2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Consistent Identification of Hammerstein Systems Using an Ersatz Nonlinearity

Asad A. Ali[1], A. M. D'Amato[1], M. S. Holzel[1], S. L. Kukreja[2], and Dennis S. Bernstein[3]

***Abstract*— We develop a method for identifying SISO Hammerstein systems with an unknown static nonlinearity, linear dynamics, white input noise and colored output noise. We use least squares with a μ-Markov model to estimate the Markov parameters of the linear time-invariant dynamical system. Since the input to the linear system is not available, we use a substitute (ersatz) nonlinearity to transform the input for use in the regressor matrix. We prove that the Markov parameters of the system can be estimated consistently up to a constant scalar as the amount of data increases. This method is demonstrated with several numerical examples.**

I. INTRODUCTION

Block-structured nonlinear models, also known as gray-box models, are widely used for system identification [1], [2], [3]. Hammerstein and Wiener models, which consist of a linear system cascaded with an input and output nonlinearity, respectively, continue to attract considerable attention [4]-[9]. The range of techniques applied to these models is vast and diverse [10], [11].

In the present paper we develop an alternative approach to identifying Hammerstein systems. We assume that the input nonlinearity and the linear system are both unknown; no assumptions are made on either component of the model except that the linear system is asymptotically stable. We inject a realization of a white noise signal into the Hammerstein system, and we measure the resulting output. The initial state of the linear system is arbitrary and unknown. We then perform linear system identification by using the input and output data.

The presence of the unknown input nonlinearity degrades the accuracy of linear system identification. If the input nonlinearity were known, then the input data could be transformed by the nonlinear mapping, allowing the transformed data to be used to identify the linear system. Since we assume that the input nonlinearity is unknown, however, we use an ersatz nonlinearity to transform the input data before performing linear system identification. The purpose of the present paper is to show that, if the ersatz nonlinearity is chosen to be neither even nor odd and has zero integral with respect to the probability distribution used to generate the input signal, then the linear system can be identified consistently as the amount of data increases. In the special case that the input nonlinearity is known to be odd and the probability density function of the input and its noise are symmetric, then the ersatz nonlinearity can be chosen to be odd, for example, u. We emphasize that the ersatz nonlinearity need not be an approximation for the true nonlinearity. The true input nonlinearity can subsequently be identified by means of input reconstruction [12], [13].

The contents of the paper are as follows: In Section II, we formulate the problem. In Section III, we define the μ-Markov structure and the identification architecture. In Section IV, we analyze the consistency of the Markov parameters obtained from the proposed method. In Section V, we give some numerical examples. Finally, in Section V, we give conclusions.

II. PROBLEM FORMULATION

Consider the Hammerstein structure shown in Figure 1, where u_0 is the input signal, $\mathcal{H} : \mathbb{R} \to \mathbb{R}$ is the static nonlinearity, $\mathcal{H}(u_0)$ is the intermediate signal, y_0 is the output signal, and G is the asymptotically stable, SISO, causal, discrete-time system

$$A(\mathbf{q})y_0(k) = B(\mathbf{q})\mathcal{H}\Big(u_0(k)\Big), \tag{1}$$

where $\mathbf{q}$ is the forward shift operator, and A and B are polynomials in $\mathbf{q}$. Throughout the paper, we assume that the nonlinearity $\mathcal{H}$, and hence the intermediate signal $\mathcal{H}(u_0)$, is unknown. Furthermore, we assume that the measurement of u_0 is corrupted with additive white noise, and the measurement of y_0 is corrupted with additive white or colored noise.

The ARX model of (1) is given by

$$y_0(k) = \sum_{j=0}^{n} b_j \mathcal{H}\Big(u_0(k-j)\Big) - \sum_{j=1}^{n} a_j y_0(k-j). \tag{2}$$

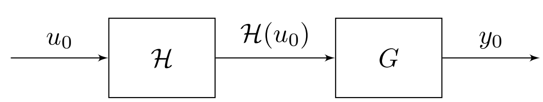

Fig. 1. Block-structured Hammerstein model where u_0 is the input, $\mathcal{H}\Big(u_0(k)\Big)$ is the intermediate signal, and y_0 is the output.

III. μ-MARKOV MODEL AND LEAST SQUARES ESTIMATES

For all $k \geq 0$, $n_{\rm mod} \geq n$, $\mu \geq 1$, the signals $y_0(k)$ and $\mathcal{H}(u_0(k))$ satisfy the μ-Markov model

$$\begin{aligned} y_0(k) = &\sum_{j=0}^{\mu-1} H_j \mathcal{H}(u_0(k-j)) + \sum_{j=\mu}^{n_{\rm mod}+\mu-1} b'_j \mathcal{H}(u_0(k-j)) \\ &- \sum_{j=\mu}^{n_{\rm mod}+\mu-1} a'_j y_0(k-j), \end{aligned} \tag{3}$$

[1]Graduate students, Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, asadali@umich.edu, amdamato@umich.edu, mholzel@umich.edu.

[2]Research Engineer, NASA Dryden Flight Research Center, Edwards, California 93523, sunil.l.kukreja@gmail.com.

[3]Professor, Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, dsbaero@umich.edu.

where $H_0, \ldots, H_{\mu-1}$ are the first μ Markov parameters of G. Note that the μ-Markov model is an overparameterization of the ARX model (2), where μ Markov parameters are explicitly displayed [14]. Furthermore, setting $n_{\rm mod} = n$ and $\mu = 1$ gives the ARX model (2).

Next, let $n_{\rm mod} \geq n$ and $\mu \geq 1$. Then the μ-Markov model (3) of (1) can be expressed as

$$y_0(k) = \theta_\mu \mathcal{H}(\phi_{\mu_0})(k) + \theta_u \mathcal{H}(\phi_{u_0})(k) - \theta_y \phi_{y_0}(k), \quad (4)$$

where

$$\begin{aligned}
\theta_\mu &\triangleq \left[\begin{array}{ccc} H_0 & \cdots & H_{\mu-1} \end{array} \right], \\
\theta_u &\triangleq \left[\begin{array}{ccc} b_\mu & \cdots & b_{n_{\rm mod}+\mu-1} \end{array} \right], \\
\theta_y &\triangleq \left[\begin{array}{ccc} a_\mu & \cdots & a_{n_{\rm mod}+\mu-1} \end{array} \right], \\
\mathcal{H}(\phi_{\mu_0})(k) &\triangleq \left[\begin{array}{ccc} \mathcal{H}(u_0(k)) & \cdots & \mathcal{H}(u_0(k-\mu+1)) \end{array} \right]^{\rm T}, \\
\mathcal{H}(\phi_{u_0})(k) &\triangleq \left[\mathcal{H}(u_0(k-\mu)) \cdots \mathcal{H}(u_0(k-n_{\rm mod}-\mu+1)) \right]^{\rm T}, \\
\phi_{y_0}(k) &\triangleq \left[y_0(k-\mu) \ \cdots \ y_0(k-n_{\rm mod}-\mu+1) \right]^{\rm T}.
\end{aligned}$$

Furthermore, least squares estimates $\hat{\theta}^{\mathcal{H}}_{\mu,\ell}$, $\hat{\theta}^{\mathcal{H}}_{u,\ell}$, $\hat{\theta}^{\mathcal{H}}_{y,\ell}$ of θ_μ, θ_u, θ_y are given by

$$\begin{aligned}
&\left[\begin{array}{ccc} \hat{\theta}^{\mathcal{H}}_{\mu,\ell} & \hat{\theta}^{\mathcal{H}}_{u,\ell} & \hat{\theta}^{\mathcal{H}}_{y,\ell} \end{array} \right] \quad (5) \\
&= \underset{[\,\theta_\mu\ \theta_u\ \theta_y\,]}{\arg\min} \|\Psi_{y_0,\ell} - \theta_\mu \mathcal{H}(\Phi_{\mu_0,\ell}) - \theta_u \mathcal{H}(\Phi_{u_0,\ell}) + \theta_y \Phi_{y_0,\ell}\|_{\rm F},
\end{aligned}$$

where $||\ .\ ||_{\rm F}$ denotes the Frobenius norm,

$$\begin{aligned}
\Psi_{y_0,\ell} &\triangleq \left[\begin{array}{ccc} y_0(n_{\rm mod}+\mu-1) & \ldots & y_0(\ell) \end{array} \right], \\
\Phi_{y_0,\ell} &\triangleq \left[\begin{array}{ccc} \phi_{y_0}(n_{\rm mod}+\mu-1) & \ldots & \phi_{y_0}(\ell) \end{array} \right], \\
\mathcal{H}(\Phi_{\mu_0,\ell}) &\triangleq \left[\begin{array}{ccc} \mathcal{H}(\phi_{\mu_0})(n_{\rm mod}+\mu-1) & \ldots & \mathcal{H}(\phi_{\mu_0})(\ell) \end{array} \right], \\
\mathcal{H}(\Phi_{u_0,\ell}) &\triangleq \left[\begin{array}{ccc} \mathcal{H}(\phi_{u_0})(n_{\rm mod}+\mu-1) & \ldots & \mathcal{H}(\phi_{u_0})(\ell) \end{array} \right],
\end{aligned}$$

and ℓ is the number of samples. In this case, if $\mathcal{H}(u_0)$ is persistently exciting, then $\lim_{\ell\to\infty} \hat{\theta}^{\mathcal{H}}_{\mu,\ell} \overset{\rm wp1}{=} \theta_\mu$.

IV. Ersatz Nonlinearity

Since $\mathcal{H}$ is assumed to be unknown, we cannot construct $\mathcal{H}(\Phi_{\mu_0,\ell})$ and $\mathcal{H}(\Phi_{u_0,\ell})$. Hence, it is not possible to solve the least squares problem (5) for the coefficients of the μ-Markov model. We thus consider a least squares problem of the form (5) in which we replace the unknown nonlinearity $\mathcal{H}$ by a nonlinearity $\mathcal{N} : \mathbb{R} \to \mathbb{R}$. The identification architecture is shown in Figure 2. The ersatz nonlinearity $\mathcal{N}$ is not intended to be an approximation of $\mathcal{H}$. Rather, $\mathcal{N}$ serves as a substitute for $\mathcal{H}$ that has the ability to render the solution of the least squares problem useful for estimating the coefficients of the μ-Markov model.

V. Consistency Analysis

In this section, we investigate the effect of the choice of the ersatz nonlinearity on the consistency of the estimates of the μ-Markov parameters of G obtained from least squares with a μ-Markov model structure. Here consistency refers to the accuracy of the estimates as the number of data points increases.

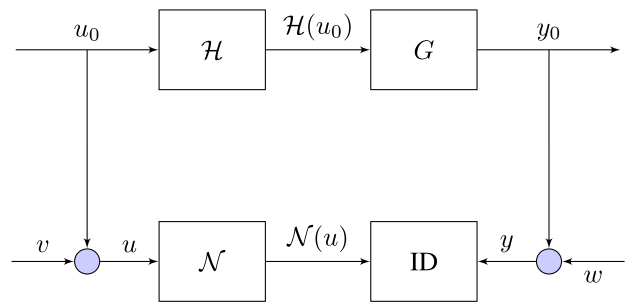

Fig. 2. Identification of a Hammerstein system using the ersatz nonlinearity $\mathcal{N}$.

Let u_0 and y_0 satisfy (1). We use the following assumptions.

Assumption 5.1: u_0 and v are realizations of the stationary white processes U_0 and V, respectively.

Assumption 5.2: w is a realization of the stationary white or colored process W.

Assumption 5.3: U_0, W, and V are mutually independent.

Assumption 5.4: $U(k) = U_0(k) + V(k)$.

Assumption 5.5: $Y(k) = Y_0(k) + W(k)$.

Assumption 5.6: $\theta_\mu \neq 0_{1\times\mu}$.

Assumption 5.7: For all $k \geq 0$, $\mathbb{E}\Big[\mathcal{N}\Big(U(k)\Big)\Big] = 0$ and $\mathbb{E}\Big[\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k)\Big)\Big] \neq 0$.

Assumption 5.8: For all $k \geq 0$ and $p \geq -k$, $\mathcal{H}\big(U(k)\big)$, $\mathcal{N}\big(U(k)\big)$, $\mathcal{H}\big(U(k)\big)\mathcal{N}\big(U(k+p)\big)$, $\mathcal{N}\big(U(k)\big)\mathcal{N}\big(U(k+p)\big)$, and $W(k)$ have finite mean and variance.

Next, consider the least squares estimates $\hat{\theta}_{\mu,\ell}$, $\hat{\theta}_{u,\ell}$, $\hat{\theta}_{y,\ell}$ of θ_μ, θ_u, θ_y, given by

$$\begin{aligned}
&\left[\begin{array}{ccc} \hat{\theta}_{\mu,\ell} & \hat{\theta}_{u,\ell} & \hat{\theta}_{y,\ell} \end{array} \right] \quad (6) \\
&= \underset{[\,\theta_\mu\ \theta_u\ \theta_y\,]}{\arg\min} \|\Psi_{y,\ell} - \theta_\mu \mathcal{N}(\Phi_{\mu,\ell}) - \theta_u \mathcal{N}(\Phi_{u,\ell}) + \theta_y \Phi_{y,\ell}\|_{\rm F},
\end{aligned}$$

where $\mathcal{N}$, u, and y replace $\mathcal{H}$, u_0, and y_0, respectively in (5). Specifically

$$\begin{aligned}
\Psi_{y,\ell} &\triangleq \left[\begin{array}{ccc} y(n_{\rm mod}+\mu-1) & \ldots & y(\ell) \end{array} \right], \\
\mathcal{N}(\phi_\mu)(k) &\triangleq \left[\begin{array}{ccc} \mathcal{N}(u(k)) & \ldots & \mathcal{N}(u(k-\mu+1)) \end{array} \right]^{\rm T}, \\
\mathcal{N}(\phi_u)(k) &\triangleq \left[\mathcal{N}(u(k-\mu)) \ \ldots \ \mathcal{N}(u(k-n_{\rm mod}-\mu+1)) \right]^{\rm T}, \\
\mathcal{N}(\Phi_{\mu,\ell}) &\triangleq \left[\begin{array}{ccc} \mathcal{N}(\phi_\mu)(n_{\rm mod}+\mu-1) & \ldots & \mathcal{N}(\phi_\mu)(\ell) \end{array} \right], \\
\mathcal{N}(\Phi_{u,\ell}) &\triangleq \left[\begin{array}{ccc} \mathcal{N}(\phi_u)(n_{\rm mod}+\mu-1) & \ldots & \mathcal{N}(\phi_u)(\ell) \end{array} \right].
\end{aligned}$$

Definition 5.1: $\hat{\theta}_{\mu,\ell}$ is *semi-consistent* if there exists nonzero $\alpha \in \mathbb{R}$ such that

$$\lim_{\ell\to\infty} \hat{\theta}_{\mu,\ell} \overset{\rm wp1}{=} \alpha\theta_\mu.$$

Theorem 5.1: Let assumptions 5.1-5.8 hold. Then $\hat{\theta}_{\mu,\ell}$ is semi-consistent.

Proof 5.1: The μ-Markov model (4) of (1), can be expressed in terms of the regression matrices by

$$\Psi_{y_0,\ell} = \theta_\mu \mathcal{H}(\Phi_{\mu_0,\ell}) + \theta_u \mathcal{H}(\Phi_{u_0,\ell}) - \theta_y \Phi_{y_0,\ell}. \quad (7)$$

By assumption 5.5, we have that

$$\Psi_{y,\ell} = \theta_\mu \mathcal{H}(\Phi_{\mu_0,\ell}) + \theta_u \mathcal{H}(\Phi_{u_0,\ell}) - \theta_y \Phi_{y_0,\ell} + \Psi_{w,\ell}, \quad (8)$$

where

$$\Psi_{w,\ell} \stackrel{\triangle}{=} \left[\begin{array}{ccc} w(n_{\rm mod}+\mu-1) & \ldots & w(\ell) \end{array}\right].$$

Furthermore, $\hat\theta_{\mu,\ell}$, $\hat\theta_{u,\ell}$, and $\hat\theta_{y,\ell}$ satisfy

$$\begin{aligned}&\Psi_{y,\ell}\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \quad (9)\\ &\quad= \Big[\hat\theta_{\mu,\ell}\mathcal{N}(\Phi_{\mu,\ell}) + \hat\theta_{u,\ell}\mathcal{N}(\Phi_{u,\ell}) - \hat\theta_{y,\ell}\Phi_{y,\ell}\Big]\mathcal{N}(\Phi_{\mu,\ell})^{\rm T}.\end{aligned}$$

Hence, from (8) and (9), it follows that

$$\begin{aligned}&\Big[\theta_\mu\mathcal{H}(\Phi_{\mu_0,\ell}) + \theta_u\mathcal{H}(\Phi_{u_0,\ell}) - \theta_y\Phi_{y_0,\ell} + \Psi_{w,\ell}\Big]\mathcal{N}(\Phi_{\mu,\ell})^{\rm T}\\ &= \Big[\hat\theta_{\mu,\ell}\mathcal{N}(\Phi_{\mu,\ell}) + \hat\theta_{u,\ell}\mathcal{N}(\Phi_{u,\ell}) - \hat\theta_{y,\ell}\Phi_{y,\ell}\Big]\mathcal{N}(\Phi_{\mu,\ell})^{\rm T}. \quad (10)\end{aligned}$$

Next, assumptions 5.1-5.5 imply that, for all $p > -k$, we have that

$$\begin{aligned}\mathbb{E}\Big[\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k+p)\Big)\Big] &= \mathbb{E}\Big[\mathcal{H}\Big(U_0(k)\Big)\Big]\mathbb{E}\Big[\mathcal{N}\Big(U(k+p)\Big)\Big]\\ &= 0.\end{aligned}$$

Hence,

$$\begin{aligned}&\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{H}(\Phi_{\mu_0,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T}\\ &\stackrel{\rm wp1}{=} \mathbb{E}\begin{bmatrix}\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k)\Big) & \ldots & \mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k')\Big)\\ \vdots & \ddots & \vdots\\ \mathcal{H}\Big(U_0(k')\Big)\mathcal{N}\Big(U(k)\Big) & \ldots & \mathcal{H}\Big(U_0(k')\Big)\mathcal{N}\Big(U(k')\Big)\end{bmatrix},\\ &= \mathbb{E}\left[\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k)\Big)\right] I_\mu, \quad (11)\end{aligned}$$

and

$$\begin{aligned}&\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{N}(\Phi_{\mu,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T}\\ &\stackrel{\rm wp1}{=} \mathbb{E}\begin{bmatrix}\mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k)\Big) & \ldots & \mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k')\Big)\\ \vdots & \ddots & \vdots\\ \mathcal{N}\Big(U(k')\Big)\mathcal{N}\Big(U(k)\Big) & \ldots & \mathcal{N}\Big(U(k')\Big)\mathcal{N}\Big(U(k')\Big)\end{bmatrix},\\ &= \mathbb{E}\left[\mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k)\Big)\right] I_\mu, \quad (12)\end{aligned}$$

where $k' \stackrel{\triangle}{=} k-\mu+1$. Similarly, from independence and causality it follows that

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{H}(\Phi_{u_0,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{n_{\rm mod}\times\mu}, \quad (13)$$

$$\lim_{\ell\to\infty}\frac{1}{\ell}\Phi_{y_0,\ell}\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{n_{\rm mod}\times\mu}, \quad (14)$$

$$\lim_{\ell\to\infty}\frac{1}{\ell}\Psi_{w,\ell}\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{1\times\mu}, \quad (15)$$

$$\lim_{\ell\to\infty}\frac{1}{\ell}\Phi_{y,\ell}\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{n_{\rm mod}\times\mu}, \quad (16)$$

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{N}(\Phi_{\mu,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{\mu\times\mu}. \quad (17)$$

Hence, from (10)-(17), it follows that

$$\begin{aligned}&\mathbb{E}\left[\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k)\Big)\right]\theta_\mu \quad (18)\\ &\qquad\stackrel{\rm wp1}{=} \mathbb{E}\left[\mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k)\Big)\right]\lim_{\ell\to\infty}\hat\theta_{\mu,\ell},\end{aligned}$$

where from Assumption 5.7, $\mathbb{E}\left[\mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k)\Big)\right]$ is the variance of $\mathcal{N}\big(U(k)\big)$. □

Let $P\Big(U_0(k)\Big)$ denote the probability density function of $U_0(k)$. We consider two cases in the absence of noise:

Case A. $\mathcal{H}$ odd, $\mathcal{N}$ odd, and $P\Big(U_0(k)\Big)$ symmetric.

In this case,

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{N}(\Phi_{\mu,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} \beta I_\mu,$$

where $\beta \stackrel{\triangle}{=} \mathbb{E}\Big[\mathcal{N}\Big(U(i)\Big)\mathcal{N}\Big(U(i)\Big)\Big] \neq 0$. Next, if

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{H}(\Phi_{\mu_0,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} \alpha I_\mu \neq 0_{\mu\times\mu},$$

then it follows from (18) that

$$\lim_{\ell\to\infty}\hat\theta_{\mu,\ell} \stackrel{\rm wp1}{=} \frac{\alpha}{\beta}\theta_\mu.$$

Case B. $\mathcal{H}$ even and $P\Big(U_0(k)\Big)$ symmetric.

In this case,

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{N}(\Phi_{\mu,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} \beta I_\mu,$$

where $\beta \stackrel{\triangle}{=} \mathbb{E}\Big[\mathcal{N}\Big(U(i)\Big)\mathcal{N}\Big(U(i)\Big)\Big] \neq 0$. Next, If $\mathcal{N}$ is odd, then

$$\lim_{\ell\to\infty}\frac{1}{\ell}\mathcal{H}(\Phi_{\mu_0,\ell})\mathcal{N}(\Phi_{\mu,\ell})^{\rm T} \stackrel{\rm wp1}{=} 0_{\mu\times\mu}$$

Hence, it follows from (18) that

$$\lim_{\ell\to\infty}\hat\theta_{\mu,\ell} \stackrel{\rm wp1}{=} 0_{1\times\mu}.$$

If $\mathcal{N}$ is neither even nor odd and $\mathbb{E}\Big[\mathcal{N}\Big(U(k)\Big)\Big] = 0$, then $\alpha \stackrel{\triangle}{=} \mathbb{E}\Big[\mathcal{H}\Big(U_0(i)\Big)\mathcal{N}\Big(U(i)\Big)\Big]$ can be nonzero. For the case in which $\alpha \neq 0$, it follows from (18) that

$$\lim_{\ell\to\infty}\hat\theta_{\mu,\ell} \stackrel{\rm wp1}{=} \frac{\alpha}{\beta}\theta_\mu.$$

Example 5.1: Let $\mathcal{H}(x) = x^3$, $\mathcal{N}(x) = x$, $v = 0$ and let $U_0(k)$ be uniformly distributed with density function

$$P\Big(U_0(k)\Big) = \begin{cases}\frac{1}{2a}, & |U_0(k)| \le a,\\ 0, & |U_0(k)| > a.\end{cases}$$

Then

$$\mathbb{E}\Big[\mathcal{H}\Big(U_0(k)\Big)\mathcal{N}\Big(U(k)\Big)\Big] = \frac{1}{2a}\int_{-a}^{a} U_0^4(k)\,dU_0(k) = a^4/5,$$

$$\mathbb{E}\Big[\mathcal{N}\Big(U(k)\Big)\mathcal{N}\Big(U(k)\Big)\Big] = \frac{1}{2a}\int_{-a}^{a} U_0^2(k)\,dU_0(k) = a^2/3.$$

Finally, it follows from (18) that

$$\lim_{\ell\to\infty}\hat\theta_{\mu,\ell} \stackrel{\rm wp1}{=} \frac{3a^2}{5}\theta_\mu.$$

VI. NUMERICAL EXAMPLES

Consider the system

$$A(\mathbf{q}) = (\mathbf{q}-0.3)(\mathbf{q}-0.8)(\mathbf{q}^2+0.04),$$
$$B(\mathbf{q}) = (5\mathbf{q}-1)(\mathbf{q}^2-2.4\mathbf{q}+1.69),$$

and define the normalized Markov parameters

$$H'_i \triangleq \frac{H_i}{H_j},$$

where H_j is the first nonzero Markov parameter. For this system, $j = 2$. The estimated Markov parameters $\hat{H}_i$, obtained from $\hat{\theta}_{\mu,\ell}$, are normalized by $\hat{H}_2$ to obtain $\hat{H}'_i$. In the following simulations, we overestimate the model order and assume the relative degree is zero. The least squares estimates are computed for 200 independent realizations of U. We also define the error metric

$$\varepsilon = \frac{1}{200}\sum_{i=0}^{200}\frac{|H'_i - \hat{H}'_i|}{|H'_i|}.$$

Unless otherwise specified, $v(k)$=0 and $w(k)$=0 for all k.

A. $\mathcal{H}(u) = u^3$, $\mathcal{N}(u) = u$

Let U be white and have the uniform pdf

$$P\Big(U(k)\Big) = \begin{cases} \frac{1}{20}, & |x| < 10, \\ 0, & |x| > 10. \end{cases} \tag{19}$$

Figure 3 indicates that all of the Markov parameter estimates are consistent up to a scalar multiple.

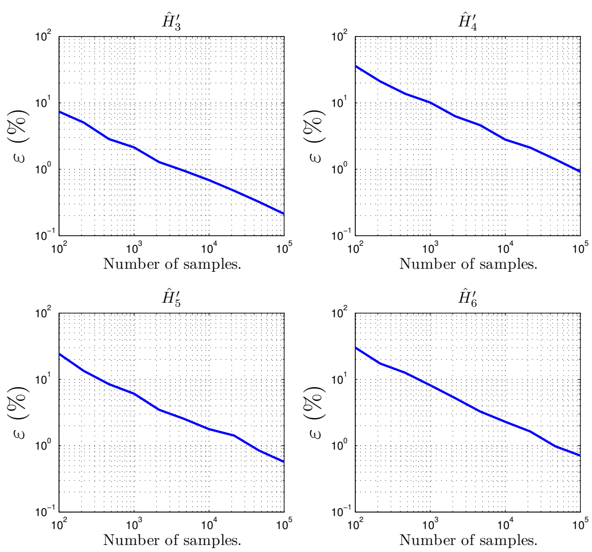

Fig. 3. Consistency of the Markov parameters obtained from the μ-Markov model with the Hammerstein nonlinearity $\mathcal{H}(u) = u^3$ and the ersatz nonlinearity $\mathcal{N}(u) = u$.

B. $\mathcal{H}(u) = \sin(u)$, $\mathcal{N}(u) = \mathrm{sign}(u)$

Let U be white and have the uniform pdf (19). Figure 4 indicates that all of the Markov parameter estimates are consistent up to a scalar multiple.

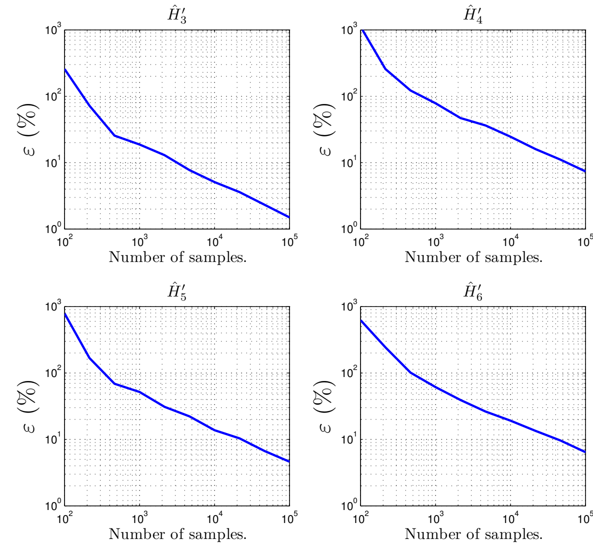

Fig. 4. Consistency of the Markov parameters obtained from the μ-Markov model with the Hammerstein nonlinearity $\mathcal{H}(u) = \sin(u)$ using the ersatz nonlinearity $\mathcal{N}(u) = \mathrm{sign}(u)$.

C. $\mathcal{H}(u) = u^2$, $\mathcal{N}(u) = -5 + |u|$

Let U be white and have the uniform pdf (19). Figure 7 indicates that all of the Markov parameter estimates are consistent up to a scalar multiple.

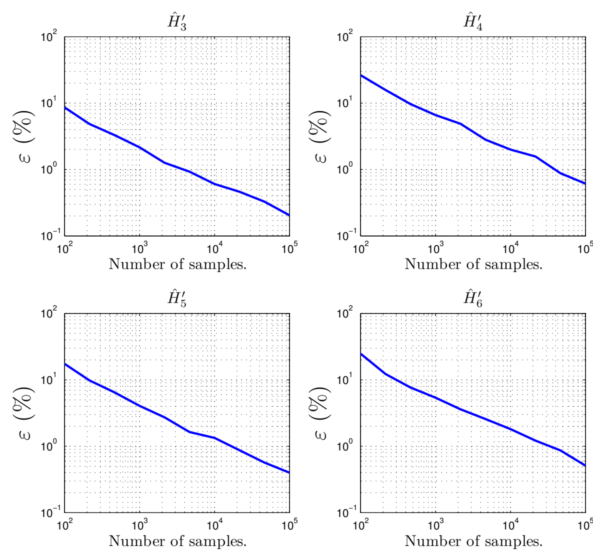

Fig. 5. Consistency of the Markov parameters obtained from the μ-Markov model with the Hammerstein nonlinearity $\mathcal{H}(u) = u^2$ using the ersatz nonlinearity $\mathcal{N}(u) = -5 + |u|$.

D. $\mathcal{H}(u) = u^3$, $\mathcal{N}(u) = ue^{u^2/2}$

Let U be white and have a Gaussian pdf with mean zero and unit variance. Figure 6 indicates that all of the Markov parameter estimates are consistent up to a scalar multiple.

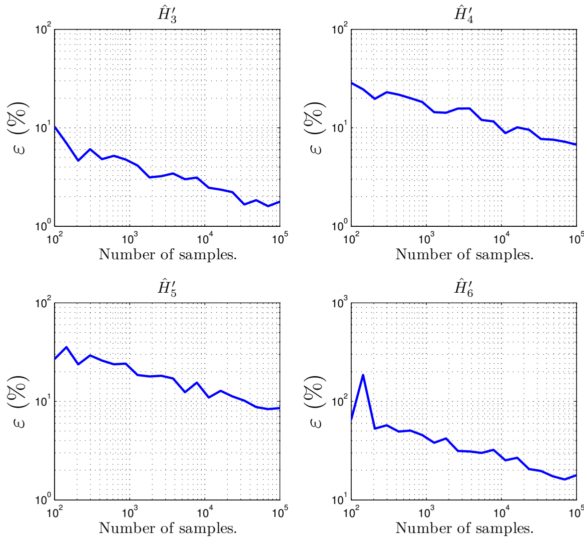

Fig. 6. Consistency of the Markov parameters obtained from the μ-Markov model with the Hammerstein nonlinearity $\mathcal{H}(u) = u^3$ using the ersatz nonlinearity $\mathcal{N}(u) = ue^{u^2/2}$.

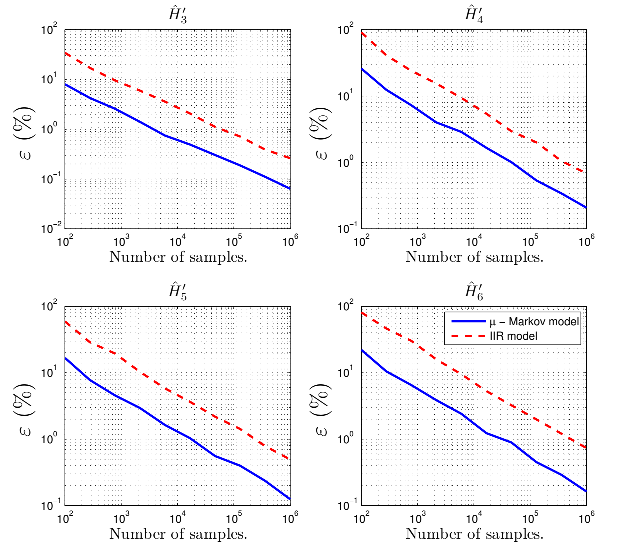

Fig. 7. Consistency of the Markov parameters obtained from the μ-Markov model and the IIR model with the Hammerstein nonlinearity $\mathcal{H}(u) = u^2$ using the ersatz nonlinearity $\mathcal{N}(u) = -5 + |u|$. This figure shows that the estimates obtained from the IIR model are biased.

E. Comparison with the standard correlation method

In this example, we compare the method proposed in this paper with the correlation method [10], where the model structure considered is the infinite impulse response(IIR) model

$$y(k) = \sum_{j=0}^{\infty} H_j u(n-j). \tag{20}$$

Assuming there is no noise in the system, the initial conditions are zero, and $U(k)$ is stationary white noise, semi-consistent estimates of the Markov parameters can be obtained using

$$\beta \hat{H}_i = \mathbb{E}\big[Y(k+i)\mathcal{N}\big(U(k)\big)\big],$$

where $\beta = \mathbb{E}\big[\mathcal{N}\big(U(k)\big)U(k)\big]$.

For this example, let $\mathcal{H} = u^2$, $\mathcal{N} = -5 + |u|$ and $v = 0$. Let U have the uniform pdf (19) and let the initial conditions be non-zero, specifically, let $y(0) = 40.93 + H_0 u(0) = 40.93$. Figure 7 indicates that the Markov parameters obtained from IIR model ((20)) are biased.

VII. Conclusions

We considered the use of a substitute nonlinearity in the least squares identification of Hammerstein systems with a μ-Markov model. We proved that, under certain assumptions, using this ersatz nonlinearity yields estimates of the Markov parameters of the system consistently up to a scalar multiple, which is semi-consistency. This method is demonstrated on several examples, including a system with a signum Hammerstein nonlinearity. Future research will focus on techniques for constructing the ersatz nonlinearity to optimize the rate of convergence of the Markov parameter estimates.

References

[1] R. Haber and L. Keviczky, *Nonlinear System Identification−Input-Output Modeling Approach*. Kluwer Academic Publishers, 1999, vol. 1.

[2] W. Greblicki and M. Pawlak, *Nonparametric System Identification*. Cambridge University Press, 2008.

[3] F. Giri and E. W. Bai, *Block-oriented Nonlinear System Identification*. Springer, 2010, vol. 1.

[4] E. W. Bai and J. Reyland, "Towards identification of wiener systems with the least amount of a *Priori* information: Iir cases," *Automatica*, vol. 45, pp. 956–964, 2009.

[5] L. A. Aguirre, M. C. S. Coelho, and M. V. Corrêa, "On the interpretation and practice of dynamical differences between hammersten and wiener models," *IEEE Proc. - Control Theory Appl.*, vol. 152, no. 4, pp. 349–356, 2005.

[6] E. W. Bai, "Frequency domain identification of wiener models," *Automatica*, vol. 39, pp. 1521–1530, 2003.

[7] S. L. Lacy, R. S. Erwin, and D. S. Bernstein, "Identification of wiener systems with known, noninvertible nonlinearities," *ASME J. Dyn. Sys. Meas. Contr.*, vol. 123, pp. 566–571, 2001.

[8] P. Crama and J. Schoukens, "Initial estimates of wiener and hammerstein systems using multisine excitation," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 6, pp. 1791–1795, 2001.

[9] A. M. D'Amato, B. O. S. Teizeira, and D. S. Bernstein, "Semiparametric identification of wiener systems using a single harmonic input and retrospective cost optimization," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 6, pp. 1791–1795, 2001.

[10] W. Greblicki and M. Pawlak, "Identification of discrete Hammerstein systems using kernel regression estimates," vol. AC-31, 1986, pp. 74–77.

[11] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P. yves Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: a unified overview," *Automatica*, vol. 31, pp. 1691–1724, 1995.

[12] S. Kirtikar, H. Palanthandalam-Madapusi, E. Zattoni, and D. S. Bernstein, "*l*-delay input reconstruction for discrete-time systems," *Proc. Conf. Dec. Contr.*, vol. 127, pp. 1848–1853, 2009.

[13] G. Marro and E. Zattoni, "Unknown-state, unknown-input reconstruction in discrete-time nonminimum-phase systems: Geometric methods," *Automatica*, vol. 46, no. 5, pp. 815–822, 2010.

[14] J. C. Akers and D. S. Bernstein, "Time domain identification using ARMARKOV / Toeplitz models," in *Proc. Amer. Cont. Conf.*, 1997, pp. 1667–1661.