2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Output Error Identification of Closed-loop Hammerstein Systems

Younghee Han and Raymond A. de Callafon
University of California at San Diego, Mechanical and Aerospace Engineering
9500 Gilman Drive, La Jolla, CA, 92093-0401, USA

***Abstract*— A feedback connection of a Hammerstein system and a linear controller provides a good approach to modeling systems with static actuator nonlinearity or input saturation during closed-loop experiments. This paper discusses the Output Error (OE) identification of such closed-loop Hammerstein systems that requires a nonlinear optimization due to the non-convexity of the output error. An iterative IV (Instrumental Variables) identification algorithm is proposed for the nonlinear optimization. The basic idea is to express the nonlinear parameter estimation as an iterative IV estimation using gradient based updates. The proposed iterative IV identification method is applied to the simulation data from a Hammerstein system with input static nonlinearity. The simulation study shows the effectiveness of the proposed identification algorithms.**

I. INTRODUCTION

A feedback connection of a Hammerstein system and a linear controller can be used to represent a feedback system with actuator nonlinearity. There has been much research on the problem of identifying Hammerstein systems in a open-loop setting [1] [2] [3] [4] [5] [6] [7] [8], while much less attention has been paid to the problem of identifying Hammerstein systems in a closed-loop setting. One of the early works dealing with closed-loop Hammerstein system identification can be found in [9]. In this work, Beyer et al. proposed a closed-loop identification method for Hammerstein systems using the LS method, the GLS method and the maximum likelihood method. In addition, Linard et al. [10] extended closed-loop identification methods (a two-stage method and using right coprime factorizations) for linear dynamic systems to nonlinear dynamic systems and De Bruyne et al. [11] presented gradient expressions for a closed-loop parametric identification scheme. However, these methods are based on linearization of a nonlinear map between time domain signals. Recently, van Wingerden and Verhaegen [12] presented an algorithm to identify MIMO Hammerstein systems under open and closed-loop conditions. They formulated an optimized predictor based subspace identification algorithm in the dual space. Laurain et al. [13] presented an IV method dedicated to closed-loop Hammerstein systems. Comprehensive studies of block-oriented nonlinear system identification can be found in [8].

In this paper, we focus on the Output Error (OE) identification of Hammerstein systems in a closed-loop setting. Closed-loop identification is often used for control-relevant identification where the goal is to estimate models suitable for robust control design. It is then often only necessary to model the plant dynamics, not noise properties. So it would be natural to use an output error model structure [14].

It is well known that the direct use of input/output data, generated from a closed-loop setting, results in biased estimation due to the correlation between input and noise. The main contribution of this paper is that we propose a method that allows us to solve a nonlinear OE minimization problem as an iterative linear optimization problem that is robust to the correlation between input and noise. The basic idea is to express the nonlinear parameter estimation as an iterative IV estimation using gradient based updates similar to the method in [15]. Convergence of the iterative steps guarantees a local minimum of the OE minimization problem.

II. STATEMENT OF THE PROBLEM

The Hammerstein system in a closed-loop setting considered in this paper is shown in Figure 1. For identification purposes, the reference input $r(t)$ and the controller $C(q)$ are known, the input $u(t)$ and output $y(t)$ are measured, whereas the intermediate signal $x_0(t)$, the static nonlinearity $f_0(\cdot)$ and the linear dynamic system $G_0(q)$ are unknown. The disturbance $v(t)$ is a filtered white noise, where the filtering properties are unknown. The purpose of this study is to propose an OE identification method for the consistent estimation of the static nonlinearity $f_0(\cdot)$ and the linear dynamic system $G_0(q)$ in a closed-loop setting on the basis of the measured signals, the input $u(t)$ and output $y(t)$, and the knowledge of the controller $C(q)$.

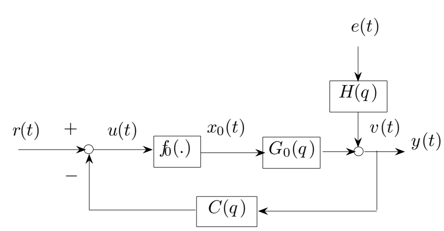

Fig. 1. Closed-loop Hammerstein system.

III. SYSTEM DESCRIPTION

A. *Modeling of static nonlinearity*

In feedback control systems, nonsmooth static nonlinearity, such as saturation, is common. A piecewise linear approximation is an excellent way to estimate such nonlinearity for feedback control systems since we can achieve good approximation with only a small number of parameters. A piecewise linear approximation of the static nonlinearity $f_0(\cdot)$ using piecewise triangle functions [16] shown in Figure 2 is used in this study. Using triangle basis functions

$f_m(\cdot)$, the static nonlinearity $f_0(\cdot)$ is assumed to satisfy the following condition

$$\sup_{u(t)\in[u_{min},u_{max}]} \lim_{M\to\infty} \sum_{m=1}^{M} |\mu_m f_m(u(t)) - f_0(u(t))| = 0 \tag{1}$$

where the center location vector $m = [m_1 \ \cdots \ m_M]^T$, specifying the center locations of triangle basis functions, spans the amplitude of the input vector $u = [u(1) \ \cdots \ u(N)]^T$ and the amplitude vector $\mu = [\mu_1 \ \cdots \ \mu_M]^T$, specifying the amplitudes of triangle basis functions at the center m, is to be estimated. The condition in (1) indicates that the static nonlinearity $f_0(\cdot)$ can be approximated arbitrarily well with a dense grid of triangular basis functions. Even a sparse grid will provide good approximation for nonsmooth nonlinearity if grid points are well chosen.

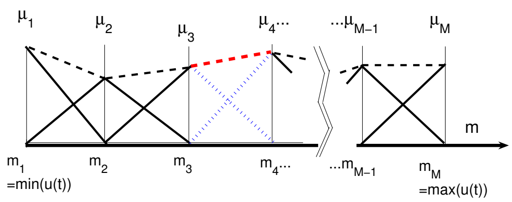

Fig. 2. Triangle basis functions.

B. Modeling of closed-loop Hammerstein systems

The output $y(t)$ generated from a closed-loop Hammerstein system is defined as

$$\begin{aligned} y(t) &= G_0(q)x_0(t) + v(t) \\ &= q^{-td}\frac{B_0(q)}{A_0(q)}x_0(t) + v(t) \end{aligned} \tag{2}$$

where $v(t)$ is colored noise, $x_0(t) = f_0(u(t))$, and td indicates the number of steps of time delay of the system. We assume at least one step time delay in $G_0(q)$. The noise free output $\hat{y}(t)$ is defined by

$$\hat{y}(t) = \frac{B_0(q)}{A_0(q)}x(t-td), \ \text{where}$$
$$x(t) = f\left(\hat{u}(t)\right), \ \hat{u}(t) = r(t) - \hat{y}_c(t), \ \hat{y}_c(t) = C(q)\hat{y}(t). \tag{3}$$

In order to define $\hat{y}(t)$, one needs $x(t-td), \cdots, x(t-n_b-td)$ and $\hat{y}(t-1), \cdots, \hat{y}(t-n_a)$. Subsequently, in order to define $\hat{y}_c(t) = C(q)\hat{y}(t)$, where

$$C(q) = \frac{d_0 + \cdots + d_{n_d}q^{-n_d}}{1 + c_1q^{-1} + \cdots + c_{n_c}q^{-n_c}}, \tag{4}$$

one needs $\hat{y}(t), \cdots, \hat{y}(t-n_d)$ and $\hat{y}_c(t-1), \cdots, \hat{y}_c(t-n_c)$.

IV. Parametrization

In order to define a piecewise linear approximation of the static nonlinearity $f_0(\cdot)$, a finite value M in (1) can be chosen, whereas the points $m_1, \cdots, m_M$ of a grid over $[u_{min}, u_{max}]$ can be chosen linearly spaced or at strategic locations. Each triangle function $f_m(u(t))$ in (1) has nonzero values through two segments and zeros elsewhere except for the first and the last intervals of the grid.

$$f_m(u(t)) = \begin{cases} \dfrac{u(t) - m_{l-1}}{m_l - m_{l-1}} & \text{for } m_{l-1} \le u(t) < m_l \\ \dfrac{m_{l+1} - u(t)}{m_{l+1} - m_l} & \text{for } m_l \le u(t) < m_{l+1} \\ 0 & \text{Otherwise} \end{cases}$$

$$f_1(u(t)) = \begin{cases} \dfrac{m_2 - u(t)}{m_2 - m_1} & \text{for } m_1 \le u(t) < m_2 \\ 0 & \text{Otherwise} \end{cases}$$

$$f_M(u(t)) = \begin{cases} \dfrac{u(t) - m_{M-1}}{m_M - m_{M-1}} & \text{for } m_{M-1} < u(t) \le m_M \\ 0 & \text{Otherwise} \end{cases}$$

In each segment of the m-axis in Figure 2, the resulting linear function as indicated by the (red/shaded) dashed line is defined by two overlapping triangle functions in the segment, as indicated by the two (blue) dotted lines. This piecewise linear parametrization using triangle basis functions can approximate not only a smooth but also a nonsmooth static nonlinearity function $f_0(\cdot)$. Let $\hat{x}(t,\mu) = f(\hat{u}(t), \mu)$ be the approximation of $x(t)$, where $\hat{u}(t)$ is the noise free input and μ is the amplitude parameter

$$\mu = \left[\ \mu_1 \quad \cdots \quad \mu_M \ \right]^T. \tag{5}$$

Then, $\hat{x}(t,\mu)$ can be written as

$$\hat{x}(t,\mu) = \rho(\hat{u}(t))\mu \tag{6}$$

where $\rho(\hat{u}(t))$ is defined as

$$\rho(\hat{u}(t)) = \left[\cdots \ 0 \ \frac{m_{k+1} - \hat{u}(t)}{m_{k+1} - m_k} \ \frac{\hat{u}(t) - m_k}{m_{k+1} - m_k} \ 0 \ \cdots \right] \tag{7}$$
$$\text{for } m_k \le \hat{u}(t) < m_{k+1}$$

where m_k and m_{k+1} are the center locations of the triangle basis functions. Let $G(q,\phi)$ be the estimation of $G_0(q)$ with the system parameter

$$\phi = [a_1 \ \cdots \ a_{n_a} \ b_0 \ \cdots \ b_{n_b}] \tag{8}$$

such that

$$\begin{aligned} &G(q,\phi) = q^{-td}\frac{B(q,\phi)}{A(q,\phi)} \\ &\text{where} \\ &A(q,\phi) = 1 + a_1q^{-1} + \cdots + a_{n_a}q^{-n_a}, \\ &B(q,\phi) = b_0 + b_1q^{-1} + \cdots + b_{n_b}q^{-n_b}. \end{aligned} \tag{9}$$

With the parameters ϕ in (8) and μ in (5), the noise free OE model output $\hat{y}(t)$ now can be written as

$$\hat{y}(t,\phi,\mu) = \frac{B(q,\phi)\hat{x}(t-td,\mu)}{A(q,\phi)} = \frac{B(q,\phi)\rho(\hat{u}(t-td))\mu}{A(q,\phi)}. \tag{10}$$

Realizing that $\rho(\hat{u}(t-td))\mu$ in (10) is a linear combination of the time shifted noise free input signal weighted by $\mu_k, k = 1, \cdots, M$, it can be verified that $B(q,\phi)\rho(\hat{u}(t-td))\mu$ in (10) can be written in a linear combination of time shifted inputs weighted by the parameter

$$\tilde{\theta} = [b_0\mu_1 \ \cdots \ b_0\mu_M \ \cdots \ b_{n_b}\mu_1 \ \cdots \ b_{n_b}\mu_M]^T.$$

In this parametrization, an arbitrary gain may be distributed between the static nonlinearity and the linear dynamic system [17] [18]. In order to avoid an ambiguous gain, the scaling of either the linear dynamic system or the static nonlinearity can be fixed. In this paper, we choose to normalize the scaling of the linear dynamic system by fixing $b_0 = 1$. As a result, we will define the parameter

$$\theta = [a_1 \ \cdots \ a_{n_a} \ \mu_1 \ \cdots \ \mu_M \ \cdots \ b_{n_b}\mu_1 \\ \cdots \ b_{n_b}\mu_M]^T \in R_{s\times 1}, \ s = n_a + M \cdot (n_b + 1). \tag{11}$$

as the parameter to be identified, leading to the shorthand notation

$$\hat{y}(t,\theta) = \frac{T(q, \hat{u}(t-td), \theta)}{A(q,\theta)}. \tag{12}$$

With the chosen system parameter θ in (11), the output error is defined as

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t,\theta). \tag{13}$$

A. IV estimation

An output error (OE) model requires a nonlinear optimization (iterative search) due to the non-convexity of $\varepsilon(t,\theta)$ in (13). Let

$$E(\theta) = [\varepsilon(1,\theta) \ \cdots \ \varepsilon(N,\theta)]^T \tag{14}$$

where $\varepsilon(t,\theta)$ is given in (13). Then, the parameter estimation is given by

$$\begin{aligned} \hat{\theta}^N_{OE} &= \arg\min_{\theta} V^N(\theta) \\ V^N(\theta) &= \frac{1}{2N} E^T(\theta)E(\theta). \end{aligned} \tag{15}$$

The next theorem shows that the nonlinear optimization problem in (15) can be rewritten as an IV estimation problem.

Theorem 1: Let the filtered *regressor* Φ^L be defined as

$$\Phi^L = \left[\begin{array}{cc} \Phi^L_a & \Phi^L_{b\mu} \end{array} \right] \tag{16}$$

where

$$\Phi^L_a = \left[\begin{array}{ccc} -y^L(0) & \cdots & -y^L(1-n_a) \\ -y^L(1) & \cdots & -y^L(2-n_a) \\ \vdots & \ddots & \vdots \\ -y^L(N-1) & \cdots & -y^L(N-n_a) \end{array} \right] \tag{17}$$

and

$$\Phi^L_{b\mu} = \left[\begin{array}{ccc} \rho^L(\hat{u}(1-td)) & \cdots & \rho^L(\hat{u}(1-td-n_b)) \\ \vdots & \vdots & \vdots \\ \rho^L(\hat{u}(N-td)) & \cdots & \rho^L(\hat{u}(N-td-n_b)) \end{array} \right]. \tag{18}$$

$y^L(t)$ in (17) is defined as

$$y^L(t) = L(q,\theta)y(t), \ L(q,\theta) = \frac{1}{A(q,\theta)}$$

where $A(q,\theta)$ is given in (9). In (18), $\rho^L(\hat{u}(t)) = L(q)\rho(\hat{u}(t))$, where $\rho(\hat{u}(t)$ is given in (7) and $\hat{u}(t)$ is given in (3) as the noise free input. Subsequently, let the *instrument* $\psi^T(\theta)$ be defined as

$$\psi^T(\theta) = -\frac{dE(\theta)}{d\theta}^T \tag{19}$$

where $E(\theta)$ is given in (14) and the filtered output vector Y^L is defined as

$$Y^L = [y^L(1) \ \cdots \ y^L(N)]^T. \tag{20}$$

Then, the minimizing argument in (15) can be written as

$$\widehat{\theta}^N_{OE} = (\psi^T(\theta)\Phi^L)^{-1}\psi^T(\theta)Y^L.$$

Proof: The minimum of $V^N(\theta)$ in (15) can be obtained by solving

$$\frac{dV^N(\theta)}{d\theta} = \frac{1}{N}E^T(\theta)\frac{dE(\theta)}{d\theta} = \vec{0} \tag{21}$$

where

$$\frac{dV^N(\theta)}{d\theta} = \left[\frac{dV^N(\theta)}{d\theta_1} \cdots \frac{dV^N(\theta)}{d\theta_s}\right]$$

and $\vec{0}$ represents a zero vector. With (19), (21) can be rewritten as

$$\frac{1}{N}E^T(\theta)\psi(\theta) = \vec{0}. \tag{22}$$

Since $E(\theta) = Y - \hat{Y}(\theta)$, where $Y = [y(1) \ \cdots \ y(N)]^T$ and $\hat{Y} = [\hat{y}(1,\theta) \ \cdots \ \hat{y}(N,\theta)]^T$, (22) can be rewritten as

$$\psi^T(\theta) \cdot [Y - \hat{Y}(\theta)] = \vec{0}. \tag{23}$$

(23) can be rewritten as

$$\psi^T(\theta) \cdot [Y^L_A - \hat{Y}^L_T(\theta)] = \vec{0} \tag{24}$$

where

$$\begin{aligned} Y^L_A &= \left[\begin{array}{ccc} Y^L_A(1) & \cdots & Y^L_A(N) \end{array} \right]^T, \\ Y^L_A(t) &= A(q,\theta)y^L(t) \end{aligned}$$

and

$$\begin{aligned} \hat{Y}^L_T(\theta) &= \left[\begin{array}{ccc} \hat{Y}^L_T(1,\theta) & \cdots & \hat{Y}^L_T(N,\theta) \end{array} \right]^T, \\ \hat{Y}^L_T(t,\theta) &= T^L(q, \hat{u}(t-td), \theta). \end{aligned} \tag{25}$$

$T^L(q, \hat{u}(t-td), \theta)$ in (25) is filtered $T(q, \hat{u}(t-td), \theta)$ in (12) with the filter $L(q,\theta)$. Finally, (24) can be written as

$$\psi^T(\theta) \cdot [Y^L - \Phi^L\theta] = \vec{0}$$

where Φ^L is given in (16), $\psi^T(\theta)$ is given in (19), θ is given in (11), and Y^L is given in (20). As a result, $\widehat{\theta}^N_{OE}$ in (15) can be written in IV expression characterized as

$$\widehat{\theta}^N_{OE} = (\psi^T(\theta)\Phi^L)^{-1}\psi^T(\theta)Y^L. \tag{26}$$

■

B. Calculation of the instrument

In this section, the instrument $\psi^T(\theta)$ in (19) is calculated as

$$\psi^T(\theta) = -\frac{dE(\theta)}{d\theta}^T = \frac{d\hat{Y}(\theta)}{d\theta}^T$$

where

$$\frac{d\hat{Y}(\theta)}{d\theta} = \left[\begin{array}{ccc} \frac{d\hat{y}(1,\theta)}{d\theta_1} & \cdots & \frac{d\hat{y}(N,\theta)}{d\theta_1} \\ \vdots & \cdots & \vdots \\ \frac{d\hat{y}(1,\theta)}{d\theta_s} & \cdots & \frac{d\hat{y}(N,\theta)}{d\theta_s} \end{array} \right]^T.$$

The following lemma summarizes the calculation of the instrument $\psi^T(\theta)$.

Lemma 1: The instrument $\psi^T(\theta)$ in (19) is defined by

$$\psi^T(\theta) = [\psi_a \ \psi_{b\mu}]$$

where

$$\psi_a = \begin{bmatrix} -\hat{y}^L(0) & \cdots & -\hat{y}^L(1-n_a) \\ -\hat{y}^L(1) & \cdots & -\hat{y}^L(2-n_a) \\ \vdots & \ddots & \vdots \\ -\hat{y}^L(N-1) & \cdots & -\hat{y}^L(N-n_a) \end{bmatrix} \tag{27}$$

$$+ \begin{bmatrix} d\hat{y}_1(0) & \cdots & d\hat{y}_{n_a}(1-td) \\ d\hat{y}_1(1) & \cdots & d\hat{y}_{n_a}(2-td) \\ \vdots & \ddots & \vdots \\ d\hat{y}_1(N-td) & \cdots & d\hat{y}_{n_a}(N-td) \end{bmatrix} \tag{28}$$

and

$$\psi_{b\mu} = \Phi^L_{b\mu} + \begin{bmatrix} d\Phi_1(1) & \cdots & d\Phi_{M\times(n_b+1)}(1) \\ \vdots & \vdots & \vdots \\ d\Phi_1(N) & \cdots & d\Phi_{M\times(n_b+1)}(N) \end{bmatrix}. \tag{29}$$

$\hat{y}^L(t)$ in (27) is defined as $\hat{y}^L(t) = L(q,\theta)\hat{y}(t,\theta)$. $d\hat{y}_i(t)$ in (28) is defined as

$$d\hat{y}_i(t) = \frac{1}{A(q,\theta)} \frac{dT(q,\hat{u}(t),\theta)}{d\theta_i} \tag{30}$$

where $\dfrac{dT(q,\hat{u}(t),\theta)}{da_i}$ in (30) is defined as

$$\frac{dT(q,\hat{u}(t),\theta)}{da_i} = \sum_{k=1}^{(n_b+1)\cdot M} \frac{d\Phi_{b\mu}(t,k)}{d\hat{u}(t-p)} \frac{d\hat{u}(t-p)}{da_i} \tilde{\theta}_k. \tag{31}$$

$\Phi^L_{b\mu}$ in (29) is given in (18) and $d\Phi_i(t)$ in (29) is defined as

$$d\Phi_i(t) = \frac{1}{A(q,\theta)} \sum_{k=1}^{(n_b+1)\cdot M} \frac{d\Phi_{b\mu}(t,k)}{d\hat{u}(t-td-p)} \frac{d\hat{u}(t-td-p)}{d\tilde{\theta}_i} \tilde{\theta}_k. \tag{32}$$

$\dfrac{d\Phi_{b\mu}(t,:)}{d\hat{u}(t-p)}$ in (31) and (32) is defined as

$$\frac{d\Phi_{b\mu}(t,:)}{d\hat{u}(t-p)} = \left[\frac{d\rho(\hat{u}(t))}{d\hat{u}(t)} \ \cdots \ \frac{d\rho(\hat{u}(t-n_b))}{d\hat{u}(t-n_b))} \right]$$

where

$$\frac{d\rho(\hat{u}(t))}{d\hat{u}(t)} = \left[\cdots \ 0 \ \frac{-1}{m_{k+1}-m_k} \ \frac{1}{m_{k+1}-m_k} \ 0 \ \cdots \right]$$

when $m_k \leq \hat{u}(t) < m_{k+1}$, $p = ceil\left(\dfrac{k}{M}\right)$ that rounds the elements of $\dfrac{k}{M}$ to the nearest integers greater than or equal to $\dfrac{k}{M}$, and $\tilde{\theta} = \theta(n_a+1:s)$.

Proof: (i) Calculation of $\left.\dfrac{d\hat{y}(t,\theta)}{d\theta_i}\right|_{i=1:n_a}$. From (12), $\hat{y}(t,\theta) = \dfrac{T(q,\hat{u}(t-td),\theta)}{A(q,\theta)}$. For brevity of notation, hereafter θ will be omitted if it does not lead to confusion in notation. If we take the derivative of $\hat{y}(t)$ with respect to $\theta_{i=1:n_a}$, we obtain

$$\left.\frac{d\hat{y}(t)}{d\theta_i}\right|_{i=1:n_a} = -q^{-i}\frac{T(q,\hat{u}(t-td))}{A(q)^2} + \frac{1}{A(q)}\frac{dT(q,\hat{u}(t-td))}{d\theta_i}. \tag{33}$$

For brevity of notation, let $a_i = \theta_{i=1:n_a}$. Then, with (30), (33) is written as

$$\frac{d\hat{y}(t)}{da_i} = -\hat{y}^L(t-i) + d\hat{y}_i(t-td).$$

With the noise free input $\hat{u}(t)$ in (3) and the controller $C(q)$ in (4), the derivative of $\hat{u}(t)$ with respect to a_i is defined as

$$\frac{d\hat{u}(t)}{da_i} = -\sum_{j=1}^{n_c} c_j \frac{d\hat{u}(t-j)}{da_i} - \sum_{k=0}^{n_d} d_k \frac{d\hat{y}(t-k)}{da_i}. \tag{34}$$

(ii) Calculation of $\dfrac{d\hat{y}(t)}{d\tilde{\theta}_i}$. From (12) and (18),

$$\hat{y}(t,\theta) = \frac{T(q,\hat{u}(t-td),\theta)}{A(q)} = \Phi^L_{b\mu}(t,:)\tilde{\theta}$$

where $\tilde{\theta} = \theta(n_a+1:s)$. If we take the derivative of $\hat{y}(t)$ with respect to $\tilde{\theta}_i$, we obtain

$$\begin{aligned} \frac{d\hat{y}(t)}{d\tilde{\theta}_i} &= \Phi^L_{b\mu}(t,i) + \\ &\frac{1}{A(q)} \sum_{k=1}^{(n_b+1)\cdot M} \frac{d\Phi_{b\mu}(t,k)}{d\hat{u}(t-td-p)} \frac{d\hat{u}(t-td-p)}{d\tilde{\theta}_i} \tilde{\theta}_k \end{aligned} \tag{35}$$

where $\Phi^L_{b\mu}(t,i)$ is an element of $\Phi^L_{b\mu}$ defined in (18). Then, with (32), (35) is written as

$$\frac{d\hat{y}(t)}{d\tilde{\theta}_i} = \Phi^L_{b\mu}(t,i) + d\Phi_i(t).$$

Similar to (34), the derivative of $\hat{u}(t)$ with respect to $\tilde{\theta}_i$ in (35) is defined as

$$\frac{d\hat{u}(t)}{d\tilde{\theta}_i} = -\sum_{j=1}^{n_c} c_j \frac{d\hat{u}(t-j)}{d\tilde{\theta}_i} - \sum_{k=0}^{n_d} d_k \frac{d\hat{y}(t-k)}{d\tilde{\theta}_i}.$$

■

C. Parameter separation

As explained in Section IV, since the static nonlinearity and the dynamic linear system are parameterized independently, the system is over-parametrized by the modified parameter vector given in (11). As the parameter estimate $\hat{\theta}^N_{OE}$ in (26) has the same structure, we need to separate the parameters of the linear dynamic system in $\eta = [\ a_1 \cdots a_{n_a} \quad b_1 \cdots b_{n_b}\ ]^T$, and the parameters for the piecewise linear approximation of $f(\cdot)$ in $\mu = [\ \mu_1 \ \cdots \ \mu_M\ ]^T$. Once $\hat{\theta}^N_{OE}$ in (26) is estimated, $\hat{a} = [\hat{a}_1 \ \cdots \ \hat{a}_{n_a}]^T$ is easily obtained. The parameter vectors $\hat{b} = \left[\hat{b}_1 \ \cdots \ \hat{b}_{n_b}\right]^T$ and $\hat{\mu}$ can be separated using the singular value decomposition (SVD) [3]. First the parameter vector $\hat{\theta}_{OE}^{N}$ is reorganized into $\Gamma_{b\mu}$ given by

$$\Gamma_{b\mu} = \begin{bmatrix} \hat{\mu}_1 & \cdots & \hat{\mu}_M \\ \vdots & \vdots & \vdots \\ \hat{b}_{n_b}\hat{\mu}_1 & \cdots & \hat{b}_{n_b}\hat{\mu}_M \end{bmatrix}.$$

The singular value decomposition of $\Gamma_{b\mu}$ is given as $\Gamma_{b\mu} = U\Sigma V^T$, where $U_{(n_b+1)\times(n_b+1)}$ and $V_{M\times M}$ are orthogonal matrices, and $\Sigma_{(n_b+1)\times M}$ is a rectangular diagonal matrix. The positive diagonal entries of Σ are called singular values. With the constraint, $b_0 = 1$, the parameter vectors $\hat{\eta}$ and $\hat{\mu}$ can be calculated by

$$\begin{aligned} \hat{\eta} &= \begin{bmatrix} \hat{a}^T & \hat{b}^T \end{bmatrix} \\ &= \begin{bmatrix} \hat{\theta}_{IV}^{N}(1:n_a)^T & U(:,1)/U(1,1) \end{bmatrix} \\ \hat{\mu}^T &= \sigma_1 V^T(1,:)\cdot U(1,1) \end{aligned} \tag{36}$$

where σ_1 is the largest singular value and $U(1,1)$ denotes the first nonzero element of $U(:,1)$, where the notation (1,:) and (:,1) are used to denote the first row and the first column in a matrix respectively.

D. Iterative IV estimation

Obviously, the IV solution in (26) cannot be used to compute $\hat{\theta}_{OE}^{N}$ as done in the IV estimate, because the right hand side also depends on the solution θ. However, the (parameter dependent) instrument $\psi_k(\theta)$ can be calculated based on the previous parameter estimate $\hat{\theta}_{k-1}^{N}$. Based on information of a previous parameter estimate $\hat{\theta}_{k-1}^{N}$, the actual filtered input/output signals can be computed as

$$\hat{\theta}_k^N = [\psi^T(\theta_{k-1})\Phi^L(\theta_{k-1})]^{-1}\psi^T(\theta_{k-1})Y_{k-1}^L. \tag{37}$$

The above expressions can be combined to summarize the iterative IV procedure to compute an OE parameter estimate $\hat{\theta}_{OE}^{N}$. With an initial parameter estimate $\hat{\theta}_{OE}^{N} = \hat{\theta}_k^N$ to model the static nonlinearity $\hat{f}$ and the linear dynamic system $\hat{G}$, one could employ an iterative solution that consists of the following computational steps:

Step 1 : Separate $\hat{\theta}_k^N$ into $\hat{\eta}$ and $\hat{\mu}$ in (36) and generate noise free input $\hat{y}(t)$ using (10) and noise free output $\hat{u}(t)$ using (3).
Step 2 : Define the filter $L(q, \hat{\theta}_{k-1}^N) = A(q, \hat{\theta}_{k-1}^N)^{-1}$. If the filter $L(q, \hat{\theta}_{k-1}^N)$ is unstable, project the poles outside the unit circle inside the unit circle.
Step 3 : Define Φ_{k-1}^L in (16), ψ_{k-1}^T in (19) , and filtered output vector Y_{k-1}^L in (20)
Step 4 : Compute the IV estimate in (37) $\hat{\theta}_k^N = (\psi_{k-1}^T\Phi_{k-1}^L)^{-1}\psi_{k-1}^T Y_{k-1}^L$
Step 5 : Stopping criterion of the algorithm. If $\|\hat{\theta}_k^N - \hat{\theta}_{k-1}^N\|/\|\hat{\theta}_{k-1}^N\| < \varepsilon$, stop. Otherwise, go to **Step 1**

In the above steps, the stable filter $L(q)$, the filtered output vector Y^L, the filtered regressor Φ^L and the instrument ψ^T are updated using $\hat{\theta}_{k-1}^N$ during the iterations over k. **Step 1** creates the noise free signals generated from closed-loop simulations. In **Step 2** the filter $L(q)$ is updated to provide the correct filtering for signals used in **Step 3**. In **Step 3**, the regressor and the instrument are calculated based on the gradient expression. **Step 4** is the actual computation of the IV estimate and **Step 5** formulates a stopping criterion for the algorithm by looking at the relative parameter error.

V. Numerical Example

In this section, numerical examples ($Case$ 1 and $Case$ 2) using the proposed iterative IV method are presented. The configuration in $Case$ 1 is the same as the example that appeared in [13] except that the controller gain is reduced for the closed-loop system stability. In $Case$ 2, only the input static nonlinearity is replaced by saturation from $Case$ 1 in order to compare the efficiency of the proposed method for different static nonlinearities. An excitation signal $r(t)$ follows a uniform distribution with values between -2 and 2. The output disturbance $v(t)$ is filtered white noise. Twenty sets of estimation data with 2000 samples are generated for the system identification. The $M = 19$ grid points are equally spaced between $min(u(t))$ and $max(u(t))$ to model static nonlinearity for $Case$ 1. $M = 5$, with $m = [min(u(t)) \;\; -1 \;\; 0 \;\; 1 \;\; max\,(u(t))]^T$ is used for $Case$ 2. An 2^{nd} order model with 1 step time delay is used to model the linear dynamic system. The configuration of the simulation is shown in Figure 1.

$$\begin{cases} Case\ 1 \\ \quad f_0(u(t)) = sin(u(t)) - 0.5sin(2u(t)) + 0.4sin(3u(t)) \\ Case\ 2 \\ \quad f_0(u(t)) = \begin{cases} 1 & \text{if} \quad u(t) > 1 \\ u(t) & \text{if} \quad |u(t)| \le 1 \\ -1 & \text{if} \quad u(t) < -1 \end{cases} \\ G_0(q) = \dfrac{0.0997q^{-1} - 0.0902q^{-2}}{1 - 1.8858q^{-1} + 0.9048q^{-2}} \\ C(q) = 0.1\dfrac{10.75 - 9.25q^{-1}}{1 - q^{-1}} \\ H(q) = \dfrac{1 + 0.5q^{-1}}{1 - 0.85q^{-1}} \end{cases}$$

The results of applying the proposed iterative IV identification method to the closed-loop time domain data are shown in Figure 3 and Figure 4. The results show that the proposed method is very efficient in not only identifying nonsmooth static nonlinearity due to the use of triangle basis functions, but also in identifying smooth static nonlinearity.

VI. Conclusions

An iterative IV method minimizing the output error based on gradient expression is proposed to identify a Hammerstein system in a closed-loop setting. This method allows us to solve a nonlinear OE minimization problem as an iterative linear optimization problem. In the proposed IV method, an instrument is calculated with filtered noise-free signals and their gradients where the filter is derived by a priori knowledge of the pole locations of the linear dynamic system, and where the noise free signals are computed from simulated closed-loop input and output signals generated by the known reference signal. For accurate computation of the

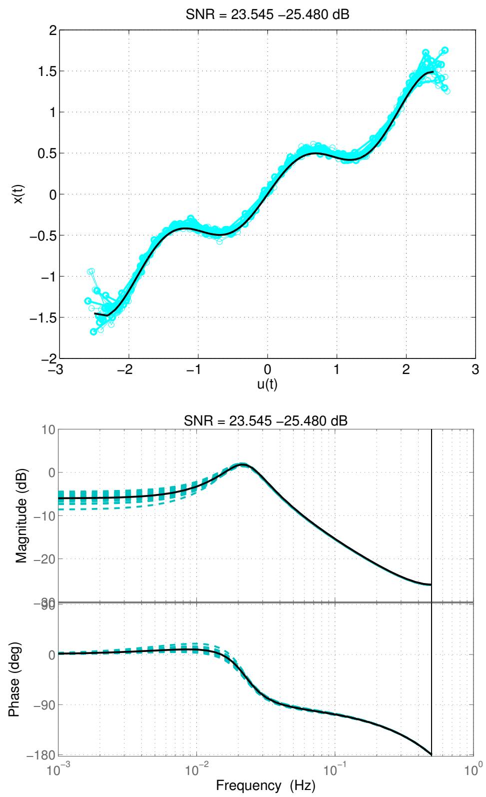

Fig. 3. *Case* 1: The plot of the identified static nonlinearity function (top figure). The Bode plot of the identified linear dynamic system (bottom figure). The black solid line indicates the real Hammerstein system. The (colored) dashed lines indicate estimated systems by using twenty different sets of data. The SNR of each data set spans the values between 23.545 and $25.480dB$.

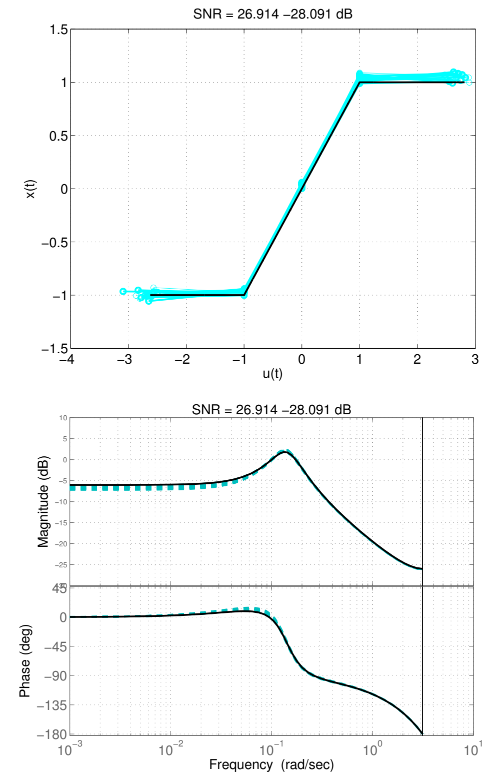

Fig. 4. *Case* 2: The plot of the identified static nonlinearity function (top figure). The Bode plot of the identified linear dynamic system (bottom figure). The black solid line indicates the real Hammerstein system. The (colored) dashed lines indicate estimated systems by using twenty different sets of data. The SNR of each data set spans the values between 26.914 and $28.091dB$.

closed-loop signals and the filter, an iterative procedure that updates the knowledge of the static nonlinearity and the linear dynamic system is used. Convergence of the iterative steps guarantees a local minimum of the OE minimization problem. The simulation study shows the effectiveness of the proposed algorithms in closed-loop identification of Hammerstein systems.

References

[1] K. Narendra and P. Gallman, “An iterative method for the identification of nonlinear systems using a Hammerstein model,” *IEEE Transactions on Automatic Control*, vol. 11, pp. 546–550, 1966.

[2] F. Chang and R. Luus, “A noniterative method for identification using Hammerstein models,” *IEEE Transactions on Automatic Control*, vol. 16, pp. 464–468, 1971.

[3] E. W. Bai, “An optimal two-stage identification algorithm for Hammerstein-wiener nonlinear systems,” *Automatica*, vol. 34, pp. 333–338, 1998.

[4] N. D. Haist, F. Chang, and R. Luus, “Nonlinear identification in the presence of correlated noise using a Hammerstein model,” *IEEE Transactions on Automatic Control*, vol. 18, pp. 553–555, 1973.

[5] T. Hsia, *System Identification, Least-Squares Methods*. Lexington, MA: Lexington Books, 1977.

[6] R. Haber, “Parametric identification of nonlinear dynamic systems based on nonlinear crosscorrelation functions,” in *IEE Proceedings*, Pt. D, No.6, 1988.

[7] W. Greblicki and M. Pawlak, “Identification of discrete Hammerstein systems using kernel regression estimates,” *IEEE Trans. Automatic Control*, vol. 31, pp. 74– 77, 1986.

[8] F. Giri and E. Bai, *Block Oriented Nonlinear System Identification*. Berlin: Springer-Verlag, 2010.

[9] J. Beyer, G. Gens, and J. Wernstedt, “Identification of nonlinear systems in closed loop,” in *Proc. 5th IFAC Symposium on Identification and System Parameter Estimation*, vol. 2, Darmstadt, Germany, 1979, pp. 661– 667.

[10] N. Linard, B. Anderson, and F. De Bruyne, “Closed loop identification of nonlinear systems,” in *Proc. 36th IEEE Conference on Decision and Control*, San Diego, CA, 1997, pp. 2998–3003.

[11] F. De Bruyne, B. Anderson, N. Linard, and M. Gevers, “On closed identification with a tailor-made parametrization,” in *Proc. American Control Conference*, Philadelphia, PA, 1998, pp. 3177–3181.

[12] J. van Wingerden and M. Verhaegen, “Closed loop identification of MIMO Hammerstein models using ls-svm,” in *15th IFAC Symposium on System Identification*, Saint-Malo, France,, 2009.

[13] V. Laurain, M. Gilson, and H. Garnier, “Refined instrumental variable methods for identifying Hammerstein models operating in closed loop,” in *48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 3614–3619.

[14] U. Forssell and L. Ljung, “Identification of unstable systems using output error and box-jenkins model structures,” *IEEE Transactions on Automatic Control*, vol. 45, pp. 137–141, 2000.

[15] R. S. Blom and P. M. J. Van den Hof, “Multivariable frequency domain identification using IV-based linear regression,” in *49th IEEE Conference on Decision and Control*, Atlanta, GA, USA, 2010, pp. 1148–1153.

[16] W. Gibson, *The Method of Moments in Electromagnetics*. Boca Raton, FL: Chapman & Hall/CRC, 2008.

[17] C. Chou and M. Verhaegen, “An indirect approach to closed-loop identification of Wiener models,” in *Proc. American Control Conference*, Philadelphia, PA, 1999, pp. 3451–3455.

[18] T. Wigren, “Recursive prediction error identification using the nonlinear Wiener model,” *Automatica*, vol. 29, pp. 1011–1025, 1993.