \text{if } w_i = 0, \\
-\bar{\eta}_i & \text{if } w_i < 0,
\end{cases}
\]

for \( i = 0, 1, \ldots, \ell \), where \( \bar{\eta}_i \in \mathbb{R}_+ \). By choosing \( \ell \) large, the accuracy of the approximation increases. However, such a 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

\[
h(\theta) = \sum_{k=0}^{N} \left( \bar{y}(kt_k) - \hat{y}(kt_k) \right)^2.
\]
for the design variable

$$\theta = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_{m-1} \\ \hat{a}_0 \\ \vdots \\ \hat{a}_{n-1} \\ \hat{h}_0 \\ \vdots \\ \hat{h}_\ell \end{bmatrix} \in \mathbb{R}^p$$

(10)

where $p = m + n + \ell + 1$. 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 $\theta$ by setting $\theta_i = 10^{x_i}$ ($i = 1, 2, \ldots, p$) with the objective function $h(10^{x_1} 10^{x_2} \cdots 10^{x_p})$. Then, our design procedure is summarized as follows:

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

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

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

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

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

(11)

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

**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,

$$x(k + 1) = x(k) - \text{sat}_\delta(u(k)g(x(k), \triangle(k))).$$

(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:

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

(13)

$$A(p) = p^6 + 10.0000p^5 + 54.7700p^4 + 156.8000p^3 + 87.0843p^2 + 25.2810p + 4.0197, B(p) = p^3,$$

$$f(u(t)) = 125(u(t) + 0.5u^2(t) + 0.25u^3(t)).$$

(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 $\hat{y}(t)$ 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 $w := (-1, -0.8, -0.6, -0.4, -0.2, 0.01, 0.2, 0.4, 0.6, 0.8, 1)^T \in \mathbb{R}^{11}$. The design variables $\theta \in \mathbb{R}^{17}$ 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 $k_{\text{max}} = 5000$. The initial values $x(0)$ 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 $\xi$. In this table, $\hat{\theta} \in \mathbb{R}^p$ is a vector, whose elements consist of true identification parameters, and $\xi$ is given by

$$\xi = \left\| \frac{\hat{\theta}_1 - \tilde{\theta}_1}{\tilde{\theta}_1}, \ldots, \frac{\hat{\theta}_p - \tilde{\theta}_p}{\tilde{\theta}_p} \right\|_2$$

(15)

where $\tilde{\theta}_i$ is the $i$-th element of the vector $\tilde{\theta}$. It shows that the resultant identified parameters are close to the
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 tangent hyperbolic function given by

$$G(p) = \frac{B(p)}{A(p)}.$$  

(16)

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

(17)

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

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
This paper presents the identification of parameters in both slow and fast frequency modes. This is achieved through the use of a modified SPSA algorithm, as seen in Table 3, which tabulates the best identified parameters of Example 1. The algorithm 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 parameter identification.

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

<table>
<thead>
<tr>
<th>$k$</th>
<th>100</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
<th>4000</th>
<th>5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>6.1660</td>
<td>0.8844</td>
<td>0.5472</td>
<td>0.2605</td>
<td>0.1064</td>
<td>0.1328</td>
<td>0.2442</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>107.1519</td>
<td>336.4346</td>
<td>1753.7017</td>
<td>1462.7104</td>
<td>1567.0862</td>
<td>1471.5679</td>
<td>1669.2042</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>70.7946</td>
<td>230.7463</td>
<td>494.7244</td>
<td>428.9361</td>
<td>395.1450</td>
<td>494.9514</td>
<td>417.6439</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>154.8817</td>
<td>291.5834</td>
<td>421.2388</td>
<td>414.8130</td>
<td>407.0802</td>
<td>408.5309</td>
<td>404.4647</td>
</tr>
<tr>
<td>$\theta_5$</td>
<td>12.3027</td>
<td>5.5955</td>
<td>4.5178</td>
<td>4.3391</td>
<td>5.3482</td>
<td>5.5801</td>
<td>5.2192</td>
</tr>
<tr>
<td>$\theta_6$</td>
<td>123.3190</td>
<td>81.1749</td>
<td>199.6981</td>
<td>226.4612</td>
<td>224.1082</td>
<td>250.6631</td>
<td>239.6720</td>
</tr>
<tr>
<td>$\theta_7$</td>
<td>29.6524</td>
<td>35.7899</td>
<td>47.4979</td>
<td>60.8397</td>
<td>91.3870</td>
<td>101.7111</td>
<td>80.0356</td>
</tr>
<tr>
<td>$\theta_8$</td>
<td>0.0155</td>
<td>0.0217</td>
<td>0.0325</td>
<td>0.0401</td>
<td>0.0780</td>
<td>0.0554</td>
<td>0.0266</td>
</tr>
<tr>
<td>$\theta_9$</td>
<td>42.8608</td>
<td>54.5442</td>
<td>55.1930</td>
<td>62.5358</td>
<td>67.8905</td>
<td>43.4431</td>
<td>76.3260</td>
</tr>
<tr>
<td>$\theta_{10}$</td>
<td>118.3287</td>
<td>92.5177</td>
<td>133.3952</td>
<td>154.7946</td>
<td>117.0218</td>
<td>105.3911</td>
<td>119.4875</td>
</tr>
<tr>
<td>$\theta_{11}$</td>
<td>117.2685</td>
<td>92.0525</td>
<td>153.4238</td>
<td>183.2179</td>
<td>189.8304</td>
<td>183.1949</td>
<td>179.0745</td>
</tr>
<tr>
<td>$\theta_{12}$</td>
<td>119.2530</td>
<td>109.8968</td>
<td>232.6424</td>
<td>226.4605</td>
<td>191.6792</td>
<td>219.5686</td>
<td>229.4416</td>
</tr>
<tr>
<td>$\theta_{13}$</td>
<td>84.9535</td>
<td>61.3159</td>
<td>135.0453</td>
<td>150.9329</td>
<td>184.8741</td>
<td>230.2703</td>
<td>185.4966</td>
</tr>
<tr>
<td>$\theta_{14}$</td>
<td>12.3027</td>
<td>5.5955</td>
<td>4.5178</td>
<td>4.3391</td>
<td>5.3482</td>
<td>5.5801</td>
<td>5.2192</td>
</tr>
<tr>
<td>$\theta_{15}$</td>
<td>123.3190</td>
<td>81.1749</td>
<td>199.6981</td>
<td>226.4612</td>
<td>224.1082</td>
<td>250.6631</td>
<td>239.6720</td>
</tr>
<tr>
<td>$\theta_{16}$</td>
<td>29.6524</td>
<td>35.7899</td>
<td>47.4979</td>
<td>60.8397</td>
<td>91.3870</td>
<td>101.7111</td>
<td>80.0356</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Noise variance, $\sigma_v^2$</th>
<th>0.01</th>
<th>0.25</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective function, $J(\hat{G}, f)$</td>
<td>Mean</td>
<td>0.1601</td>
<td>3.9949</td>
</tr>
<tr>
<td></td>
<td>Best</td>
<td>0.1594</td>
<td>3.9749</td>
</tr>
<tr>
<td></td>
<td>Worst</td>
<td>0.1607</td>
<td>4.0838</td>
</tr>
<tr>
<td></td>
<td>Std.</td>
<td>0.0003</td>
<td>0.0093</td>
</tr>
<tr>
<td>Parameter identification error, $\xi$</td>
<td>Mean</td>
<td>1.1579</td>
<td>1.4794</td>
</tr>
<tr>
<td></td>
<td>Best</td>
<td>1.0205</td>
<td>1.0583</td>
</tr>
<tr>
<td></td>
<td>Worst</td>
<td>1.3569</td>
<td>1.7621</td>
</tr>
<tr>
<td></td>
<td>Std.</td>
<td>0.0984</td>
<td>0.1698</td>
</tr>
</tbody>
</table>

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 viewpoints 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.

References


