Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-10 (2016) 019–024

# An Optimal Two Stage Identification Algorithm for Discrete Hammerstein Time Delay Systems

**Asma Atitallah, Saïda Bedoui, Kamel Abderrahim**

*National Engineering school of Gabes, Research Unit: Numerical Control of Industrial Processes, Gabes University, Tunisia, (e-mail: atitallah.asma@yahoo.fr, saida.bedoui@enig.rnu.tn and kamelabderrahim@yahoo.fr)*

**Abstract:** This paper deals with the problem of identification of Hammerstein time delay systems. In fact, we propose a Hierarchical optimization based approach allowing to estimate simultaneously the time delay, the linear dynamic parameters and the nonlinear static parameters of such systems. This approach consists, firstly, in decomposing a complex nonlinear cost function into two simple cost functions and secondly, in using the gradient method to minimize each function (HH-G). The convexity condition to authorize the use of the rounding property is also developed and the convergence analysis of the proposed algorithm indicates that the parameter estimation errors converge to zero under persistent excitation conditions and the estimated time delay converge to a finite value. The simulation results are presented to illustrate the effectiveness of the proposed method.



*Keywords:* Identification; Hammerstein systems; Time delay; Gradient approach; Hierarchical principle; Convexity; Convergence analysis.

## 1. INTRODUCTION

Block-oriented models are often used to describe many nonlinear systems (Giri and Bai (2010); Billings (2013)). The most well known are the Hammerstein and the Wiener models (Ding and Chen (2005); Billings (2013)). They consist of a single linear dynamic block and static nonlinear subsystem. In general, these models are used to decompose a complex system in simpler subsystems and their main advantages over other nonlinear models lie in their structured and parsimonious nature.

System identification which contains generally parameter and time delay estimations is a technique that can be used to construct system models from experimental data. Numerous approaches have been presented in the literature to estimate only the parameters of Hammerstein systems (Ding and Chen (2005); Ding et al. (2007, 2013); Xiao et al. (2010); Wang et al. (2014)). However, time delay estimation remains until now a discussing problem for this kind of nonlinear systems. In fact, identification of time delay is one of the most important topics in the field of time delay system and it should be treated as first task during system analysis and control design (Boukas and Liu (2002); Richard (2003); Atitallah et al. (2015b); Bedoui et al. (2013); Atitallah et al. (2015a)). Indeed, identifying time delay is not an easy task for systems with input or/and state delays or/and output delay. Thus due to their specific structure parameter that differs both from the system order and from the system parameters.

In the following, we propose an estimation algorithm based on the hierarchical gradient method for Hammerstein time delay systems to identify both the parameters and the time delay. We develop a condition to authorize the use of the rounding property and we discuss the convergence analysis of the proposed algorithm using the martingale convergence theorem.

The remaining parts of this paper are organized as follows. The identification problem is described in Section 2. In the third Section, the proposed algorithm is presented. Section 4 is devoted to the convexity problem and the convergence analysis. An illustrative example to show the effectiveness of the proposed algorithm is provided in Section 5. Finally, some concluding remarks are offered in section 6.

## 2. SYSTEM DESCRIPTION AND IDENTIFICATION MODEL

Consider the Hammerstein time delay system expressed as:

$$y(k) = \left[1 - A(q^{-1})\right] y(k) + q^{-d} B(q^{-1}) \bar{u}(k) + v(k), \qquad (1)$$

where $u(k)$ and $y(k)$ are the system input and output, respectively, $v(k)$ is a white noise with zero mean and finite variance, $\bar{u}(k)$ is the output of nonlinear part, $d$ is the unknown time delay, and $A(q^{-1})$ and $B(q^{-1})$ are two coprime polynomials of orders $n_a$ and $n_b$ in the shift operator $q^{-1}$ with:

$$A(q^{-1}) := 1 + \sum_{i=1}^{n_a} a_i q^{-i},\ B(q^{-1}) := \sum_{j=1}^{n_b} b_j q^{-j}.$$

The nonlinear part in the Hammerstein model is a nonlinear function of known basis $f := (f_1, f_2, ..., f_{n_c})$ with coefficients $(c_1, c_2, ..., c_{n_c})$ given by:

$$\begin{aligned} \bar{u}(k) = f(u(k)) &= \sum_{i=1}^{n_c} c_i f_i(u(k)) \\ &= c_1 f_1(u(k)) + c_2 f_2(u(k)) + \cdots + c_{n_c} f_{n_c}(u(k)). \end{aligned} \qquad (2)$$

Assume that the base functions $f_i(u(k)$ are polynomials, the degrees $n_a$, $n_b$ and $n_c$ are known and $y(k) = 0$, $u(k) = 0$, $\bar{u}(k) = 0$ and $v(k) = 0$ for $k \leq 0$.
Then, the Hammerstein time delay system can be expressed as:


10.1016/j.ifacol.2016.07.466

$$y(k) = -\sum_{i=1}^{n_a} a_i y(k-i) + q^{-d} \sum_{i=1}^{n_b} b_i \sum_{j=1}^{n_c} c_j f_j(u(k-i)) + v(k) \tag{3}$$

Equation (3) can be rewritten as:

$$y(k) = -\sum_{i=1}^{n_a} a_i y(k-i) + q^{-d} \sum_{j=1}^{n_c} c_j \sum_{i=1}^{n_b} b_i f_j(u(k-i)) + v(k) \tag{4}$$

Notice that the Hammerstein model shown in (4) is not unique. Indeed, any pair $\beta b$ and $c/\beta$ for some nonzero and finite constants $\beta$ would produce the same input and output measurements. Therefore, to get a unique parameterization, one of the elements of $b$ or $c$ has to be fixed. There are several ways to normalize the elements see (Cerone and Regruto (2003); Bai and Li (2004); Liu and Bai (2007); Ding et al. (2011b)).
In fact, we have considered here the following assumption:

**Uniqueness assumption:** The first elements of the vectors $c$ is equal 1, i.e., $c_1 = 1$ (Bai and Li (2004)).

Define the parameter vector $\theta$ and the information vector $\varphi(k)$ as:

$$\theta = [a^T, c_1 b^T, c_2 b^T, ..., c_{n_c} b^T]^T \in \mathbb{R}^{n_p}, \tag{5}$$

$$\varphi(k) = [\psi_0(k); \psi_1(k); \psi_2(k); ...; \psi_{n_c}(k)]^T \in \mathbb{R}^{n_p}, \tag{6}$$

$$\psi_0(k) = [-y(k-1), -y(k-2), ..., -y(k-n_a)]^T \in \mathbb{R}^{n_a}, \tag{7}$$

$$\psi_j(k) = q^{-d}[f_j(u(k-1)), ..., f_j(u(k-n_b))]^T = q^{-d}\xi_j(k), \tag{8}$$

$$\xi_j(k) = [f_j(u(k-1)), ..., f_j(u(k-n_b))]^T \in \mathbb{R}^{n_b}, \tag{9}$$

where $j = 1, 2, ..., n_c$, $q^{-d}$ is a delay operator and $n_p := n_a + n_b n_c$ is defined as the dimension of the parameter vector.
The parameters $a$, $b$ and $c$ are given by:

$$a := [a_1, a_2, ..., a_{n_a}]^T \in \mathbb{R}^{n_a},$$
$$b := [b_1, b_2, ..., b_{n_b}]^T \in \mathbb{R}^{n_b},$$
$$c := [c_1, c_2, ..., c_{n_c}]^T \in \mathbb{R}^{n_c},$$

Then, equation (4) can be rewritten as:

$$y(k) = \psi_0^T(k) a + q^{-d} \sum_{j=1}^{n_c} c_j \xi_j^T(k) b + v(k), \tag{10}$$

or, it can be defined as:

$$y(k) = \varphi^T(k)\theta + v(k). \tag{11}$$

*Our goal is to estimate the system parameters a, b and c and the time delay d from sampled data of the input and the noisy output of the system. It is well known that the time delay, in the discrete time domain, is an integer. However, the proposed approach will return a real value. This value will be rounding using the rounding property, see (Bedoui et al. (2013); Hübner and Schöbel (2014))*

## 3. HIERARCHICAL IDENTIFICATION ALGORITHM

In this section, we develop a method based on the hierarchical principle to estimate separately the parameters and the time delay (HH-G).
Consider the cost function:

$$J(\theta, d) = \frac{1}{2} e^2(k) \tag{12}$$

where $e(k)$ is the prediction error given by:

$$e(k) = y(k) - \hat{y}(k) = y(k) - \varphi^T(k)\hat{\theta}(k) \tag{13}$$

The model in (12) contains the product of two parameter vectors $b$ and $c$ and the delay operator $q^{-d}$. This cost function $J$ is a nonlinear cost function. Hence, the parameter vector and the time delay identification are difficult in the same step. So to solve this problem, an alternative approach which consists in using the hierarchical identification principle (Bedoui et al. (2013); Wang et al. (2014); Atitallah et al. (2015b)) is used. The principle is to decompose the cost function $J$ into two cost functions $J(\theta, \hat{d}(k-1)$ for fixed $d = \hat{d}(k-1)$ and $J(\hat{\theta}(k), d)$ for fixed $\theta = \hat{\theta}(k)$. This is equivalent to minimize the following two optimization problems:

**Problem 1.** The optimization of $\theta$:

$$\hat{\theta}(k) = \arg\min[(J^2(k)\Big|_{\hat{d}(k-1)})] \text{ for fixed } \hat{d}(k-1), \quad \theta \in \mathbb{R}^{n_p}$$

**Problem 2.** The optimization of $d$:

$$\hat{d}(k) = \arg\min[(J^2(k)\Big|_{\hat{\theta}(k)})] \text{ for fixed } \hat{\theta}(k), \qquad d \in \mathbb{Z}^+$$

***Parameter estimation*** Problem 1 leads to define the following cost function as:

$$J_1(\theta) = \frac{1}{2} e^2(k)\Big|_{\hat{d}(k-1)} \quad \theta \in \mathbb{R}^{n_p} \tag{14}$$

The negative gradient search method is then used to estimate the parameter vector $\hat{\theta}(k)$. Solving the optimization problem (14) leads to the following algorithm of computing $\hat{\theta}(k)$:

$$\hat{\theta}(k) = \hat{\theta}(k-1) - \mu_1(k) \frac{\partial J_1(k)}{\partial \hat{\theta}(k-1)} \tag{15}$$

where $\mu_1(k)$ is the convergence factor.
Then equation (15) can be written as:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \mu_1(k)\hat{\varphi}(k) e_1(k) \tag{16}$$

where

$$e_1(k) = y(k) - \hat{\varphi}^T(k)\hat{\theta}(k-1) \tag{17}$$

$$\hat{\varphi}(k) = [-y(k-1), ..., -y(k-n_a), \hat{\psi}_1(k), ..., \hat{\psi}_{n_c}(k)]^T \tag{18}$$

and

$$\hat{\psi}_j(k) = q^{-\hat{d}(k-1)} \xi_j(k) \tag{19}$$

for $j = 1, 2, ..., n_c$.
According to the uniqueness assumption ($c_1 = 1$), the estimates $\hat{a} := [\hat{a}_1, \hat{a}_2, ..., \hat{a}_{n_a}]^T$ and $\hat{b} := [\hat{b}_1, \hat{b}_2, ..., \hat{b}_{n_b}]^T$ are derived from the first elements $n_a$ and $n_b$ of $\theta$ respectively. To obtain the estimated parameters $\hat{c}_j (j = 2, 3, ..., n_c)$, we use the average method which consists in separating these parameters from the parameter vector $\theta$:

$$\hat{c}_j = \frac{\hat{\theta}_{n_a + (j-1)n_b + i}}{\hat{b}_i}, j = 2, ..., n_c, i = 1, ..., n_b \tag{20}$$

From the above equation, we know that there are $n_b$ estimates $\hat{c}_j$, $(j = 2, 3, ..., n_c)$ for each $c_j$. Obviously, great redundancy exists in this processing of $\hat{c}_j$. To solve this problem, one option is choosing their mean value as the estimate of $c_j$:

$$\hat{c}_j = \frac{1}{n_b} \sum_{i=1}^{n_b} \frac{\hat{\theta}_{n_a + (j-1)n_b + i}}{\hat{b}_i}, j = 2, ..., n_c. \tag{21}$$

Note that the convergence factor $\mu_1(k)$ can be chosen as follows (Atitallah et al. (2015b)):

$$\mu_1(k) = \frac{1}{r(k)}, \tag{22}$$

with

$$r(k) = r(k-1) + \|\hat{\varphi}(k)\|^2, \quad r(0) = 1, \tag{23}$$