

## A Blind Approach to Hammerstein Model Identification

Er-Wei Bai\* and Minyue Fu%

\*Dept. of ECE, University of Iowa, Iowa City, Iowa 52242

%Dept. of ECE, The University of New Castle, Callaghan, Australia

### Abstract

This paper discusses discrete Hammerstein model identification using blind system identification approach. By sampling faster at the output for the sampled Hammerstein systems, it is shown that identification of the linear part can be achieved based only on the output measurements that makes Hammerstein model identification possible without knowing the structure of the nonlinearity and the internal variable. The fundamental identifiability problem is solved and several schemes are presented.

### 1. Introduction

The Hammerstein model is a special kind of nonlinear system [7] where a nonlinear block is followed by a linear system. The Hammerstein model has applications in many engineering problems and therefore, identification of Hammerstein model has been a research topic over three decades. There exists a large number of research papers on the topic of Hammerstein model identification in the literature. Existing methods can be roughly divided into four categories: the iterative method [9], the over-parameterization method [2], the correlation method [5, 6] and the separable least squares method [1].

In this paper, we consider identification of discrete Hammerstein systems. In particular, we focus our study and explain our ideas in detail for the sampled Hammerstein systems as shown in Figure 1. Then, we will extend results to non-sampled discrete Hammerstein systems. Consider the sampled Hammerstein system in Figure 1 with sampling interval  $T$ , the goal of the sampled Hammerstein system identification is to estimate the transfer function of the equivalent sampled linear system for the given sampling interval, and to estimate the unknown nonlinear function  $f$  based only on the measurement of  $u$  and  $y$ . The internal signal  $x$  is not assumed to be available.

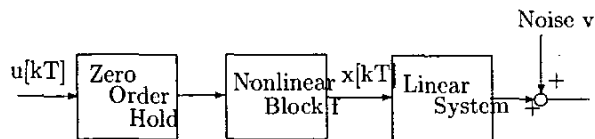


Figure 1: Sampled Hammerstein system

Our approach in this paper is different from all four methods discussed above and is based on the idea of our previous work on blind system identification [3]. We identify the linear part using the output measurements only and no information on the input  $u[kT]$  and the interval variable  $x[kT]$  are needed. In general, blind system identification is not possible only based on the output measurements because different systems coupled with appropriate inputs can produce identical outputs at the sampling instants  $kT$ . However, by faster sampling at the output and keeping the input sampling interval unchanged, blind identification based on the output measurements is possible. Once the linear part is obtained, identification of the nonlinear part can be carried out in a number of ways. We remark that the blind system identification techniques for IIR system was first developed in our early work [3]. The techniques were introduced to the identification of the sampled Hammerstein model in a recent paper [11]. However, some important questions were left unanswered. The most important one is the identifiability question. Recall that the goal of Hammerstein model identification is to find the equivalent sampled linear system for a given sampling interval  $T$ . Because of faster sampling required by the blind system identification technique, the transfer function at the given sampling interval  $T$  was not obtained in [11] and instead a transfer function at a much higher sampling rate is calculated. It was not clear whether the transfer function at the given sampling interval  $T$  is obtainable without reconstructing the internal variable  $x[kT]$ . This fundamental identifiability question is investigated with a positive answer in this paper which shows that such a transfer function is obtainable without reconstructing  $x[kT]$ .

### 2. Problem statement and preliminaries

Consider a sampled Hammerstein model in Figure 1, which consists of a Zero Order Hold, a nonlinear block and a continuous linear time-invariant system. For a given sampling interval  $T$ , the input  $u[kT]$  is a discrete sequence. The output of the nonlinear block  $x$ , which is also the input to the linear system, is also assumed to be a piecewise

constant signal represented by

$$\begin{aligned} x(t) &= x[kT], \quad kT \leq t < (k+1)T \\ x[kT] &= f(u[kT], \dots, u[(k-m)T], \gamma) \end{aligned} \quad (2.1)$$

where  $f$  is a nonlinear function with known or unknown structure parameterized by an unknown parameter vector  $\gamma \in \mathbb{R}^l$ . Let the continuous time system be represented by an  $n$ th order state space equation. Then, the discrete transfer function from  $x[kT]$  to  $y[kT]$  is given by

$$G(z) = \frac{\beta(z)}{\alpha(z)} = \frac{b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_n z^{-n}} \quad (2.2)$$

form some  $a_i$ 's and  $b_j$ 's. The goal of the Hammerstein system identification is to estimate  $G(z)$  in terms of its parameters  $a_i$ 's and  $b_j$ 's, as well as to estimate the unknown nonlinear function  $f$  based only on the measurement of  $u$  and  $y$ .

We now make an assumption on the sampled system throughout the paper.

**Assumption 1:** It is assumed that the sampled system at the sampling interval  $T$  is minimal (reachable and observable).

This assumption actually implies that [3] the sampled system is minimal at any sampling interval  $h = T/p$  for any integer  $p \geq 1$ .

Our approach in this paper is based on blind system identification, i.e., to estimate  $G(z)$  using only the output measurements. Let the output sampling interval be

$$h = T/p, \quad p \geq 1$$

for some positive integer  $p$ , referred to as the over-sampling ratio. For given  $T$  and  $p$ , consider the following sequences

$$\begin{aligned} \{y[kh]\} &\iff Y_p(z) = \sum_{k=0}^{\infty} y[kh]z^{-k} \\ \{y[kT]\} &\iff Y(z) = \sum_{k=0}^{\infty} y[kT]z^{-k} \\ \{y[kT+h]\} &\iff Y_{T,1}(z) = \sum_{k=0}^{\infty} y[kT+h]z^{-k} \\ &\dots \end{aligned} \quad (2.3)$$

Although the input sampling interval is fixed at  $T$  and

$$\begin{aligned} \{u[kT]\} &\iff U(z) = \sum_{k=0}^{\infty} u[kT]z^{-k} \\ \{x[kT]\} &\iff X(z) = \sum_{k=0}^{\infty} x[kT]z^{-k}, \end{aligned} \quad (2.4)$$

$X(z)$  can be written as

$$\begin{aligned} X_p(z) &= \sum_{k=0}^{\infty} x[kh]z^{-k} \\ &= x[0] + x[h]z^{-1} + \dots + x[T-h]z^{-(p-1)} \\ &+ x[T]z^{-p} + x[T+h]z^{-(p+1)} + \dots + x[2T-h]z^{-(2p-1)} + \dots \\ &= (1 + z^{-1} + \dots + z^{-(p-1)})X(z^p). \end{aligned} \quad (2.5)$$

Denote by  $G_p(z)$  the discrete transfer function from  $x[kh]$  to  $y[kh]$  and  $G_{T,i}(z)$ ,  $0 \leq i \leq (p-1)$ , the discrete transfer functions from  $x[kT]$  to  $y[kT+ih]$ , i.e.,

$$Y_p(z) = G_p(z)X_p(z), \quad Y_{T,i}(z) = G_{T,i}(z)X(z).$$

The transfer functions of  $G_{T,i}(z)$  can be easily derived

$$G_{T,i}(z) = \frac{\beta_i(z)}{\alpha_i(z)} = \frac{b_0^i + b_1^i z^{-1} + \dots + b_n^i z^{-n}}{1 - a_1^i z^{-1} - a_2^i z^{-2} - \dots - a_n^i z^{-n}}. \quad (2.6)$$

It is interesting to note that

- All  $G_{T,i}(z)$ 's share the same denominator, i.e.,  
 $a_1^i = a_1, a_2^i = a_2, \dots, a_n^i = a_n, \quad 0 \leq i \leq p-1.$
- $G_{T,0}(z) = G(z)$  as in (2.2) and this implies  
 $b_0^0 = b_0, b_1^0 = b_1, \dots, b_n^0 = b_n.$
- $G_{T,0}(z) = G(z)$  is strictly proper and from (2.6),  $G_{T,i}(z)$ 's,  $i = 1, \dots, p-1$ , are proper but not strictly proper.

**Lemma 2.1** Let  $T = ph$  for some integer  $p \geq 1$ . Suppose the transfer function  $G_p(z)$  at the sampling interval  $h = T/p$  is in the form of

$$G_p(z) = \frac{\beta_p(z)}{\alpha_p(z)} = \frac{\beta_p(z)}{(1 - s_1 z^{-1})(1 - s_2 z^{-1}) \dots (1 - s_n z^{-1})}.$$

Write

$$(1 + \dots + z^{-(p-1)})\beta_p(z) \prod_{k=1}^n \left( \sum_{m=1}^{p-1} s_k^m z^{-m} \right) = \sum_{k=1}^{(n+1)p-1} q_k z^{-k}.$$

Then, the transfer function  $G(z)$  at the sampling interval  $T$  is given by

$$G(z) = \frac{\beta(z)}{\alpha(z)} = \frac{q_p z^{-1} + q_{2p} z^{-2} + \dots + q_{np} z^{-n}}{(1 - s_1^p z^{-1})(1 - s_2^p z^{-1}) \dots (1 - s_n^p z^{-1})}. \quad (2.7)$$

Before closing this section, we observe that the parameterization of the Hammerstein model is actually not unique. Suppose the nonlinear block and the linear block are represented by some function  $f$  and the transfer function  $G(z)$  respectively. Then, any pair of  $c \cdot f$  and  $c^{-1}G(z)$  for some non-zero constant  $c$  would produce the identical input-output measurements. In other words, any identification setting can not distinguish between  $(f, G(z))$  and  $(cf, c^{-1}G(z))$ . To obtain a unique parameterization,  $G(z)$  needs to be normalized, e.g., set  $b_1 = 1$ . The problem with this approach is that it indirectly presumes  $b_1 \neq 0$  which may not be the case. To avoid this problem, we assume the following assumption throughout the paper.

**Assumption 2:** Consider  $G(z)$  of (2.2). Assume that  $\|(b_1, b_2, \dots, b_n)\|_2 = 1$  and the sign of the first non-zero element of  $b_i$  is positive.

### 3. Identification of linear block $G(z)$

To avoid unnecessary complications, we assume in Sections 3.1-3.2 that noise  $v(\cdot) = 0$ . The convergence in the presence of noise will be discussed in Section 3.4.

#### 3.1. Estimation of the denominator $\alpha_{n+1}(z)$

Given input sampling interval  $T$ , let the output sampling interval be  $h = T/(n+1)$ . Write  $G_{n+1}(z)$  as

$$G_{n+1}(z) = \frac{\beta_{n+1}(z)}{\alpha_{n+1}(z)} = \frac{\bar{b}_1 z^{-1} + \bar{b}_2 z^{-2} + \dots + \bar{b}_n z^{-n}}{1 - \bar{a}_1 z^{-1} - \bar{a}_2 z^{-2} - \dots - \bar{a}_n z^{-n}}$$

for some unknown  $\bar{b}_i$ 's and  $\bar{a}_j$ 's. Its time domain expression is given by

$$y[kh] = \sum_{i=1}^n \bar{a}_i y[(k-i)h] + \sum_{j=1}^n \bar{b}_j x[(k-j)h].$$

In particular, at  $k = l(n+1)$ ,  $l = 0, 1, 2, \dots$ , or  $kh = l(n+1)\frac{T}{(n+1)} = lT$ , we have

$$y[kh] = y[lT] = \sum_{i=1}^n \bar{a}_i y[lT - i\frac{T}{n+1}] + \sum_{j=1}^n \bar{b}_j x[lT - j\frac{T}{n+1}]$$

$$y[kh - h] = y[lT - h] = \sum_{i=1}^n \bar{a}_i y[lT - (i+1)\frac{T}{n+1}] +$$

$$\sum_{j=1}^n \bar{b}_j x[lT - (j+1)\frac{T}{n+1}].$$

Since the input sampling interval is fixed at  $T$  and

$$\begin{aligned} (l-1)T &= lT - (n+1)\frac{T}{n+1} \leq lT - nT/(n+1) \\ &\leq \dots \leq lT - \frac{T}{n+1} < lT, \end{aligned}$$

it follows that

$$x[(l-1)T] = x[lT - n\frac{T}{(n+1)}] = \dots = x[lT - \frac{T}{(n+1)}].$$

This implies that at  $k = l(n+1)$ ,  $l = 0, 1, 2, \dots$

$$y[lT] - y[lT - h] = \sum_{i=1}^n \bar{a}_i (y[lT - ih] - y[lT - (i+1)h]). \quad (3.1)$$

Define

$$\Delta y[l] = y[lT] - y[lT - h],$$

$$\phi'_1[l] = (y[lT - h] - y[lT - 2h], \dots, y[lT - nh] - y[(l-1)T])$$

$$\bar{a}' = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n).$$

We have

$$\Delta y[l] = \phi'_1[l] \bar{a}. \quad (3.2)$$

This equation is linear in unknown  $\bar{a}$  and thus the denominator coefficients  $\bar{a}$  can be estimated by many standard algorithms, e.g., the recursive LMS or recursive least squares method.

#### 3.2. Estimation of the numerator $\beta_{n+1}(z)$

To estimate the numerator of  $G_{n+1}(z)$  at the sampling interval  $h = T/(n+1)$ , consider two sequences

$$\{y[kh]\} \iff Y_{n+1}(z) = \sum_{k=1}^{\infty} y[kh]z^{-k} = G_{n+1}(z)X_{n+1}(z)$$

$$\begin{aligned} \{y[kh + h/2]\} &\iff \tilde{Y}_{n+1}(z) = \sum_{k=1}^{\infty} y[kh + h/2]z^{-k} \\ &= \tilde{G}_{n+1}(z)X_{n+1}(z) \end{aligned}$$

where  $X_{n+1}(z) = \sum_{k=0}^{\infty} x[kh]z^{-k} = (1 + z^{-1} + \dots + z^{-n})X(z^{n+1})$ . As discussed in (2.2) and (2.6),  $G_{n+1}(z)$  is strictly proper and  $\tilde{G}_{n+1}(z)$  is proper but not strictly proper. Also  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  share the same denominator. Hence,

$$\tilde{G}_{n+1}(z) = \frac{\tilde{b}_0 + \tilde{b}_1 z^{-1} + \dots + \tilde{b}_n z^{-n}}{1 - \bar{a}_1 z^{-1} - \dots - \bar{a}_n z^{-n}} = \frac{\tilde{\beta}_{n+1}(z)}{\alpha_{n+1}(z)}.$$

Clearly

$$\tilde{G}_{n+1}(z)Y_{n+1}(z) - G_{n+1}(z)\tilde{Y}_{n+1}(z) = 0$$

and this implies

$$\tilde{\beta}_{n+1}(z)Y_{n+1}(z) - \beta_{n+1}(z)\tilde{Y}_{n+1}(z) = 0$$

$$(y[kh], \dots, y[kh - nh], -y[kh - h + h/2], \dots, -y[kh - nh + h/2]) \cdot$$

$$(\tilde{b}_0, \dots, \tilde{b}_n, \bar{b}_1, \dots, \bar{b}_n)' = 0, \quad k = 1, 2, \dots$$

Again from (2.6),  $\tilde{G}_{n+1}(z)$  is proper but not strictly proper and this implies  $\tilde{b}_0 \neq 0$ . Define

$$\phi'_2[k] = (-y[kh - h], \dots, -y[kh - nh],$$

$$y[kh - h + h/2], \dots, y[kh - nh + h/2])$$

$$\bar{b}' = (\bar{b}_1, \dots, \bar{b}_n, \bar{b}_1, \dots, \bar{b}_n) / \tilde{b}_0.$$

It follows that

$$y[kh] = \phi'_2[k] \bar{b}. \quad (3.3)$$

This is again linear in the unknown variable  $\bar{b}$  and all other variables  $y[kh]$  and  $\phi'_2[k]$  are available. The unknown numerator coefficients  $\bar{b}$  can be estimated by any standard algorithms.

#### 3.3. Algorithm for estimating $G(z)$

Now, we are in a position to provide the algorithm estimating  $G(z)$  based only on the output measurements. **Blind Identification Algorithm for Estimating  $G(z)$ :**

1. Given input sampling interval  $T$ . Set  $h = T/(n+1)$ .
2. Sample and collect output measurements  $\{y[kh]\}$  and  $\{y[kh + h/2]\}$ . Define  $\Delta y[l]$ ,  $\phi'_1[l]$  and  $\phi'_2[k]$ .

- At each  $k$ , apply either the recursive least squares or the recursive LMS algorithm to estimate  $\bar{b}$  using (3.3). At each  $k = l(n + 1)$ ,  $l = 0, 1, 2, \dots$ , apply either the recursive least squares or the recursive LMS algorithm to estimate  $\bar{a}$  using (3.2). The estimate  $\hat{G}_{n+1}(z)$  of  $G_{n+1}(z)$  is defined as

$$\hat{G}_{n+1}(z) = \frac{\hat{b}_1 z^{-1} + \hat{b}_2 z^{-2} + \dots + \hat{b}_n z^{-n}}{1 - \hat{a}_1 z^{-1} - \hat{a}_2 z^{-2} - \dots - \hat{a}_n z^{-n}}.$$

- Compute the estimate  $\hat{G}(z)$  of  $G(z)$ , which is the transfer function at the sampling interval  $T$ , in terms of its coefficient estimates  $\hat{a} = (\hat{a}_1, \dots, \hat{a}_n)'$  and  $\hat{b} = (\hat{b}_1, \dots, \hat{b}_n)'$  based on  $\hat{G}_{n+1}(z)$  by using equation (2.7). Because  $\|\hat{b}\|_2 = 1$  and the first nonzero element of  $\hat{b}$  is positive, we normalize  $\hat{b}$  by  $\hat{b} = \frac{\hat{b}}{\|\hat{b}\|_2}$  and set  $\hat{b} = -\hat{b}$  if the first nonzero element of  $\hat{b}$  is negative. Finally, the estimate  $\hat{G}(z)$  of  $G(z)$ , at time  $k$ , is obtained.

$$\hat{G}(z) = \frac{\hat{b}_1 z^{-1} + \hat{b}_2 z^{-2} + \dots + \hat{b}_n z^{-n}}{1 - \hat{a}_1 z^{-1} - \hat{a}_2 z^{-2} - \dots - \hat{a}_n z^{-n}}.$$

- Set  $k = k + 1$  and go to Step 2.

Only output measurements are needed to implement the algorithm. The algorithm is recursive and produces the estimate  $\hat{G}(z)$ .

### 3.4. Convergence analysis

Whether  $\hat{G}(z)$  converges to  $G(z)$  depends on whether  $\hat{G}_{n+1}(z)$  converges to  $G_{n+1}(z)$ . Therefore, it boils down to the parameter convergence of  $\hat{a}_i \rightarrow \bar{a}_i$  and  $\hat{b}_j \rightarrow \bar{b}_j$ . It is well known that both parameter estimates converge asymptotically if  $\phi_1[l]$  and  $\phi_2[k]$  are persistently exciting (PE) at least in the absence of noise. In fact, the PE condition on  $\phi_2[k]$  has been developed in our early work on the subject of blind system identification [3]. The PE condition on  $\phi_1[l]$  is however much involved.

**Lemma 3.1** • Suppose the spectral measure of  $\Delta x[l] = x[lT] - x[(l - 1)T]$  is not concentrated on  $m < n$  points. Then,  $\phi_1[l]$  is PE.

- Suppose the numerators of  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  do not share any common factor and the spectral measure of  $x[kT]$  is not concentrated on  $m < 2n$  points. Then,  $\phi_2[k]$  is PE.

Clearly, if noise is absent, the parameter estimates converge to the true values asymptotically. In the presence of noise, the parameter estimation errors converge to a ball centered at the origin with radius  $c_2$  for some constant  $c_2 \geq 0$ , where  $c_2$  relies on the level of the PE and the bound of the noise.

### 3.5. Sufficient richness of $x[kT]$

From the convergence analysis, we see that the parameter convergence depends on the PE conditions of  $\phi_1[l]$  and  $\phi_2[k]$  which rely on the spectral contents of  $x[kT]$  and  $\Delta x[l]$ , often referred to as the sufficient richness condition in the adaptive literature [3, 8]. However,  $x$  is the internal variable which is not measurable and directly controllable. It is desirable to have the richness conditions in terms of the input  $u[kT]$  which we may have control. Translation of the richness condition from  $x$  to  $u$  is actually difficult because of the nonlinearity  $f$ . For instance, let  $u[kT]$  be a Pseudo-Random Binary Noise Sequence (PRBS) taking values  $\pm 1$  and  $x[kT] = u[kT]^4$ .  $u[kT]$  contains infinitely many spectral lines and  $x[kT] \equiv 1$ , containing only one spectral line. Thus,  $u[kT]$  is sufficiently rich but  $x[kT]$  may be not. On the other hand, let  $u[kT] = \sin(0.1k)$  that has two spectral lines and  $x[kT] = u[kT]^4 = [\cos(0.4k) - 4\cos(0.2k) + 3]/8$  which contains 5 spectral lines because of harmonics generated the nonlinearity  $f$ . Therefore, without knowing the exact form of  $f$  and/or  $u[kT]$ , sufficient richness  $u[kT]$  does not necessarily gives rise to a sufficiently rich  $x[kT]$  and vice versa.

However, in the following two practical situations, richness relationship between  $u[kT]$  and  $x[kT]$  can be established. The first case is the i.i.d. random input  $u[kT]$ . Suppose that  $x[kT] = f(u[kT])$  is static and assumes at least two distinctive values with nonzero probability. Then,  $x[kT]$  is also i.i.d. and its autocorrelation function is given by  $R_x(\tau) = \delta(\tau)\sigma^2$ , where  $\sigma^2$  is the variance of  $x[kT]$ . Moreover, the autocorrelation function of  $\Delta x[l] = x[lT] - x[(l - 1)T]$  is in the form of

$$R_{\Delta x}(\tau) = \begin{cases} -\sigma^2 & \tau \pm 1 \\ 2 \cdot \sigma^2 & \tau = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Clearly, both  $x[kT]$  and  $\Delta x[l]$  have infinitely many spectral lines and this implies that  $\phi_1[l]$  and  $\phi_2[k]$  are PE. The second case is the polynomial nonlinearity  $x = \sum_{i=1}^q r_i u^i$  and sinusoidal input  $u[kT] = \sum_{i=1}^m c_i \cos(\Omega_i k)$ . This is the case when input is periodic by the Fourier series representation. If the input has a single spectral line  $u[kT] = e^{j\Omega k}$ ,  $x[kT] = \sum_{i=1}^q r_i e^{j\Omega_i k}$  has  $q$  spectral lines unless some frequencies  $\Omega_i$ 's are the same module  $2\pi$ . If

$$u[kT] = \sum_{i=1}^m 2c_i \cos(\Omega_i k) = \sum_{i=1}^m c_i (e^{j\Omega_i k} + e^{-j\Omega_i k})$$

has  $2m$  spectral lines,  $x[kT]$  has all the frequencies

$$\pm \Omega_{i1} \pm \Omega_{i2} \pm \dots \pm \Omega_{iq}, \quad \Omega_{il} \in [0, \Omega_1, \Omega_2, \dots, \Omega_m]$$

unless in a pathological case where either the coefficients are zeros or the frequencies are the same module  $2\pi$ .

## 4. Identification of the nonlinear block

### 4.1. Direct approach

Once the linear part  $G(z)$  is identified, we have the estimates  $\hat{a}_i$ 's and  $\hat{b}_j$ 's of  $a_i$ ' and  $b_j$ 's. The unknown parameter vector  $\gamma$  that parameterizes the nonlinear block can be estimated directly by minimizing

$$\hat{\gamma} = \arg \min_{\tilde{\gamma}} \sum_k (y[kT] - \sum_{i=1}^n \hat{a}_i y[kT - iT]) - \sum_{i=1}^n \hat{b}_i f(u[kT - iT], \dots, u[kT - iT - mT], \tilde{\gamma})^2. \quad (4.1)$$

The convergence and computational complexity of the minimization depends, of course, on the nonlinearity  $f$ . Here, we are particularly interested in the linear parameterization structure

$$f(u[kT], \dots, u[kT - mT], \gamma) = \sum_{i=1}^l \gamma_i f_i(u[kT], \dots, u[kT - mT]) \quad (4.2)$$

with known  $f_i$ 's and unknown  $\gamma_i$ 's. Then, by defining

$$\phi'_3[k] = \left( \sum_{i=1}^n \hat{b}_i f_i(u[kT - iT], \dots, u[kT - iT - mT]), \dots, \sum_{i=1}^n \hat{b}_i f_i(u[kT - iT], \dots, u[kT - iT - mT]) \right)$$

$$\xi[k] = y[kT] - \sum_{i=1}^n \hat{a}_i y[kT - iT].$$

Equation (4.1) can be re-written as

$$\hat{\gamma} = \arg \min_{\tilde{\gamma}} \sum_k (\xi[k] - \phi'_3[k] \tilde{\gamma})^2 \quad (4.3)$$

All variables  $\phi_3[k]$  and  $\xi[k]$  are available and  $\hat{\gamma}$  can be estimated by many standard algorithms, e.g., the recursive LMS algorithm

$$\hat{\gamma}[k] = \hat{\gamma}[k-1] + \frac{\phi_3[k]}{1 + \phi'_3[k] \phi_3[k]} (\xi[k] - \phi'_3[k] \hat{\gamma}[k-1]).$$

We remark that the common polynomial nonlinearity

$$x[kT] = \sum_{i=1}^l \gamma_i u^i[kT]$$

in the Hammerstein model representation is a special case of (4.2) with  $f_j(u[kT], \dots, u[kT - mT]) = u[kT]^j$ .

### 4.2. Indirect approach

In this approach, our goal is to recover the unknown internal signal  $x[kT]$  first and then to estimate the nonlinear block using the information of  $u[kT]$  and  $x[kT]$ . This approach is particularly useful when the nonlinear block is static  $x[kT] = f(u[kT])$  but lacks of structure. Because of unknown structure, it is not possible to estimate the nonlinear function  $f$  in terms of parameter estimation. However, if the data  $u[kT]$  and  $x[kT]$  become available, the complete picture of  $f$  can be easily graphed. This graphical picture provides us accurate information on the unknown  $f$  as long as there is enough pair  $(u[kT], x[kT])$  in the range of interests.

The first step of this indirect approach is to recover unknown  $x[kT]$ . This can be done by recovering  $x[kT]$  or  $x[kh]$  because  $x$  is held constant between  $kT \leq t < (k+1)T$ . To easy notation, suppose the transfer functions  $G(z)$ ,  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  are known. If not, their estimates  $\hat{G}(z)$ ,  $\hat{G}_{n+1}(z)$  and  $\hat{\tilde{G}}_{n+1}(z)$  can be obtained by applying the Blind Identification Algorithm presented in the previous section.

Now, recall

$$Y(z) = G(z)X(z), \quad Y_{n+1}(z) = G_{n+1}(z)X_{n+1}(z),$$

$$\tilde{Y}_{n+1}(z) = \tilde{G}_{n+1}(z)X_{n+1}(z)$$

where  $Y(z)$ ,  $Y_{n+1}(z)$ ,  $\tilde{Y}_{n+1}(z)$  and  $X_{n+1}(z)$  are Z-transforms of  $\{y[kT]\}$ ,  $\{y[kh]\}$ ,  $\{y[kh + h/2]\}$  and  $\{x[kh]\}$  respectively. Also note all  $y[kh]$ ,  $y[kh + h/2]$  and  $y[kT]$  are available. Suppose one of  $G(z)$ ,  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  is minimum phase,  $x[kT]$  or  $x[kh]$  can be recovered by taking the inverse

$$X(z) = G^{-1}(z)Y(z), \text{ or } X_{n+1}(z) = (G_{n+1}(z))^{-1}Y_{n+1}(z),$$

$$X_{n+1}(z) = (\tilde{G}_{n+1}(z))^{-1}\tilde{Y}_{n+1}(z)$$

or in time domain

$$x[kT] = \frac{1}{b_1} (-b_2 x[(k-1)T] - \dots - b_n x[(k-n)T])$$

$$+ y[(k+1)T] - a_1 y[kT] - \dots - a_n y[(k-i-n)T])$$

$$x[kh] = \frac{1}{b_1} (-\bar{b}_2 x[(k-1)h] - \dots - \bar{b}_n x[(k-n)h])$$

$$+ y[(k+1)h] - \bar{a}_1 y[kh] - \dots - \bar{a}_n y[(k-i-n)h])$$

$$x[kh] = \frac{1}{\bar{b}_0} (-\bar{b}_1 x[(k-1)h] - \dots - \bar{b}_n x[(k-n)h])$$

$$+ y[kh + h/2] - \bar{a}_1 y[(k-1)h + h/2] - \dots - \bar{a}_n y[(k-n)h + h/2])$$

If, however,  $G(z)$ ,  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  are all non-minimum phase, inversion becomes problematic. To this end, suppose  $G_{n+1}(z)$  and  $\tilde{G}_{n+1}(z)$  do not share any common zeros. Then, from the Bezout identity, there exist two stable transfer functions  $F(z)$  and  $\tilde{F}(z)$  such that

$$F(z)\tilde{G}_{n+1}(z) + \tilde{F}(z)G_{n+1}(z) = 1. \quad (4.4)$$

This implies

$$F(z)\hat{Y}_{n+1}(z) + \tilde{F}(z)Y_{n+1}(z) = \quad (4.5)$$

$$(F(z)\tilde{G}_{n+1}(z) + \tilde{F}(z)G_{n+1}(z))X_{n+1}(z) = X_{n+1}(z).$$

Therefore,  $x[kh]$  and consequently  $x[kT]$  can be obtained by filtering  $y[kh]$  and  $y[kh + h/2]$  using  $\tilde{F}(z)$  and  $F(z)$ . Note that calculations of  $F(z)$  and  $\tilde{F}(z)$  are straightforward if  $G(z)$  and  $\tilde{G}(z)$  are available.

Once  $x[kT]$  is obtained, the nonlinear block  $f$  can be estimated by using the information of  $u[kT]$  and  $x[kT]$ . We consider two cases:

1. The nonlinear function  $f$  is static  $x[kT] = f(u[kT])$  and non-parametric. In this case, the function  $f$  can be graphed using pairs of  $u[kT]$ 's and the estimated  $x[kT]$ 's. From the graph, the nonlinear function  $f$  can be estimated and approximated. We remark that to extract useful informations about  $f$  from the graph, enough pairs of  $(u[kT], x[kT])$  in the range of interests are preferable. Clearly, the Pseudo-Random Binary Noise Sequence (PRBS) which can generate only two pairs is certainly not a good choice. This is a well known fact in the literature.
2. The nonlinear function is in the general form of (2.1) and then  $\gamma$  can be estimated by minimizing

$$\hat{\gamma} = \arg \min_{\tilde{\gamma}} \sum_k (x[kT] - f(u[kT], \dots, u[kT-mT], \tilde{\gamma}))^2. \quad (4.6)$$

In particular, if  $f$  is linear in the unknown  $\gamma$ ,

$$f(u[kT], \dots, u[kT-mT], \gamma) = \sum_{i=1}^l \gamma_i f_i$$

with known  $f_i$ 's and unknown  $\gamma_i$ 's, then, by defining

$$\phi'_4[k] = (f_1, \dots, f_l).$$

we have

$$\hat{\gamma} = \arg \min_{\tilde{\gamma}} \sum_k (x[kT] - \phi'_4[k]\tilde{\gamma})^2 \quad (4.7)$$

and  $\hat{\gamma}$  can be estimated by many standard algorithms.

$$\hat{\gamma}[k] = \hat{\gamma}[k-1] + \frac{\phi_4[k]}{1 + \phi'_4[k]\phi_4[k]} (x[kT] - \phi'_3[k]\hat{\gamma}[k-1]).$$

## 5. Concluding remarks

In this paper, we have proposed blind approaches for Hammerstein model identification. By sampling faster at the output or holding the input over a period of time, the

linear part may be obtained using only the output measurements. Convergence results in terms of PE conditions are also derived which apply to a large class of signals. Our focus in this paper is to present the idea and therefore, not much effort was devoted to study the performance of the proposed algorithms under various type of model uncertainties and noises. This issue is certainly an interesting one for further study and we expect that results will be quite different from the traditional linear system identification because of noise structure in the error equation. It is also interesting to characterize the sufficient richness condition for some specific nonlinear functions and inputs.

## References

- [1] Bai, E.W. (2001), "Identification of systems with hard input nonlinearities", in Perspectives in Control, R. Moheimani(ed), Springer Verlag, 2001
- [2] Bai, E.W. (1998), An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems, *Automatica*, **34**, No. 3. pp.333-338.
- [3] Bai, E.W. and M. Fu (1999), "Blind system identification and channel equalization of IIR systems without statistical information", *Proc of IEEE CDC*, 1996 also in *IEEE Trans. on Signal Processing*, **47**, No.7, pp.1-13.
- [4] Bai, E.W. and M. Fu (2001), "A blind approach to Hammerstein model identification", Tech Report, Dept of ECE, Univ. of Iowa
- [5] S.A. Bilings and S.Y. Fakhouri (1978), Identification of a class of nonlinear systems using correlation analysis, *Proc. of IEE* **125**, No. 7. pp.691-697.
- [6] Greblicki, W. (2000), "Continuous time Hammerstein system identification", *IEEE Trans. on Auto. Contr.*, **45**, pp.1232-1236
- [7] Haber, R. and H. Unbehauen (1990), "Structure identification of nonlinear dynamic systems-a survey of input/output approaches", *Automatica*, **24**, pp.651-677
- [8] L. Ljung (1987) *System Identification*, Englewood Cliffs, NJ: Prentice -Hall,
- [9] Narendra, K.S. and P.G. Gallman (1966), "An iterative method for the identification of nonlinear systems using a Hammerstein model", *IEEE Trans. on Auto. Contr.*, **11**, pp.546-550
- [10] Stoica, P. (1981), "On the convergence of an iterative algorithm used for Hammerstein system identification", *IEEE Trans. on Auto. Contr.*, **26**, pp.967-969
- [11] Sun, L., W. Liu and A. Sano (1998), "Identification of dynamical system with input nonlinearity", *IEE Proc-Control Theory Appl*, **146**, No.1, pp.41-51