

Parameter Identification of Hammerstein Models using Elimination Theory

Kaiyu Wang, Marc Bodson, John Chiasson, and Leon M. Tolbert

Abstract—A Hammerstein model is a system model in which the inputs go through a static nonlinearity followed by a linear time-invariant system. Often the static nonlinearity is modeled as a polynomial nonlinearity in the inputs or as a piecewise constant nonlinearity. Such models are nonlinear in the unknown parameters and therefore present a challenging identification problem. In this work, the authors show that elimination theory can be used to solve exactly for parameter values that minimize a least-square criterion. Thus, the approach guarantees the minimum can be found in a finite number of steps, unlike iterative methods that are currently used.

Index Terms—Hammerstein models, Parameter identification, Nonlinear least-squares, Resultants.

I. INTRODUCTION

Methods to identify the parameters of Hammerstein models have been considered for many years; for example, see Refs. [1][2][3][4] for some of the early work in the controls literature. In the last few years, there has been a renewed interest in the area, as seen in the recent publications [5][6][7][8][9][10][11][12]. With the exception of [3], these all involve iterative methods to solve for the parameters. The noniterative method in [3] (see also [13] pp. 584-585) is achieved by rewriting the system model so that it is linear in the parameters, but overparameterized. In this work, we present a noniterative method that does not overparameterize the system model and guarantees that the parameter vector that minimizes the least-squares error criterion is found. The method exploits the fact that the squared error is rational in the unknown parameters. The optimal parameter vector is found using the method of resultants from elimination theory [14]. The approach here is based on previous work by the authors on identifying the parameters of an induction motor [15].

II. HAMMERSTEIN MODEL

A Hammerstein model consists of the input going through a static nonlinear block followed by a linear time-invariant system, as shown in Figure 1 (see [7][13]). Mathematically,

K. Wang, J. Chiasson and L. M. Tolbert are with the ECE Department, The University of Tennessee, Knoxville, TN 37996. *wkaiyu@utk.edu*, *chiasson@utk.edu*, *tolbert@utk.edu*

M. Bodson is with the EE Department, The University of Utah, Salt Lake City, UT 84112. *bodson@ece.utah.edu*

L. M. Tolbert is also with Oak Ridge National Laboratory, NTRC, 2360 Cherahala Boulevard, Knoxville, TN 37932. *tolbertlm@ornl.gov*

Drs. Chiasson and Tolbert would like to thank Oak Ridge National Laboratory for partially supporting this work through the UT/Battelle contract no. 4000023754. Dr. Tolbert would also like to thank the National Science Foundation for partially supporting this work through contract NSF ECS-0093884.

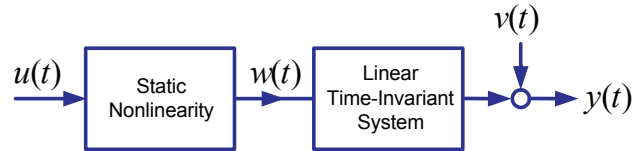


Fig. 1. Hammerstein model.

a Hammerstein system may be represented as [7]

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=1}^n b_i w(t-i) + v(t) \quad (1)$$

$$w(t) = g(u(t)) \quad (2)$$

where $g(\cdot)$ is a static nonlinearity and $v(t)$ is (typically) a discrete-time random process of independent, identically distributed (*i.i.d.*) random variables. The static nonlinearity $g(\cdot)$ is often taken to have a polynomial form, that is,

$$g(u) = \sum_{i=1}^m c_i u^i. \quad (3)$$

The unknowns are then

$$a = (a_1, \dots, a_n)$$

$$b = (b_1, \dots, b_n)$$

$$c = (c_1, \dots, c_m)$$

where upper bounds on n and m are assumed to be known. Substituting the polynomial nonlinearity (3) into (2), the Hammerstein model (1) and (2) may be rewritten as

$$\begin{aligned} y(t) &= - \sum_{i=1}^n a_i y(t-i) + \sum_{i=1}^n \sum_{j=1}^m b_i c_j u^j(t-i) + v(t) \\ &= - [a_1 \quad \dots \quad a_n] \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-n) \end{bmatrix} \\ &\quad + [b_1 \quad \dots \quad b_n] \begin{bmatrix} u(t-1) & \dots & u^m(t-1) \\ \vdots & \ddots & \vdots \\ u(t-n) & \dots & u^m(t-n) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \\ &\quad + v(t). \end{aligned}$$

This model is linear in the unknown parameters a_i , but nonlinear in the unknown parameters b_i, c_i . It can be made to “look” linear in the parameters, as in [3] and [13], by

letting

$$k_i = \begin{cases} -a_i & \text{for } i = 1, \dots, n \\ b_1 c_{i-n} & \text{for } i = n+1, \dots, n+m \\ \vdots & \\ b_n c_{i-n-(n-1)m} & \text{for } i = n+(n-1)m+1, \\ & \dots, n+nm \end{cases}$$

$$W_i^T(t) = \begin{cases} y(t-i) & \text{for } i = 1, \dots, n \\ u^{i-n}(t-1) & \text{for } i = n+1, \dots, n+m \\ \vdots & \\ u^{i-n-(n-1)m}(t-1) & \text{for } i = n+(n-1)m \\ & +1, \dots, n+nm \end{cases}$$

so that

$$y(t) = W(t)K.$$

However, in this representation, the k_i 's for $i > n$ are not independent of each other so that the system is overparameterized.

To simplify the notation for the remainder of the paper, consider the specific model

$$y(t) = \sum_{i=1}^2 b_i w(t-i) + v(t) \quad (4)$$

$$w(t) = g(u(t)) = c_1 u(t) + c_2 u^2(t) \quad (5)$$

or

$$y(t) = \sum_{i=1}^2 \sum_{j=1}^2 b_i c_j u^j(t-i) + v(t)$$

$$= \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} u(t-1) & u^2(t-1) \\ u(t-2) & u^2(t-2) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + v(t).$$

Therefore

$$y(t) = b_1 c_1 u(t-1) + b_1 c_2 u^2(t-1) + b_2 c_1 u(t-2) + b_2 c_2 u^2(t-2) + v(t) \quad (6)$$

or, as in [3][13], write

$$y(t) = k_1 u(t-1) + k_2 u^2(t-1) + k_3 u(t-2) + k_4 u^2(t-2) + v(t) \quad (7)$$

where

$$k_1 = b_1 c_1, k_2 = b_1 c_2, k_3 = b_2 c_1, k_4 = b_2 c_2. \quad (8)$$

The model (7) is overparameterized as

$$k_4 = k_2 k_3 / k_1. \quad (9)$$

In the next section, it is shown how to solve this identification problem explicitly taking into account the constraint (9).

III. LEAST-SQUARES IDENTIFICATION

Let

$$W(t) \triangleq [u(t-1) \quad u^2(t-1) \quad u(t-2) \quad u^2(t-2)]$$

and

$$K \triangleq \begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix}$$

so that equation (7) is written compactly as

$$y(t) = W(t)K + v(t). \quad (10)$$

Collecting the data set $\{y(t), u(t)\}$ for $t = 1, \dots, N$, the least-squares criterion requires that the parameter vector K be found that minimizes the squared error

$$E^2(K) = \sum_{t=1}^N |y(t) - W(t)K|^2 \quad (11)$$

subject to the constraint (9) and assuming $k_1 \neq 0$. With

$$R_{WW} \triangleq \sum_{t=1}^N W^T(t)W(t)$$

$$R_{Wy} \triangleq \sum_{t=1}^N W^T(t)y(t)$$

$$R_y \triangleq \sum_{t=1}^N y^T(t)y(t),$$

the squared error may be rewritten

$$E^2(K) = \sum_{t=1}^N |y(t) - W(t)K|^2$$

$$= R_y - 2R_{Wy}^T K + K^T R_{WW} K. \quad (12)$$

For the moment, assume that $k_1 = b_1 c_1$ is not zero and the problem is then to find $K_P \triangleq [k_1 \quad k_2 \quad k_3]$ that minimizes

$$E^2(K_P) \triangleq \sum_{t=1}^N \left| y(t) - W(t)K \right|_{k_4=k_2 k_3 / k_1}^2$$

$$= R_y - 2R_{Wy}^T K \Big|_{k_4=k_2 k_3 / k_1} + (K^T R_{WW} K) \Big|_{k_4=k_2 k_3 / k_1}. \quad (13)$$

The standard procedure is then to compute the extrema points, that is, the solutions to

$$r_1(K_P) \triangleq \frac{\partial E^2(K_P)}{\partial k_1} = 0$$

$$r_2(K_P) \triangleq \frac{\partial E^2(K_P)}{\partial k_2} = 0 \quad (14)$$

$$r_3(K_P) \triangleq \frac{\partial E^2(K_P)}{\partial k_3} = 0.$$

These are rational functions of k_1, k_2 , and k_3 . The solutions to this set of rational functions are contained in the solution set of

$$\begin{aligned} p_1(K_P) &\triangleq k_1^{\ell_1} \frac{\partial E^2(K_P)}{\partial k_1} = 0 \\ p_2(K_P) &\triangleq k_1^{\ell_2} \frac{\partial E^2(K_P)}{\partial k_2} = 0 \\ p_3(K_P) &\triangleq k_1^{\ell_3} \frac{\partial E^2(K_P)}{\partial k_3} = 0 \end{aligned} \quad (15)$$

where ℓ_1, ℓ_2 , and ℓ_3 are the *smallest* positive integers that make the p_i polynomials in the k_1 . As $k_1 \neq 0$ is assumed, the zero sets of (14) and (15) are identical. Systems of polynomial equations can be solved for all of their roots using resultants. This is summarized in the next subsection.

A. Solving Systems of Polynomial Equations via Resultants [14][16]

Given two polynomial equations $a(k_1, k_2) = 0$ and $b(k_1, k_2) = 0$, how does one solve them simultaneously to eliminate (say) k_2 ? A systematic procedure to do this is known as *elimination theory* and uses the notion of *resultants*. Briefly, one considers $a(k_1, k_2)$ and $b(k_1, k_2)$ as polynomials in k_2 whose coefficients are polynomials in k_1 . Then, for example, letting $a(k_1, k_2)$ and $b(k_1, k_2)$ have degrees 3 and 2, respectively in k_2 , they may be written in the form

$$\begin{aligned} a(k_1, k_2) &= a_3(k_1)k_2^3 + a_2(k_1)k_2^2 + a_1(k_1)k_2 \\ &\quad + a_0(k_1) \\ b(k_1, k_2) &= b_2(k_1)k_2^2 + b_1(k_1)k_2 + b_0(k_1). \end{aligned}$$

The $n \times n$ Sylvester matrix, where $n = \deg_{k_2} \{a(k_1, k_2)\} + \deg_{k_2} \{b(k_1, k_2)\} = 3 + 2 = 5$, is defined by

$$S_{a,b}(k_1) \triangleq \begin{bmatrix} a_0(k_1) & 0 & b_0(k_1) & 0 & 0 \\ a_1(k_1) & a_0(k_1) & b_1(k_1) & b_0(k_1) & 0 \\ a_2(k_1) & a_1(k_1) & b_2(k_1) & b_1(k_1) & b_0(k_1) \\ a_3(k_1) & a_2(k_1) & 0 & b_2(k_1) & b_1(k_1) \\ 0 & a_3(k_1) & 0 & 0 & b_2(k_1) \end{bmatrix}. \quad (16)$$

The *resultant polynomial* is then defined by

$$r(k_1) = \text{Res} \left(a(k_1, k_2), b(k_1, k_2), k_2 \right) \triangleq \det S_{a,b}(k_1) \quad (17)$$

and is the result of *eliminating* the variable k_2 from $a(k_1, k_2)$ and $b(k_1, k_2)$. In fact, the following is true.

Theorem 1: [14][16] Any solution (k_1^0, k_2^0) of $a(k_1, k_2) = 0$ and $b(k_1, k_2) = 0$ must satisfy $r(k_1^0) = 0$.

See the Appendix for more discussion on resultants.

B. Solving the Polynomial System (15)

Consider the system (15) to be polynomials in k_3 with coefficients polynomials in (k_1, k_2) . Compute

$$\begin{aligned} r_{p_1 p_2}(k_1, k_2) &\triangleq \text{Res} \left(p_1, p_2, k_3 \right) = \det S_{p_1 p_2}(k_1, k_2) \\ r_{p_1 p_3}(k_1, k_2) &\triangleq \text{Res} \left(p_1, p_3, k_3 \right) = \det S_{p_1 p_3}(k_1, k_2). \end{aligned}$$

to eliminate k_3 . Next, k_2 is eliminated to obtain

$$r(k_1) \triangleq \text{Res} \left(r_{p_1 p_2}, r_{p_1 p_3}, k_2 \right) = \det S_{r_{p_1 p_2} r_{p_1 p_3}}(k_1)$$

which is a polynomial in one variable.

The roots k_{1i} for $i = 1, \dots, n_1$ of $r(k_1) = 0$ are found, which are then substituted into $r_{p_1 p_2}$. For each fixed k_{1i} , solve $r_{p_1 p_2}(k_{1i}, k_2) = 0$ (or $r_{p_1 p_3}(k_{1i}, k_2) = 0$) for its roots k_{2ij} $j = 1, \dots, n_{2i}$ to obtain the partial solutions (k_{1i}, k_{2ij}) for $i = 1, \dots, n_1, j = 1, \dots, n_{2i}$. For each i and j , the partial solutions (k_{1i}, k_{2ij}) are then substituted into p_1 (or into p_2 or p_3) and $p_1(k_{1i}, k_{2ij}, k_3) = 0$ is solved to obtain the solutions $(k_{1i}, k_{2ij}, k_{3ijk})$ for $i = 1, \dots, n_1, j = 1, \dots, n_{2i}, k = 1, \dots, n_{3ij}$. These solutions are then checked to see which ones satisfy the complete system of polynomial equations (15) and, therefore, constitute the *candidate* solutions for the minimization. From the set of candidate solutions, the one that results in the smallest squared error is chosen. These computations presented in this work were carried out using MATHEMATICA [17].

Remark

The assumption that $k_1 \neq 0$ was made above. However, one does not need to know this *a priori*. Simply, one solves the problem twice: once assuming that $k_1 \neq 0$ and the second time assuming $k_4 \neq 0$ [substituting $k_1 = k_2 k_3 / k_4$ into (12)]. A parameter vector K_P^* is found in each case. From these two solutions, the one that gives the smallest residual (squared) error is chosen.

C. Sufficient Richness

The system of polynomial equations (15) can be solved using resultants as outlined in the previous subsection and one may then check each of the finite number of solutions for the one that gives the minimal value for $E^2(K_P)$. However, one needs to know if the solution makes sense. For example, in the *linear* least-squares problem, there is a unique, well-defined solution provided that the regressor matrix R_W is nonsingular (or in practical terms, if its condition number is not too large). In the nonlinear case here, a Taylor series expansion about the computed minimum point $K_P^* = [k_1^*, k_2^*, k_3^*]^T$ is done to obtain $(i, j = 1, 2, 3)$

$$\begin{aligned} E^2(K_P) &= E^2(K_P^*) \\ &\quad + \frac{1}{2} [K_P - K_P^*]^T \frac{\partial^2 E^2(K_P^*)}{\partial k_i \partial k_j} [K_P - K_P^*] + \dots \end{aligned} \quad (18)$$

One then checks that the Hessian matrix with elements $\frac{\partial^2 E^2(K_P^*)}{\partial k_i \partial k_j}$ is positive definite and that its condition number is not too large, to ensure that the data was sufficiently rich to identify the parameters.

D. Solving for the Model Parameters

Once the k_i are found, the b_i, c_i are found. If $k_1 \neq 0$ is the minimizing case, then the system (see (8))

$$\begin{aligned} k_1 &= b_1 c_1 \\ k_2 &= b_1 c_2 \\ k_3 &= b_2 c_1 \end{aligned} \quad (19)$$

is solved using resultants while if $k_4 \neq 0$ is the minimizing case, then

$$\begin{aligned} k_2 &= b_1 c_2 \\ k_3 &= b_2 c_1 \\ k_4 &= b_2 c_2 \end{aligned} \quad (20)$$

is solved via resultants.

IV. SIMULATION EXAMPLES

Here we consider the use of the above methodology to estimate the parameter values of some examples presented in the literature.

A. Example 1

The first example is from [4][8]

$$y(t) = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} u(t-1) & u^2(t-1) \\ u(t-2) & u^2(t-2) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + v(t)$$

where $b_1 = 1, b_2 = -2, c_1 = 0, c_2 = 2, u(t)$ is zero mean Gaussian white noise with unity variance, and $v(t)$ is also zero mean Gaussian white noise with variance 1.56.

Note that $k_1 = b_1 c_1 = 0$. However, one would not know this a priori so this example is worked out assuming first that $k_1 \neq 0$, and then assuming $k_4 \neq 0$.

Case 1 $k_1 \neq 0$

A simulation is performed and the input-output data is collected from $t = 1, \dots, 1000$ to compute R_W, R_{W_y} , and R_y . Using equations (13), (14) and (15), the polynomials $[K_P \triangleq (k_1, k_2, k_3)]$

$$\begin{aligned} p_1(K_P) &\triangleq k_1^3 \frac{\partial E^2(K_P)}{\partial k_1} = 0 \\ p_2(K_P) &\triangleq k_1^2 \frac{\partial E^2(K_P)}{\partial k_2} = 0 \\ p_3(K_P) &\triangleq k_1^2 \frac{\partial E^2(K_P)}{\partial k_3} = 0 \end{aligned} \quad (21)$$

are computed, where

	deg k_1	deg k_2	deg k_3
p_1	4	2	2
p_2	3	1	2
p_3	3	2	1

Using resultants, $r_{p_1 p_2} \triangleq \text{Res}(p_1, p_2, k_2)$ and $r_{p_1 p_3} \triangleq \text{Res}(p_1, p_3, k_2)$ are computed where, it happens that $r_{p_1 p_2} = k_1^2 \bar{r}_{p_1 p_2}$ and $r_{p_1 p_3} = k_1^2 \bar{r}_{p_1 p_3}$. The factors k_1^2 in $r_{p_1 p_2}$ and k_1^2 in $r_{p_1 p_3}$ can be removed without changing the solution set as k_1 is assumed to be nonzero. The degrees of $\bar{r}_{p_1 p_2}$ and $\bar{r}_{p_1 p_3}$ are

	deg k_1	deg k_3
$\bar{r}_{p_1 p_2}$	5	5
$\bar{r}_{p_1 p_3}$	6	6

Finally, k_3 is eliminated from $\bar{r}_{p_1 p_2}$ and $\bar{r}_{p_1 p_3}$ to obtain

$$r(k_1) \triangleq \text{Res}(\bar{r}_{p_1 p_2}, \bar{r}_{p_1 p_3}, k_3) = \det S_{\bar{r}_{p_1 p_2} \bar{r}_{p_1 p_3}}(k_1)$$

where $\deg\{r(k_1)\} = 18$. The optimal solution is $K_{P_1}^* = (0.036, 2.008, -0.072)$ so that $k_4 = k_2 k_3 / k_1 = -4.016$ and the error is computed to be $E^2(K_{P_1}^*) = 1931.6$.

For this optimal value of $K_{P_1}^*$ and choosing the norm constraint $\|b\| = \sqrt{b_1^2 + b_2^2} = \sqrt{5}$ as in [8], the solutions to $k_1^* = b_1 c_1, k_2^* = b_1 c_2$ and $k_3^* = b_2 c_1$ are

$$(b_1, b_2, c_1, c_2) = \begin{cases} (0.999, -2.001, 0.036, 2.009) \\ (-0.999, 2.001, -0.036, -2.009) \end{cases}$$

Zero measurement noise. If this problem is redone with no measurement noise, then the optimal solution is $K_{P_1}^* = (0.234, 1.975, -0.470)$ so that $k_4 = k_2 k_3 / k_1 = -3.967$ and the error is computed to $E^2(K_{P_1}^*) = 338.19$. For this optimal value of $K_{P_1}^*$ and again choosing $\|b\| = \sqrt{b_1^2 + b_2^2} = \sqrt{5}$, the solutions to $k_1^* = b_1 c_1, k_2^* = b_1 c_2$ and $k_3^* = b_2 c_1$ are

$$(b_1, b_2, c_1, c_2) = \begin{cases} (0.994, -2.001, 0.235, 1.987) \\ (-0.994, 2.001, -0.235, -1.987) \end{cases}$$

Case 2. $k_4 \neq 0$

Using the same data as in case 1, equation (13) is used with $k_1 = k_2 k_3 / k_4$ and $K_P \triangleq (k_2, k_3, k_4)$. The corresponding polynomial system whose zeros are the extremal points of the least-squares error are of the form

$$\begin{aligned} p_1(K_P) &\triangleq k_4^2 \frac{\partial E^2(K_P)}{\partial k_2} = 0 \\ p_2(K_P) &\triangleq k_4^2 \frac{\partial E^2(K_P)}{\partial k_3} = 0 \\ p_3(K_P) &\triangleq k_4^3 \frac{\partial E^2(K_P)}{\partial k_4} = 0 \end{aligned} \quad (22)$$

where

	deg k_2	deg k_3	deg k_4
p_1	1	2	3
p_2	2	1	3
p_3	2	2	4

Using resultants, $r_{p_1 p_2} \triangleq \text{Res}(p_1, p_2, k_2)$ and $r_{p_1 p_3} \triangleq \text{Res}(p_1, p_3, k_2)$ are computed, where it happens that $r_{p_1 p_2} = k_4^2 \bar{r}_{p_1 p_2}$ and $r_{p_1 p_3} = k_4^2 \bar{r}_{p_1 p_3}$, where again, the factors k_4^2 in $r_{p_1 p_2}$ and k_4^2 in $r_{p_1 p_3}$ can be removed without changing the solution set as k_4 is assumed to be nonzero. The degrees of $\bar{r}_{p_1 p_2}$ and $\bar{r}_{p_1 p_3}$ are

	deg k_3	deg k_4
$\bar{r}_{p_1 p_2}$	5	5
$\bar{r}_{p_1 p_3}$	6	6

Finally, k_3 is eliminated to obtain

$$r(k_4) \triangleq \text{Res}(\bar{r}_{p_1 p_2}, \bar{r}_{p_1 p_3}, k_3) = \det S_{\bar{r}_{p_1 p_2} \bar{r}_{p_1 p_3}}(k_4)$$

where $\deg\{r(k_4)\} = 18$. The optimal solution is $K_{P_2}^* = (2.008, -0.072, -4.020)$ so that $k_1 = k_2 k_3 / k_4 = 0.036$ and the error is computed to $E^2(K_{P_2}^*) = 1931.6$. For this optimal value of $K_{P_2}^*$ and choosing the norm constraint $\|b\| = \sqrt{b_1^2 + b_2^2} = \sqrt{5}$ as in [8], the solutions to $k_2^* = b_1 c_2, k_3^* = b_2 c_1$ and $k_4^* = b_2 c_2$ are

$$(b_1, b_2, c_1, c_2) = \begin{cases} (0.999, -2.001, 0.036, 2.010) \\ (-0.999, 2.001, -0.036, -2.010) \end{cases}$$

The errors as well as the estimated parameters b_i and c_i are the same in both cases which is not a surprise as neither of the estimated values for k_1 and k_4 are zero

Zero measurement noise. The optimal solution is $K_{P2}^* = (2.00, 0.00, -4.00)$ so that $k_1 = k_2 k_3 / k_4 = 0.00$ and the error is computed to be $E^2(K_{P2}^*) = 0$. For this optimal value of K_{P2}^* and again choosing $\|b\| = \sqrt{b_1^2 + b_2^2} = \sqrt{5}$, the solutions to $k_2^* = b_1 c_2$, $k_3^* = b_2 c_1$ and $k_4^* = b_2 c_2$ are

$$(b_1, b_2, c_1, c_2) = \begin{cases} (1.00, -2.00, 0.00, 2.00) \\ (-1.00, 2.00, 0.00, -2.00) \end{cases}$$

With zero measurement noise, $E^2(K_{P2}^*) = 0$ as expected and the parameter values estimated with $k_4 \neq 0$ are chosen.

B. Example 2

The second example is from [5] where the discrete-time system to be identified is

$$\begin{aligned} y(t) + a_1 y(t - T) + a_2 y(t - 2T) &= b_1 w(t - T) \\ &+ b_2 w(t - 2T) + v(t) \\ w(t) = g(u(t)) &\triangleq c_1 u + c_2 u^2 \end{aligned}$$

where the sampling time is $T = 0.6$ sec, $u(t)$ is an *i.i.d.* random process uniformly distributed in the interval $[-5, 5]$ and $v(t)$ is an *i.i.d.* random process uniformly distributed in the interval $[-0.1, 0.1]$. The parameter values in the simulation are

$$\begin{aligned} a &= (a_1, a_2) = (-1.8287, 0.8353) \\ b &= (b_1, b_2) = (0.9397, -0.3420) \\ c &= (c_1, c_2) = (1, 1). \end{aligned}$$

Data is collected from the simulation for $t = T, 2T, \dots, 1000T$. Set

$$\begin{aligned} k_1 &= -a_1 \\ k_2 &= -a_2 \\ k_3 &= b_1 c_1 \\ k_4 &= b_1 c_2 \\ k_5 &= b_2 c_1 \\ k_6 &= k_4 k_5 / k_3 \end{aligned}$$

where $k_3 \neq 0$ is assumed. The corresponding polynomial system whose zeros are the extremal points of the least-squares error are of the form $[K_P \triangleq (k_1, k_2, k_3, k_4, k_5)]$

$$\begin{aligned} p_1(K_P) &\triangleq k_3 \frac{\partial E^2(K_P)}{\partial k_1} = 0 \\ p_2(K_P) &\triangleq k_3 \frac{\partial E^2(K_P)}{\partial k_2} = 0 \\ p_3(K_P) &\triangleq k_3^3 \frac{\partial E^2(K_P)}{\partial k_3} = 0 \\ p_4(K_P) &\triangleq k_3^2 \frac{\partial E^2(K_P)}{\partial k_4} = 0 \\ p_5(K_P) &\triangleq k_3^2 \frac{\partial E^2(K_P)}{\partial k_5} = 0 \end{aligned}$$

where

	deg k_1	deg k_2	deg k_3	deg k_4	deg k_5
p_1	1	1	2	1	1
p_2	1	1	2	1	1
p_3	1	1	4	2	2
p_4	1	1	3	1	2
p_5	1	1	3	2	1

Next, k_2 is eliminated to obtain four polynomials in four unknowns as

$$\begin{aligned} q_1(k_1, k_3, k_4, k_5) &\triangleq \text{Res}(p_1, p_2, k_2) \\ q_2(k_1, k_3, k_4, k_5) &\triangleq \text{Res}(p_3, p_2, k_2) \\ q_3(k_1, k_3, k_4, k_5) &\triangleq \text{Res}(p_4, p_2, k_2) \\ q_4(k_1, k_3, k_4, k_5) &\triangleq \text{Res}(p_5, p_2, k_2) \end{aligned}$$

where $q_1 = k_3 \bar{q}_1, q_2 = k_3 \bar{q}_2, q_3 = k_3 \bar{q}_3, q_4 = k_3 \bar{q}_4$ and

	deg k_1	deg k_3	deg k_4	deg k_5
\bar{q}_1	1	2	1	1
\bar{q}_2	1	4	2	2
\bar{q}_3	1	3	1	2
\bar{q}_4	1	3	2	1

Next, k_1 is eliminated to obtain three polynomials in three unknowns as

$$\begin{aligned} s_1(k_3, k_4, k_5) &\triangleq \text{Res}(\bar{q}_1, \bar{q}_2, k_1) \\ s_2(k_3, k_4, k_5) &\triangleq \text{Res}(\bar{q}_1, \bar{q}_3, k_1) \\ s_3(k_3, k_4, k_5) &\triangleq \text{Res}(\bar{q}_1, \bar{q}_4, k_1) \end{aligned}$$

where $s_1 = k_3 \bar{s}_1, s_2 = k_3 \bar{s}_2, s_3 = k_3 \bar{s}_3$ and

	deg k_3	deg k_4	deg k_5
\bar{s}_1	4	2	2
\bar{s}_2	3	1	2
\bar{s}_3	3	2	1

Next, k_5 is eliminated to obtain two polynomials in two unknowns as

$$\begin{aligned} t_1(k_3, k_4) &\triangleq \text{Res}(\bar{s}_1, \bar{s}_2, k_5) \\ t_2(k_3, k_4) &\triangleq \text{Res}(\bar{s}_1, \bar{s}_3, k_5) \end{aligned}$$

where $t_1 = k_3^2 \bar{t}_1, t_2 = k_3^2 \bar{t}_2$ and

	deg k_3	deg k_4
\bar{t}_1	6	6
\bar{t}_2	5	5

Finally, k_4 is eliminated to obtain a single polynomial in k_3 as

$$r(k_3) = \text{Res}(\bar{t}_1(k_3, k_4), \bar{t}_2(k_3, k_4), k_4)$$

where

$$\deg_{k_3} \{r(k_3)\} = 18.$$

The calculated optimal solution is

$$K_P^* = (1.8288, -0.8354, 0.9394, 0.9397, -0.3417)$$

so that $k_6 = k_4 k_5 / k_3 = -0.3418$ and the error is computed to be $E^2(K_P^*) = 0$. For this optimal value of K_P^* and

choosing $\|b\| = \sqrt{b_1^2 + b_2^2} = 1$, the solutions to $k_2^* = b_1 c_2, k_3^* = b_2 c_1$ and $k_4^* = b_2 c_2$ are

$$(a_1, a_2, b_1, b_2, c_1, c_2) = \begin{cases} (1.8288, -0.8354, -0.9398, 0.3418, -0.9996, -0.9999) \\ (-1.8288, 0.8354, 0.9398, -0.3418, 0.9996, 0.9999) \end{cases}$$

V. CONCLUSIONS AND FUTURE RESEARCH

A procedure to identify the parameters in a Hammerstein model was presented and illustrated with examples. The method guarantees that the parameters that minimize the squared error is found. A limitation of the method is that models with larger number of parameters result in much higher degree polynomials to be solved using resultant theory. In such cases, the symbolic computation of the determinant of the Sylvester matrix for finding the resultant polynomials becomes an issue. New results in [18][19] for doing such computations will allow one to work problems with larger number of parameters.

REFERENCES

- [1] K. S. Narendra and P. G. Gallman, "An iterative method for the identification of nonlinear systems using a hammerstein model," *IEEE Transactions on Automatic Control*, vol. 11, pp. 546–550, July 1966.
- [2] P. G. Gallman, "A comparison of two hammerstein model identification algorithms," *IEEE Transactions on Automatic Control*, vol. 21, pp. 124–126, February 1976.
- [3] F. H. I. Chang and R. Luus, "A noniterative method for identification using hammerstein model," *IEEE Transactions on Automatic Control*, vol. 16, pp. 464–468, October 1971.
- [4] P. Stoica, "On the convergence of an iterative algorithm used for hammerstein identification," *IEEE Transactions on Automatic Control*, vol. 26, pp. 967–969, April 1981.
- [5] E.-W. Bai and M. Fu, "A blind approach to hammerstein model identification," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1610–1619, July 2002.
- [6] E.-W. Bai, "Frequency domain identification of hammerstein models," *IEEE Transactions on Automatic Control*, vol. 48, pp. 530–542, April 2003.
- [7] E.-W. Bai and D. Li, "Convergence of the iterative hammerstein system identification algorithm," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1929–1940, November 2004.
- [8] E.-W. Bai and D. Li, "Convergence of the iterative hammerstein system identification algorithm," in *Proceeding of the 43rd IEEE Conference on Decision and Control*, pp. 3868–3873, Paradise Island, Bahamas, December 2004.
- [9] W. Greblicki, "Continuous-time hammerstein system identification," *IEEE Transactions on Automatic Control*, vol. 45, pp. 1232–1236, June 2000.
- [10] Z. Hasiiewicz and G. Mzyk, "Combined parametric-nonparametric identification of hammerstein systems," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1370–1375, August 2004.
- [11] J. Vörös, "Iterative algorithm for parameter identification of hammerstein systems with two-segment nonlinearities," *IEEE Transactions on Automatic Control*, vol. 44, pp. 2145–2149, November 1999.
- [12] J. Vörös, "Modeling and parameter identification of systems with multisegment piecewise-linear characteristics," *IEEE Transactions on Automatic Control*, vol. 47, pp. 184–188, January 2002.
- [13] O. Nelles, *Nonlinear System Identification*. Springer-Verlag, Berlin, 2001.
- [14] D. Cox, J. Little, and D. O'Shea, *IDEALS, VARIETIES, AND ALGORITHMS An Introduction to Computational Algebraic Geometry and Commutative Algebra*. 2nd Edition, Springer-Verlag, Berlin, 1996.
- [15] K. Wang, J. Chiasson, M. Bodson, and L. M. Tolbert, "A nonlinear least-squares approach for estimation of the induction motor parameters," in *Proceedings of the IEEE Conference on Decision and Control*, pp. 3856–3861, December 2004.
- [16] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*. Cambridge University Press, Cambridge, UK, 1999.

- [17] S. Wolfram, *Mathematica, A System for Doing Mathematics by Computer*. 2nd Edition, Addison-Wesley, Reading, MA, 1992.
- [18] M. Hromcik and M. Šebek, "New algorithm for polynomial matrix determinant based on FFT," in *Proceedings of the European Conference on Control ECC'99*, August 1999, Karlsruhe Germany.
- [19] M. Hromcik and M. Šebek, "Numerical and symbolic computation of polynomial matrix determinant," in *Proceedings of the 1999 Conference on Decision and Control*, pp. 1887–1888, 1999, Tampa FL.
- [20] D. Cox, J. Little, and D. O'Shea, *Using Algebraic Geometry*. Springer-Verlag, Berlin, 1998.

A. Appendix: Resultants (see [14][16][20]).

Given two polynomials $a(x_1, x_2)$ and $b(x_1, x_2)$ how does one find their common zeros? To see how this is done, write

$$\begin{aligned} a(x_1, x_2) &= a_3(x_1)x_2^3 + a_2(x_1)x_2^2 + a_1(x_1)x_2 + a_0(x_1) \\ b(x_1, x_2) &= b_2(x_1)x_2^2 + b_1(x_1)x_2 + b_0(x_1). \end{aligned}$$

Next, see if polynomials of the form

$$\begin{aligned} \alpha(x_1, x_2) &= \alpha_1(x_1)x_2 + \alpha_0(x_1) \\ \beta(x_1, x_2) &= \beta_2(x_1)x_2^2 + \beta_1(x_1)x_2 + \beta_0(x_1). \end{aligned}$$

can be found such that

$$\alpha(x_1, x_2)a(x_1, x_2) + \beta(x_1, x_2)b(x_1, x_2) = r(x_1). \quad (23)$$

Equating powers of x_2 , this equation may be rewritten in matrix form as

$$\begin{bmatrix} a_0 & 0 & b_0 & 0 & 0 \\ a_1 & a_0 & b_1 & b_0 & 0 \\ a_2 & a_1 & b_2 & b_1 & b_0 \\ a_3 & a_2 & 0 & b_2 & b_1 \\ 0 & a_3 & 0 & 0 & b_2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} r(x_1) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The matrix on the left-hand side is called the *Sylvester* matrix and is denoted here by $S_{a,b}(x_1)$. The inverse of $S_{a,b}(x_1)$ has the form

$$S_{a,b}^{-1}(x_1) = \frac{1}{\det S_{a,b}(x_1)} \text{adj} \left(S_{a,b}(x_1) \right)$$

where $\text{adj}(S_{a,b}(x_1))$ is the adjoint matrix and is a 5×5 polynomial matrix in x_1 . Solving for $\alpha_i(x_1), \beta_i(x_1)$ gives

$$\begin{bmatrix} \alpha_0(x_1) \\ \alpha_1(x_1) \\ \beta_0(x_1) \\ \beta_1(x_1) \\ \beta_2(x_1) \end{bmatrix} = \frac{\text{adj} S_{a,b}(x_1)}{\det S_{a,b}(x_1)} \begin{bmatrix} r(x_1) \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Choosing $r(x_1) = \det S_{a,b}(x_1)$ this becomes

$$\begin{bmatrix} \alpha_0(x_1) \\ \alpha_1(x_1) \\ \beta_0(x_1) \\ \beta_1(x_1) \\ \beta_2(x_1) \end{bmatrix} = \text{adj} S_{a,b}(x_1) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and guarantees that $\alpha_0(x_1), \alpha_1(x_1), \beta_0(x_1), \beta_1(x_1), \beta_2(x_1)$ are polynomials in x_1 . That is, the *resultant polynomial* defined by $r(x_1) = \det S_{a,b}(x_1)$ is the polynomial required for (23).