

A Correlation Least-Squares Method for Hammerstein Model Identification with ARX and μ -Markov Structures

Kai-Yew Lum*, Dennis S. Bernstein**

* Temasek Laboratories, National University of Singapore,
Singapore 117411 (e-mail: kaiyew.lum@nus.edu.sg).

** The University of Michigan, MI 48109, USA
(e-mail: dsbaero@umich.edu.)

Abstract: This paper presents a two-step method for identification of the SISO Hammerstein model, which employs input autocorrelation and input-output cross-correlation functions as data for least-squares estimation. Using separable processes as input signals, the proposed method allows the linear block of a Hammerstein model to be identified up to a multiplicative constant, without *a priori* knowledge of the nonlinear model structure. Both ARX and μ -Markov structures of the linear block are considered, where the main concern is the accuracy of pole and zero estimates. The correlation least-squares method is compared numerically with a well-known nonlinear least-squares method, which shows that the correlation method is consistently accurate across different nonlinear model structures.

Keywords: System identification, nonlinear models, Markov parameters, ARMA models, least-squares estimation.

1. INTRODUCTION

The Hammerstein model is a special nonlinear model consisting of a static nonlinear block followed by a linear dynamic block. Like other block-oriented models such as the Wiener and Hammerstein-Wiener models, the Hammerstein model has been widely employed as an approximation for nonlinear dynamics in engineering problems (Billings and Fakhouri [1979], Baldelli et al. [2005], Zeng et al. [2008], Lum and Lai [2010]). Thus, methods for identification of Hammerstein (and Wiener) models have been widely studied, among which iterative (Narendra and Gallman [1966]), frequency-domain (Crama and Schoukens [2004]), nonlinear parametric (Bai [1998]), and non-iterative (Zeng et al. [2008]) methods, correlation analysis (Billings and Fakhouri [1978a, 1979]), and methods based on adaptive model refinement (Palanthandalam-Madapusi et al. [2005], D'Amato et al. [2010]).

One of the main purposes of Hammerstein model identification is correct estimation of the linear block, as this is critical for linear or linear parameter-varying analyses (Lum and Lai [2010], Zeng et al. [2008]). In most approaches including some of the above-mentioned, parametrization of the linear block in ARX model structure depends on knowledge of the true system order, whereas direct estimation of Markov parameters requires truncation by finite impulse-response (FIR) approximation. Meanwhile, the μ -Markov model structure is an over-parametrized ARX model, and provides a general parametrization that contains an arbitrary number of Markov parameters, but retains the appropriate infinite impulse-response (IIR) model structure (Van Pelt and Bernstein [1998], Venugopal and Bernstein [2000]). In par-

ticular, it has been shown that least-squares identification of the μ -Markov model yields consistent estimates under unknown system order, and in the presence of input and measurement noises (Fledderjohn et al. [2010]).

The advantage of the μ -Markov model structure can be exploited for the Hammerstein model if the identification method is of the two-step kind that identifies the linear block separately from the nonlinear block. One such approach is the two-step nonlinear least-squares algorithm of Bai [1998]. A disadvantage of this method is that a wrong assumption of the form of the nonlinear function can directly affect parameter estimates of the linear block.

The purpose of this paper is to present an alternative two-step method that is based on correlation analysis with least-squares estimates (Billings and Fakhouri [1978a], Isermann et al. [1974]). For a class of so-called *separable* input signals, Nuttall [1958] showed that the input-output cross-correlation function of a memoryless device is proportional to the autocorrelation function of the input. Thus, the correlation least-squares method with separable input allows the linear block of a Hammerstein model to be identified up to a multiplicative constant, without *a priori* knowledge of the nonlinear model structure; the constant depends on the input process and the nonlinear function (Billings and Fakhouri [1978b]). Mismodeling of the latter will only lead to a wrong estimate of the static gain.

In this paper, the correlation least-squares method will be compared, in numerical examples, with the nonlinear least-squares method (Bai [1998]) and the standard least-squares method, using both μ -Markov and ARX model structures. Our main concern is the accuracy of pole and zero estimates of the SISO Hammerstein model. In the μ -

Markov technique, the poles and zeros are obtained from the estimated Markov parameters using the eigensystem realization algorithm (Juang [1994]). For simplicity we shall assume the model orders and relative degrees are known. The investigation will focus on the effects of noise, unknown nonlinear model structure, and input signals. Two different separable signals are considered: zero-mean Gaussian white noise, and phase-modulated sinusoid or chirp.

The paper is organized as follows: Section 2 defines the Hammerstein model with ARX and μ -Markov structures; Section 3 presents the correlation-based and nonlinear least-squares methods; Section 4 presents the numerical results; concluding remarks are given in Section 5.

2. MODEL STRUCTURES

We consider the scalar Hammerstein model composed of a scalar, static input nonlinearity $F[\cdot]$ followed by a rational linear SISO block represented by its continuous-time impulse response function $h(\tau)$ (see Figure 1). We shall consider two discrete-time models for the linear block:

2.1 ARX Model

The ARX model is given by

$$y(k) = -\sum_{j=1}^n a_j y(k-j) + \sum_{i=d}^n b_i u(k-i), \quad (1)$$

where n is the model order, $d \geq 0$ is the relative degree, and $b_d = h_d$ is the first nonzero Markov parameter. Note that the Markov parameters h_k are obtained by the impulse response of (1), and are related to $h(\tau)$ by $h_k = d^k h(t)/dt^k|_{t=0}$ (Chen [1984]). Equation (1) with zero initial conditions can be equivalently represented as the transfer function

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_d z^{n-d} + \dots + b_n z}{z^n + a_1 z^{n-1} + \dots + a_n}. \quad (2)$$

2.2 μ -Markov Model

The μ -Markov model structure is the μ -step ahead predictor given by:

$$y(k) = -\sum_{j=1}^n \alpha_j y(k-\mu+1-j) + \sum_{i=d}^{d+\mu-1} h_i u(k-i) + \sum_{i=d+1}^n \beta_i u(k-\mu+1-i), \quad (3)$$

where $\mu \geq 1$ (Astrom and Wittemark [1995]). Note that (3) is an over-parametrized ARX model with μ Markov parameters $h_d, \dots, h_{d+\mu-1}$ explicitly included. In the case $\mu = 1$, equation (3) specializes to the usual ARX model (1) with $\alpha_j = a_j$ and $\beta_i = b_i$. In the general case, an equivalence can be established between (3) and (1) by calculating the coefficients α_j and β_i recursively from a_j , b_i and h_k (Venugopal and Bernstein [2000]). Equation (3) with zero initial conditions can be equivalently represented by the *ARMARKOV* transfer function

$$G(z) = \frac{B(z)}{A(z)} = z^{-d} \frac{H_d z^{n+\mu-1} + \dots + H_{d+\mu-1} z^n + \beta_{d+1} z^{n-1} + \dots + \beta_n z^d}{z^{n+\mu-1} + \alpha_1 z^{n-1} + \dots + \alpha_n}. \quad (4)$$

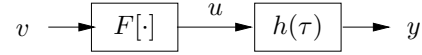


Fig. 1. Hammerstein model

3. IDENTIFICATION METHODS

We are interested in identification of the linear block in both the ARX and μ -Markov structures, particularly the poles and zeros. First, we propose a correlation least-squares method. Next we summarize a well-known nonlinear least-squares method with which the proposed method will be compared.

3.1 Correlation Least Squares

Separable Input Process.

Definition 1. A real valued random process $v : \mathbb{R} \mapsto \mathbb{R}$ is said to be *separable* if

$$\mathcal{E}\{v(t-\tau)v(t)\} = c(\tau)v(t), \quad \forall \tau, t \in \mathbb{R}, \quad (5)$$

with $c(\tau) = \phi_{vv}(\tau)/\phi_{vv}(0)$, where $\phi_{vv}(\tau)$ is the autocorrelation function of v (Enqvist and Ljung [2005]).

It has been shown that Gaussian, sine-wave, and phase-modulated sine-wave processes are separable. Furthermore, for a wide class of nonlinear functions $u = F[v]$, the input-output cross-correlation function ϕ_{uv} is related to the input auto-correlation function ϕ_{uu} by

$$\phi_{uv}(\tau) = C_F(v)\phi_{vv}(\tau), \quad C_F(v) = \frac{1}{\phi_{vv}(0)} \int_{\mathbb{R}} vF[v] \cdot p(v)dv, \quad (6)$$

provided that v is separable (Nuttall [1958]). In (6), $p(v)$ is the probability distribution function of the input process v , and $C_F(v)$ is a constant depending only on $F[\cdot]$ and the statistics of v (Billings and Fakhouri [1978b]).

With (6), one can also establish a relationship between the input and output correlation functions of the Hammerstein model. Indeed, since $h(\tau)$ is linear, one has

$$\phi_{yw}(\tau) = h(\tau) * \phi_{uw}(\tau)$$

for any signal w , where $*$ denotes the convolution operator. In particular, letting $w = v$ and using (6), one obtains the following lemma.

Lemma 2. Given a separable input process v , the cross-correlation function of v and the output y of the scalar Hammerstein model is proportional to the input auto-correlation function passed through $h(\tau)$:

$$\phi_{yv}(\tau) = C_F(v)h(\tau) * \phi_{vv}(\tau). \quad (7)$$

Calculation of the constant $C_F(v)$ for a given process v is generally cumbersome. However, by treating the pair (ϕ_{vv}, ϕ_{yv}) as input and output signals, (7) implies that these are related by a linear system whose 'effective' impulse response is $\hat{h}(\tau) := C_F(v)h(\tau)$. This observation gives rise to the following two-step identification method.

Correlation Least Squares (CLS) Identification Method

Step 1: Identification of $\hat{h}(\tau) = C_F(v)h(\tau)$.

Using a separable process as input signal, identify $\hat{h}(\tau)$ in ARX or μ -Markov representations, respectively, by standard least-squares estimation but using the correlation functions $(\phi_{vv}(\tau), \phi_{yv}(\tau))$ as “input-output” data. The respective model equations thus become

$$\begin{aligned}\phi_{yv}(k) &= -\sum_{j=1}^n a_j \phi_{yv}(k-j) + \sum_{i=d}^n b_i \phi_{vv}(k-i), \text{ or} \\ \phi_{yv}(k) &= -\sum_{j=1}^n \alpha_j \phi_{yv}(k-\mu+1-j) + \sum_{i=d}^{d+\mu-1} h_i \phi_{vv}(k-i) \\ &\quad + \sum_{i=d+1}^n \beta_i \phi_{vv}(k-\mu+1-i).\end{aligned}$$

Step 2: Correcting for $C_F(v)$.

Assume a p -order polynomial approximation of $F[\cdot]$:

$$F[v] = \gamma_1 v + \dots + \gamma_p v^p,$$

and estimate the γ_j 's according to the equation

$$\hat{A}(q^{-1})y(k) = \hat{B}(q^{-1})[\gamma_1 v(k) + \dots + \gamma_p v^p(k)] + e(k),$$

where q^{-1} is the delay operator, and $\hat{G}(z) = \hat{B}(z)/\hat{A}(z)$ is the ARX or ARMARKOV transfer function of $\hat{h}(\tau)$ similar to (2) or (4), respectively, depending on which model structure is employed. Stack the data points in rows to yield a regression equation $Z = \Phi\theta + E$ where $\theta = (\gamma_1, \dots, \gamma_p)^T$, and compute a least-squares estimate $\hat{\theta}$ of the latter. Finally, set $C_F(v) = 1/\hat{\gamma}_1$, and obtain a model of $h(\tau)$ by $B(z) = \hat{\gamma}_1 \hat{B}(z)$.

Remark 1. Step 1 of the above correlation least-squares method was applied to identification of linear systems in Isermann et al. [1974]. Presently, it follows from (7) that the poles and zeros of the linear block can be identified in Step 1 without *a priori* knowledge of the nonlinear function $F[\cdot]$. An assumption on the form of $F[\cdot]$ is only needed in Step 2 in order to estimate the static gain.

Remark 2. Application of CLS requires sampled-data approximation of the correlation functions:

$$\phi_{vv}(\tau) \approx \frac{1}{N+1} \sum_{k=0}^N v(k)v(k-\tau), \quad (8a)$$

$$\phi_{yv}(\tau) \approx \frac{1}{N+1} \sum_{k=0}^N y(k)v(k-\tau). \quad (8b)$$

To obtain correlation values for $\tau = -\tau_{\max}, \dots, 0, \dots, \tau_{\max}$, a total of $N + 2\tau_{\max} + 1$ data points in $v(k)$ and $y(k)$ are needed. Hence, CLS requires more data points than standard LS. As will be illustrated later in a numerical example, N must be sufficiently large for the identification results to be accurate.

3.2 Nonlinear Least Squares

A two-stage identification algorithm was given by Bai [1998] for multi-input multi-output Hammerstein-Wiener systems. For our purpose, this is specialized to the scalar Hammerstein model. We shall summarize the method

using the ARX model structure while that for the μ -Markov model follows similarly. The method assumes the following parametric model of the nonlinearity $F[\cdot]$:

$$u = F[v(k)] = \sum_{l=1}^p \gamma_l f_l[v(k)], \quad (9)$$

where the f_l 's are *a priori known smooth* functions, and the coefficients γ_l are to be estimated. Substituting into (1) yields the following nonlinear ARX equation:

$$y(k) = -\sum_{j=1}^n a_j y(k-j) + \sum_{i=d}^n b_i \left\{ \sum_{l=1}^p \gamma_l f_l[v(k-i)] \right\}. \quad (10)$$

One can write $y(k) = \phi^T(k)\theta$, where

$$\begin{aligned}\theta &:= (a_1, \dots, a_n, b_d \gamma_1, \dots, b_d \gamma_p, \dots, b_n \gamma_1, b_n \gamma_p)^T, \\ \phi(k) &:= (y(k-1), \dots, y(k-n), f_1[v(k-d)], \dots, \\ &\quad f_p[v(k-d)], \dots, f_1[v(k-n)], \dots, f_p[v(k-n)])^T.\end{aligned}$$

Let
$$\Theta_{b\gamma} := b\gamma^T = \begin{bmatrix} b_d \\ \vdots \\ b_n \end{bmatrix} [\gamma_1 \dots \gamma_p].$$

Two-Stage Nonlinear Least Squares (NLS) Method (Bai [1998]) Assume that $\Theta_{b\gamma}$ is non-zero, and $\|b\|_2 = 1$, then proceed as follows:

Stage 1: Estimate θ , which contains the a_j 's and the lumped parameters $b_i \gamma_l$, using standard least squares.

Stage 2: Construct $\hat{\Theta}_{b\gamma}$ using the above estimates, and compute its singular-value decomposition:

$$\hat{\Theta}_{b\gamma} = \sum_{i=1}^{\min(n,p)} \sigma_i U_i V_i^T, \quad U_i \in \mathbb{R}^n, V_i \in \mathbb{R}^p$$

where U_i and V_i are orthonormal, and the singular values are ordered: $\sigma_1 \geq \dots \geq \sigma_{\min(n,p)}$. Let s_U denote the sign of the first nonzero element of U_1 , and calculate the estimates of b and γ as:

$$\hat{b} = s_U U_1, \quad \hat{\gamma} = s_U \sigma_1 V_1. \quad (11)$$

Remark 3. For monomials $f_l[v(k)] = v^l(k)$, instead of $\|b\|_2 = 1$, it is more reasonable to assume $\gamma_1 = 1$. Thus, one has the following instead of (11):

$$\hat{b} = U_1/\sigma_1 V_{11}, \quad \hat{\gamma} = V_1/V_{11}, \quad (12)$$

where V_{11} is the first element of V_1 .

3.3 Standard Least Squares

In the standard least-squares method, one simply ignores the static nonlinearity $F[\cdot]$ and identifies the ARX and μ -Markov models respectively using the input-output pairs $(v(k), y(k))$ to form the regressor.

4. NUMERICAL EXAMPLES

In the following numerical examples, we shall compare the ability of the three identification methods in estimating the poles and zeros of the linear block with both ARX and μ -Markov model structures. The purpose is to examine the consistency of each method and model for different nonlinear functions. We shall assume that the model order n and relative degree d are known. The number of Markov parameters to be estimated is $\mu = 30$.

4.1 Model Setup

Linear Block. Two models of the linear block are considered, defined by their stable transfer functions:

$$G_1(z) = \frac{(z - 1.25)(z - 0.5)}{(z - 0.95)(z^2 - 0.49z + 0.3895)}$$

$$G_2(z) = \frac{(z - 1.25)(z + 0.2)(z^2 - 0.28z + 0.9605)}{(z - 0.25)(z^2 - 0.5z + 0.3989)(z^2 + 0.64z + 0.9124)}$$

Both models have static gain 1, relative degree 1 and a non-minimum phase zero at 1.25. Besides having a higher order than $G_1(z)$, the model $G_2(z)$ also has a pole (0.25) and a zero (-0.2) closer to the origin, simulating a relatively slow sampling rate. The pole-zero maps of the two models are shown in Fig. 2.

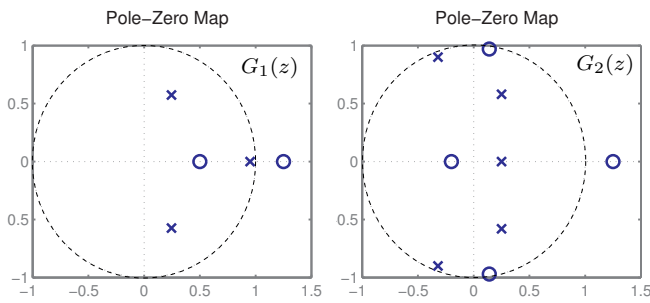


Fig. 2. Pole-zero map of $G_1(z)$ and $G_2(z)$

Nonlinear Functions. For both the CLS and NLS methods, we shall assume a monic 3rd-order polynomial:

$$F[v] = v + \gamma_2 v^2 + \gamma_3 v^3.$$

Meanwhile, in the simulations, we consider three odd nonlinear functions:

3rd-order polynomial: $F_1[v] = v + 4v^3;$

Smooth saturation: $F_2[v] = \frac{1}{8} \tanh(8v);$

High-order function: $F_3[v] = v + 0.1 \sin(4\pi v).$

Note that F_1 is perfectly matched by the nonlinear model assumed by CLS and NLS, but not F_2 and F_3 . Moreover, F_2 and F_3 have the following Taylor-series expansions:

$$F_2[v] = v - 21.33v^3 + 546.13v^5 - \dots,$$

$$F_3[v] = 2.26v - 33.07v^3 + 261.14v^5 - \dots$$

Hence, they cannot be approximated by a 3rd-order polynomial except for small input values of u . Moreover, since F_3 is not monic, estimation of the static gain will also be inaccurate. Plots of F_1 , F_2 and F_3 for the input range considered (see below) are given in Fig. 3.

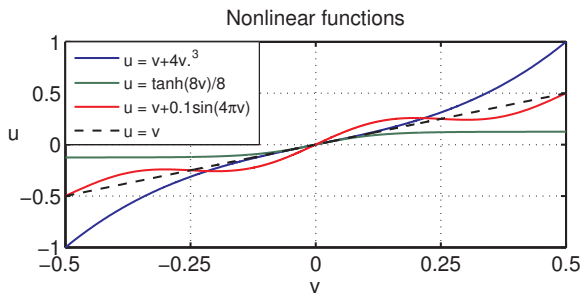


Fig. 3. Nonlinear functions

Input Signals. Tests are conducted using two different separable input processes. The total number of data points for each test is 5001. The first input signal is the discrete-time simulation of a Gaussian white noise with zero mean and standard deviation $1/2\pi$, so that 99% of the input values fall within ± 0.48 (3-sigma point).

The second separable input process employed is a phase-modulated sinusoid given by:

$$v(t) = 0.5 \sin(\omega_0 t + \Delta_\omega / \omega_s \sin(\omega_s t)), \quad (13)$$

with $\omega_0 \sim 20$ Hz, $\Delta_\omega \sim 19.5$ Hz, and $\omega_s \sim 0.02$ Hz. In practice this is a chirp signal with the central frequency ω_0 , demi-frequency band Δ_ω and sweep frequency ω_s . Fig. 4 shows the power spectral density of the input signal (13).

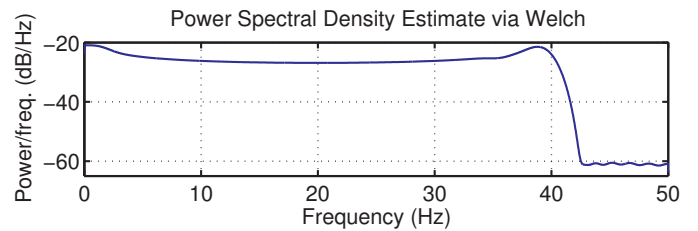


Fig. 4. Power spectral density of chirp input signal

4.2 Presentation of Results

The following sections present the estimated poles, zeros and gains of both the ARX and μ -Markov models¹ for different combinations of linear block, nonlinear function and input signal, with or without input and measurement noises. The tables also list the corresponding root-sum-squares (RSS) errors; i.e. for the poles (zeros) (z_1, \dots, z_n) and their estimates $(\hat{z}_1, \dots, \hat{z}_n)$, the RSS error is given by

$$e_{\text{rss}} = \sqrt{|z_1 - \hat{z}_1|^2 + \dots + |z_n - \hat{z}_n|^2}$$

after having matched each z_i with the closest \hat{z}_i .

4.3 $G_1(z)$ with Gaussian white-noise input

Tables 1–3 give the results for the linear block $G_1(z)$ with Gaussian white noise input. No input and output noises are added in these simulations.

First, in Table 1, it can be seen that for F_1 which is perfectly matched by the nonlinear model assumed in NLS, this latter method unsurprisingly yields perfect pole and zero estimates with both the ARX and μ -Markov models. The ARX poles and zeros identified by CLS are accurate up to the second decimal point, whereas those identified by standard LS are not acceptable.

With the μ -Markov model, both CLS and LS yield poles, zeros and gains that are accurate up to the third decimal point. However, the standard LS gain estimate of 1.309 is wrong. Fig. 5 shows that CLS and NLS both yield accurate estimates of the Markov parameters.

For the other two nonlinear functions, F_2 and F_3 (Tables 2–3), CLS yields consistently accurate poles and zeros with the ARX model, whereas the NLS and standard LS estimates are unacceptable. The μ -Markov model yields better estimates for all three methods.

¹ Note that the poles and zeros of the identified μ -Markov model are obtained via the eigensystem realization algorithm (Juang [1994]).

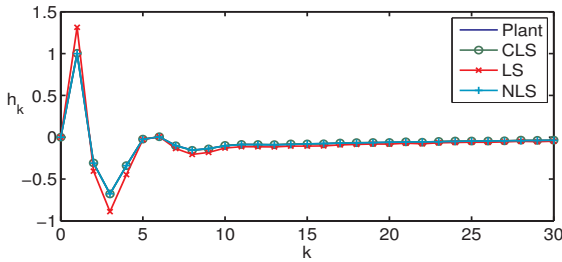


Fig. 5. Plant and estimated Markov parameters, $G_1(z)$

The above tests are repeated by adding uncorrelated white noises to both the input and output signals at noise-to-signal ratios of 4% to 6%. Table 4 gives the result for F_2 . The pole and zero errors are summarized in Fig. 6, where one observes only slight deterioration in each test compared with the noiseless case. In summary:

- For each identification method, the μ -Markov model yields more accurate estimates than the ARX model.
- Overall, CLS is the most accurate when the ARX model is used, and yields consistently accurate poles and zeros with the μ -Markov model.
- NLS is accurate when a matching model of the nonlinear function is assumed (in this case, F_1). However, it deteriorates with unmatched functions (F_2 and F_3).

Note that, in both noiseless and noisy cases, the gains estimated by CLS and NLS are not accurate for F_2 and F_3 because of the mismatch in the nonlinear block.

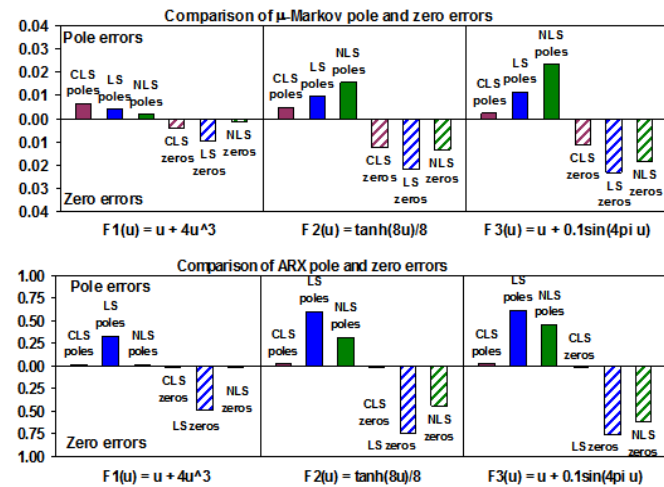


Fig. 6. Comparison of pole & zero errors, $G_1(z)$, noisy case.

4.4 Influence of correlation approximation on CLS

For the CLS method, Fig. 7 examines the RSS pole, zero and Markov parameter errors for the case $G_1(z)$ with nonlinear function F_1 , as functions of the data length N used for correlation function approximation (see equation (8)). It can be seen that estimation accuracy is consistent with increasing N . However, the ARX-model pole and zero errors are an order of magnitude larger than those of the μ -Markov model, and only improve significantly for $N > 4700$ out of 5001 data points.

From the enlarged view on the right for $4900 < N < 5000$, it can be seen that the μ -Markov errors deteriorate beyond

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.239 \pm 0.576i$	0.477	$0.249 \pm 0.573i$	0.500
0.950	1.250	0.939	1.237	0.949	1.255
RSS errors:		0.0143	0.0269	0.0059	0.0045
Gain:		1.007		0.999	
Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.107 \pm 0.501i$	0.014	$0.244 \pm 0.576i$	0.491
0.950	1.250	0.703	1.201	0.948	1.253
RSS errors:		0.3309	0.4882	0.0039	0.0097
Gain:		1.322		1.309	
Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.245 \pm 0.574i$	0.500	$0.245 \pm 0.574i$	0.500
0.950	1.25	0.950	1.250	0.950	1.250
RSS errors:		0.0000	0.0000	0.0000	0.0000
Gain:		1.000		1.000	

Table 1. $G_1(z)$, $F_1[v] = v + 4v^3$, Gaussian white-noise input, noiseless data

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.242 \pm 0.566i$	0.483	$0.241 \pm 0.577i$	0.505
0.950	1.250	0.928	1.262	0.950	1.242
RSS errors:		0.0253	0.0214	0.0064	0.0094
Gain:		0.699		0.709	
Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.051 \pm 0.380i$	-0.239	$0.249 \pm 0.569i$	0.520
0.950	1.250	0.513	1.164	0.952	1.244
RSS errors:		0.5840	0.7438	0.0091	0.0211
Gain:		0.513		0.515	
Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.132 \pm 0.514i$	0.090	$0.253 \pm 0.578i$	0.509
0.950	1.250	0.725	1.220	0.944	1.252
RSS errors:		0.2890	0.4110	0.0143	0.0094
Gain:		0.703		0.703	

Table 2. $G_1(z)$, $F_2[v] = \tanh(8v)/8$, Gaussian white-noise input, noiseless data

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.245 \pm 0.565i$	0.492	$0.243 \pm 0.575i$	0.507
0.950	1.250	0.931	1.258	0.950	1.243
RSS errors:		0.0233	0.0110	0.0030	0.0098
Gain:		1.521		1.536	
Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.051 \pm 0.370i$	-0.253	$0.251 \pm 0.568i$	0.522
0.950	1.250	0.494	1.159	0.951	1.245
RSS errors:		0.6056	0.7585	0.0114	0.0229
Gain:		1.164		1.166	
Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
$0.245 \pm 0.574i$	0.500	$0.076 \pm 0.445i$	-0.112	$0.258 \pm 0.579i$	0.516
0.950	1.250	0.610	1.208	0.938	1.250
RSS errors:		0.4542	0.6133	0.0230	0.0160
Gain:		1.518		1.522	

Table 3. $G_1(z)$, $F_3[v] = v + 0.1 \sin(4\pi v)$, Gaussian white-noise input, noiseless data

$N > 4970$. This is because, for $N > 4970$, the number of correlation coefficients available is (recall Remark 2)

$$2\tau_{\max} + 1 = 5001 - N < 31,$$

which is insufficient for identification of $\mu=30$ Markov parameters.

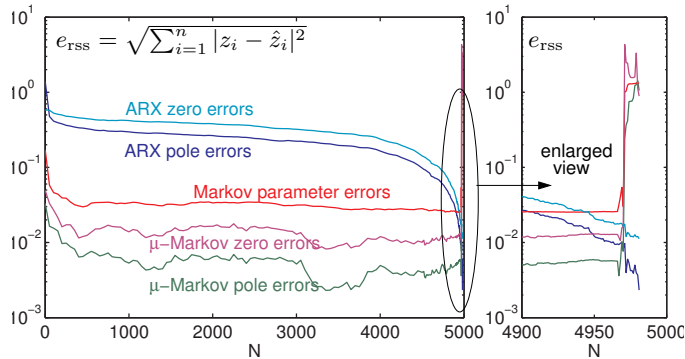


Fig. 7. CLS estimation errors for $G_1(z)$, $F_1[u] = u + 4u^3$ as functions of N ; see equation (8).

4.5 $G_1(z)$ with chirp input and uncorrelated input, output noises

Table 5 shows the results for the linear block $G_1(z)$ with nonlinear function F_2 and the chirp input signal (13) sampled at 100 Hz. At this sampling frequency, $G_1(z)$ has equivalent continuous-time poles with natural frequencies at 0.82 Hz and 20.1 Hz. Hence, the input signal bandwidth is sufficiently wide to cover the spectral range of $G_1(z)$ (Fig. 4). Performance similar to the case of Gaussian input can be observed (compare Table 4 and Table 5), i.e. the μ -Markov model is more accurate than the ARX model for all three identification methods, whereas CLS remains accurate with both models.

4.6 $G_2(z)$ with Gaussian white-noise input and uncorrelated input, output noises

The tests with Gaussian input and uncorrelated input and output noises are repeated for the linear block $G_2(z)$. Fig. 8 summarizes the resulting pole and zero errors. One can observe the same trend as in the previous cases (Fig. 6), i.e. for each method the μ -Markov model yields better estimates than the ARX model. Moreover, the accuracy of CLS with both models is consistent over different nonlinear functions. CLS is also more accurate with the ARX model, except for F_1 in which case NLS is the most accurate because the nonlinear model is matched.

As a particular case, Table 6 gives the results for the linear block $G_2(z)$ with nonlinear function F_2 . It can be seen that for all three methods, the largest errors occur for the small pole (0.25) and small zero (-0.20). Hence, sufficiently fast sampling is required, as is well known.

5. CONCLUSION

This paper presents a correlation least-squares method for identification of the SISO Hammerstein model, where both the ARX and μ -Markov model structures have been considered for the linear block. The proposed method

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.243±0.565i	0.487	0.242±0.576i	0.509
0.950	1.250	0.929	1.259	0.951	1.241
RSS errors:		0.0245	0.0153	0.0049	0.0126
Gain:		0.700		0.708	

Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.049±0.372i	-0.249	0.250±0.570i	0.521
0.950	1.250	0.505	1.162	0.952	1.243
RSS errors:		0.5966	0.7545	0.0093	0.0220
Gain:		0.514		0.515	

Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.121±0.504i	0.053	0.254±0.580i	0.513
0.950	1.250	0.706	1.218	0.946	1.253
RSS errors:		0.3161	0.4478	0.0157	0.0135
Gain:		0.703		0.700	

Table 4. $G_1(z)$, $F_2[v] = \tanh(8v)/8$, white-noise input, uncorrelated input/output noises

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.246±0.573i	0.500	0.242±0.575	0.485
0.950	1.250	0.950	1.249	0.946	1.258
RSS errors:		0.0023	0.0007	0.0066	0.0171
Gain:		0.566		0.563	

Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.045±0.515i	-0.075	0.246±0.572i	0.503
0.950	1.250	0.791	1.264	0.950	1.248
RSS errors:		0.3349	0.5756	0.0036	0.0033
Gain:		0.305		0.311	

Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
0.245±0.574i	0.500	0.158±0.567i	0.246	0.236±0.582i	0.480
0.950	1.250	0.877	1.257	0.949	1.241
RSS errors:		0.1437	0.2537	0.0168	0.0222
Gain:		0.564		0.571	

Table 5. $G_1(z)$, $F_2[v] = \tanh(8v)/8$, chirp input, uncorrelated input/output noises

has been compared with the well-known nonlinear least-squares method of Bai [1998] and, for completeness, the standard least-squares method. Our main concern has been the accuracy of pole and zero estimates of the linear block. Based on the numerical examples presented, the following conclusions can be made:

- For each identification method, the μ -Markov model yields better estimates than the ARX model. This result is consistent with both Gaussian and chirp input signals.
- Accuracy of the correlation least-squares method with both ARX and μ -Markov models is the most consistent across different nonlinear functions, compared with the nonlinear and standard least-squares methods.
- The nonlinear least-squares method assumes an *a priori* known model of the nonlinear block, in the absence of which its accuracy deteriorates significantly.
- Hence, the correlation method with separable input signal has the advantage of accurately identifying the linear block's poles and zeros without the need for a

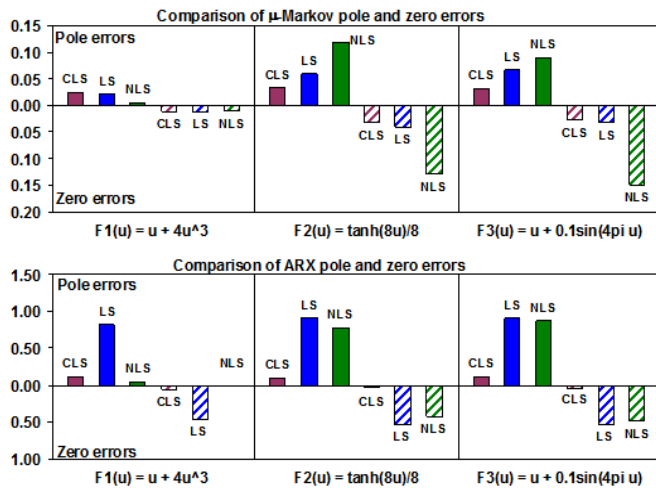


Fig. 8. Comparison of pole & zero errors, $G_2(z)$, noisy case.

Correlation LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
-0.32±0.90i	0.14±0.97i	-0.32±0.90i	0.14±0.97i	-0.32±0.90i	0.14±0.98i
0.25±0.58i	1.25	0.27±0.56i	1.25	0.25±0.57i	1.27
0.25	-0.20	0.16	-0.24	0.22	-0.22
RSS errors:		0.1010	0.0365	0.0346	0.0319
Gain:		0.700		0.688	
Standard LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
-0.32±0.90i	0.14±0.97i	-0.32±0.91i	0.11±0.92i	-0.32±0.90i	0.14±0.97i
0.25±0.58i	1.25	0.39±0.51i	1.24	0.25±0.56i	1.26
0.25	-0.20	-0.64	-0.74	0.19	-0.24
RSS errors:		0.9143	0.5438	0.0602	0.0396
Gain:		0.508		0.504	
Nonlinear LS					
Plant:		ARX model:		μ -Markov model:	
Poles:	Zeros:	Poles:	Zeros:	Poles:	Zeros:
-0.32±0.90i	0.14±0.97i	-0.32±0.91i	0.11±0.92i	-0.32±0.90i	0.14±0.96i
0.25±0.58i	1.25	0.38±0.51i	1.22	0.27±0.63i	1.22
0.25	-0.20	-0.50	-0.61	0.35	-0.07
RSS errors:		0.7748	0.4233	0.1189	0.1298
Gain:		0.723		0.724	

Table 6. $G_2(z)$, $F_2[v] = \tanh(8v)/8$, white-noise input, uncorrelated input/output noises

a priori knowledge of the nonlinear block. Mismatch in the nonlinear model leads however to a wrong estimate of the static gain.

- e. The correlation method has a slight disadvantage in that it requires more data points than the other two methods. Accuracy of the method depends on both the number of data points used for approximating each correlation value, and the number of correlation values available.

A straight-forward extension of the proposed correlation method to the MIMO case can be done if the MIMO nonlinear block is square diagonal (Lum and Lai [2010]). However, extension to the general MIMO case and other block-oriented models requires further work.

REFERENCES

K. J. Astrom and B. Wittemark. *Adaptive Control*. 1995.
E.-W. Bai. An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Proc. of the*

American Control Conference, 1998.
D. H. Baldelli, R. Lind, and M. Brenner. Nonlinear aeroelastic/aeroservoelastic modeling by block-oriented identification. *J. of Guidance, Control, and Dynamics*, 28(5):1056–1064, 2005.
S. A. Billings and S. Y. Fakhouri. Identification of nonlinear systems using correlation analysis. *Proceedings of the IEE*, 125(7):691–697, 1978a.
S. A. Billings and S. Y. Fakhouri. Identification of nonlinear unity feedback systems. In *IEEE Conference on Decision and Control*, 1978b.
S. A. Billings and S. Y. Fakhouri. Nonlinear system identification using the Hammerstein model. *International Journal of System Science*, 10(5):567–578, 1979.
C. T. Chen. *Linear System Theory and Design*. 1984.
P. Crama and J. Schoukens. Hammerstein-Wiener system estimator initialization. *Automatica*, 40(9):1543–1550, 2004.
A.M. D’Amato, B.O.S. Teixeira, and D.S. Bernstein. Semi-parametric identification of Wiener systems using a single harmonic input and retrospective cost optimization. In *American Control Conference*, Baltimore, MD, 2010.
M. Enqvist and L. Ljung. Linear approximations of nonlinear FIR systems for separable input processes. *Automatica*, 41(3):459–473, 2005.
M. S. Fledderjohn, M. S. Holzel, H. J. Palanhandalam-Madapusi, R. J. Fuentes, and D.S. Bernstein. A comparison of least squares algorithms for estimating Markov parameters. In *American Control Conference*, Baltimore, MD, 2010.
R. Isermann, U. Baur, and W. Bamberger. Comparison of six on line identification and parameter estimation methods. *Automatica*, 10(1):81–103, 1974.
J. N. Juang. *Applied System Identification*. 1994.
K.-Y. Lum and K.L. Lai. Identification of a Hammerstein model for wing flutter analysis using CFD data and correlation method. In *American Control Conference*, Baltimore, MD, 2010.
K. Narendra and P. Gallman. An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, 11(3): 546–550, 1966.
A. H. Nuttall. *Theory and application of the separable class of random processes*. PhD thesis, Research Laboratory of Electronics, MIT, 1958.
H.J. Palanhandalam-Madapusi, E.L. Renk, and D.S. Bernstein. Data-based model refinement for linear and Hammerstein systems using subspace identification and adaptive disturbance rejection. In *IEEE Conference on Control Applications*, pages 1630–1635, 2005.
T. Van Pelt and D.S. Bernstein. Least squares identification using μ -Markov parameterizations. In *IEEE Conference on Decision and Control*, 1998.
R. Venugopal and D.S. Bernstein. Adaptive disturbance rejection using ARMARKOV/Toeplitz models. *IEEE Transactions on Control Systems Technology*, 8(2):257–269, March 2000.
J. Zeng, D.H. Baldelli, and M. Brenner. Novel nonlinear Hammerstein model identification: Application to nonlinear aeroelastic/aeroservoelastic system. *J. of Guidance, Control, and Dynamics*, 31(6):1677–1686, 2008.