MIMO Hammerstein System Identification using LS-SVM and Steady State Time Response

Ricardo Castro-Garcia, Oscar Mauricio Agudelo, Johan A. K. Suyken
ESAT-STADIUS KU Leuven
Leuven, Belgium B-3001
Email: ricardo.castro@esat.kuleuven.be, mauricio.agudelo@esat.kuleuven.be, johan.suykens@esat.kuleuven.be

***Abstract*—A new methodology for identifying Multiple Input Multiple Output (MIMO) Hammerstein Systems is presented in this paper. The method consists of two stages. In the first stage, a Least Squares Support Vector Machine (LS-SVM) is used to model the nonlinear block of the Hammerstein System from its steady-state response. In the second stage, the intermediate variables are computed by using the previously estimated nonlinear block. Then, the linear block is estimated from the latter and the known outputs by using subspace identification methods. The method is very flexible concerning the class of problems it can handle and no previous knowledge about the underlying non-linearities is required except for very mild assumptions. It is particularly effective when dealing with hard to model nonlinearities where other methods often fail. Also, it can handle different numbers of inputs/outputs and performs well in the presence of white Gaussian noise. The performance of the proposed methodology is evaluated through two simulation examples with different levels of noise. The results of this evaluation are compared with those of some state of the art techniques.**

I. Introduction

The field of nonlinear system identification has been widely explored in the last decades. A common approach is to use the block structured nonlinear models introduced in the literature [1], [2]. The Hammerstein model [3] is one of such models consisting of a static nonlinear part $f(\cdot)$, followed by a linear part $G_0(q)$ containing the dynamics of the process (see Fig. 1). An overview of the extensive existing literature around Hammerstein system identification can be found in [4] and different classifications of these methods can be found in [5], [6] and [7].

Hammerstein models have been used to represent nonlinear systems in very diverse areas ranging from control [8] and power amplifiers [9] to signal processing [10], chemical processes [11] and biological processes [12] among others.

EU: The research leading to these results has received funding from the European Research Council under the European Unions Seventh Framework Programme (FP7/2007-2013) / ERC AdG A-DATADRIVE-B (290923). This paper reflects only the authors views and the Union is not liable for any use that may be made of the contained information. Research Council KUL: CoE PFV/10/002 (OPTEC), BIL12/11T; PhD/Postdoc grants Flemish Government: FWO: projects: G.0377.12 (Structured systems), G.088114N (Tensor based data similarity); PhD/Postdoc grant iMinds Medical Information Technologies SBO 2015 IWT: POM II SBO 100031 Belgian Federal Science Policy Office: IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2012-2017)



Fig. 1. SISO Hammerstein system with $G_0(q)$ a linear dynamical system, $f(u(t))$ a static nonlinearity and $v(t)$ the measurement noise. The input and output are $u(t)$ and $y(t)$ respectively and $x(t)$ is the unknown intermediate variable. Throughout this paper, the q-notation will be used. The operator q is a time shift operator of the form $q^{-1}x(t) = x(t-1)$.

Most of the works regarding Hammerstein System Identification are focused on the Single-Input Single-Output (SISO) case while the Multiple-Input Multiple-Output (MIMO) case has received much less attention. Methods dealing with the MIMO case include for instance: In [13] basis functions are used to represent both the linear and nonlinear parts of Hammerstein models; in [14], through the use of specially designed signals, the impulse response of the system is estimated and through least squares the intermediate variables are computed. Using this approximation and the known input, a mapping of the nonlinearity is done through the fitting of a polynomial; an overparametrization approach is proposed in [15] in combination with a reformulated version of LS-SVM, although the MIMO case is not actually tested. Other methods for MIMO Hammerstein system identification can be found in [16], [17] and [18].

The proposed method consists of two stages. In the first one multilevel input signals with a step duration longer than the system's settling time are applied to the process. Next, the levels of the input signals are paired with the steady-state values of each of the outputs. The scalar functions that are part of the nonlinear block (here, it is assumed that the number of intermediate variables is equal to the number of inputs) are approximated from the previously found input-output mappings using LS-SVM. In the second stage, an additional experiment is carried out in order to identify the linear part. Here the input signals are evaluated in the obtained nonlinear scalar functions in order to estimate the intermediate variables. With these estimations and the known outputs of the

978-1-5386-2726-6/17/$31.00 ©2017 IEEE

system, the linear block is identified using subspace methods (e.g., N4SID [19]).

Due to the use of LS-SVM to model the nonlinear part, the proposed method is very flexible regarding the class of systems that can be modeled. For instance, whereas the work in [17] is applicable only to the case where the nonlinearities are in terms of a polynomial, or in [13] specific basis functions have to be chosen beforehand, the methodology presented in this work is free of these limitations thanks to the good generalization properties of LS-SVM.

The proposed method was tested in two examples through several Monte Carlo simulations. It will be illustrated how the measurement noise (white Gaussian noise with zero mean) affects its behavior and also how its accuracy compares with other state of the art methodologies.

The paper is organized as follows. In Section II function estimation using LS-SVM is reviewed. In Section III, the proposed method is presented. Section IV shows the results found when applying the described methodology on two simulation examples. Finally, in Section V, the conclusions are exposed.

II. Function estimation using LS-SVM

LS-SVM [20] has been proposed within a primal-dual formulation framework. Having a data set $\{\boldsymbol{u}_i, x_i\}_{i=1}^N$, the objective is to find a model

$$\hat{x} = \boldsymbol{w}^T \boldsymbol{\varphi}(\boldsymbol{u}) + b. \tag{1}$$

Here, $\boldsymbol{\varphi}(\cdot) : \mathbb{R}^p \rightarrow \mathbb{R}^{n_h}$ is the feature map to a high dimensional (possibly infinite) feature space, $\boldsymbol{w} \in \mathbb{R}^{n_h}$ is the weight vector, $\boldsymbol{u} \in \mathbb{R}^p$ is the input (for an input with p features), $\hat{x} \in \mathbb{R}$ represents the estimated output value, and b is the bias term.

A constrained optimization problem is then formulated:

$$\begin{aligned} \min_{\boldsymbol{w},b,e_i} \quad & \tfrac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \tfrac{\gamma}{2}\sum_{i=1}^{N} e_i^2 \\ \text{subject to} \quad & x_i = \boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{u}_i) + b + e_i,\ i = 1, ..., N, \end{aligned} \tag{2}$$

with e_i the errors and γ the regularization parameter.

From the Lagrangian $\mathcal{L}(\boldsymbol{w}, b, e_i; \alpha_i) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \gamma\frac{1}{2}\sum_{i=1}^N e_i^2 - \sum_{i=1}^N \alpha_i(\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{u}_i)+b+e_i-x_i)$ with $\alpha_i \in \mathbb{R}$ the Lagrange multipliers, the optimality conditions are derived:

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial \boldsymbol{w}} = 0 \rightarrow \boldsymbol{w} = \sum_{i=1}^N \alpha_i \boldsymbol{\varphi}(\boldsymbol{u}_i) \\ \frac{\partial \mathcal{L}}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i = 0 \\ \frac{\partial \mathcal{L}}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i, i = 1, ..., N \\ \frac{\partial \mathcal{L}}{\partial \alpha_i} = 0 \rightarrow x_i = \boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{u}_i) + b + e_i, i = 1, ..., N. \end{cases} \tag{3}$$

Using Mercer's theorem [21], the kernel matrix $\boldsymbol{\Omega}$ can be represented by $\boldsymbol{\Omega}_{ij} = \boldsymbol{K}(\boldsymbol{u}_i, \boldsymbol{u}_j) = \boldsymbol{\varphi}(\boldsymbol{u}_i)^T\boldsymbol{\varphi}(\boldsymbol{u}_j)$ with $i, j = 1, ..., N$ and $\boldsymbol{K}$ a positive definite kernel function. In this representation $\boldsymbol{\varphi}(\cdot)$ does not necessarily have to be explicitly known as it is implicitly used through the positive definite kernel function. In this paper the radial basis function kernel is used:

$$K(\boldsymbol{u}_i, \boldsymbol{u}_j) = \exp\left(\frac{-\|\boldsymbol{u}_i - \boldsymbol{u}_j\|_2^2}{\sigma^2}\right), \tag{4}$$

where σ is the kernel parameter.

The dual formulation is obtained then from (3) by elimination of $\boldsymbol{w}$ and e_i:

$$\begin{bmatrix} 0 & \mathbf{1}_N^T \\ \mathbf{1}_N & \boldsymbol{\Omega} + \gamma^{-1}\boldsymbol{I}_N \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{x} \end{bmatrix} \tag{5}$$

with $\boldsymbol{x} = [x_1, ..., x_N]^T$ and $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_N]^T$. The resulting model is then:

$$\hat{x}(\boldsymbol{u}) = \sum_{i=1}^{N} \alpha_i \boldsymbol{K}(\boldsymbol{u}, \boldsymbol{u}_i) + b. \tag{6}$$

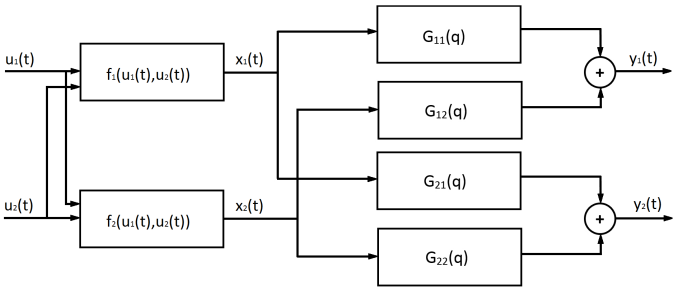

Fig. 2. MIMO Hammerstein system with 2 inputs and 2 outputs.

III. Proposed Method

The proposed methodology is an extension of the work in [22], where the authors offered a method for SISO Hammerstein identification using steady state information.

In this paper it is assumed that the system will have as many intermediate variables as inputs. Additionally, for a system with p inputs it is assumed that $f_i(\mathbf{0}_p) = 0$ for $i = 1, \ldots, p$.

For the sake of clarity and without loss of generality, let us consider a system with 2 inputs $u_1(t)$ and $u_2(t)$, and 2 outputs $y_1(t)$ and $y_2(t)$, as the one shown in Fig. 2. In order to estimate $f_1(\cdot)$ and $f_2(\cdot)$, we first excite the system with multilevel signals $u_1(t)$ and $u_2(t)$ defined as follows:

$$\begin{aligned} u_1(t) &= r_k, \quad \text{for} \quad kT_{C,1} \leq t < (k+1)T_{C,1} \\ u_2(t) &= w_i, \quad \text{for} \quad iT_{C,2} \leq t < (i+1)T_{C,2}. \end{aligned} \tag{7}$$

This means that for each of the steps k and $i \in \mathbb{N}$, $u_1(t)$ has a constant value r_k and $u_2(t)$ has a constant value w_i. $T_{C,1}$ and $T_{C,2}$ are the amount of time that $u_1(t)$ and $u_2(t)$ are kept constant and are defined as follows:

$$\begin{aligned} T_{C,1} &= T_S + \Delta_T \\ T_{C,2} &= N_s T_{C,1}. \end{aligned} \tag{8}$$

Here, T_S is the settling time of the system, Δ_T is an arbitrary additional time and N_s is the number of levels of $u_1(t)$ to be tried out per level of $u_2(t)$. This way of constructing $u_1(t)$ and $u_2(t)$ allows a proper sweep of the possible combinations

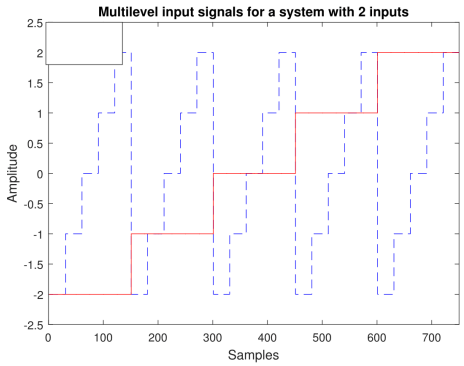

Fig. 3. Multilevel input signals for a system with 2 inputs.

of the inputs. Figure 3 shows an example of how these signals look like. Δ_T guarantees that during each step of $u_1(t)$ some samples will be taken after the system has reached steady state (i.e. those taken during Δ_T after T_S).

Now, let us define the vectors $\tilde{\boldsymbol{u}}_1 \in \mathbb{R}^{N_1 N_2}$ and $\tilde{\boldsymbol{u}}_2 \in \mathbb{R}^{N_2 N_2}$ containing the amplitude levels of the input signals

$$\tilde{\boldsymbol{u}}_1 = \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{N_1-1} \\ r_0 \\ r_1 \\ \vdots \\ r_{N_1-1} \\ \vdots \\ r_0 \\ \vdots \\ r_{N_1-1} \end{bmatrix}, \quad \tilde{\boldsymbol{u}}_2 = \begin{bmatrix} w_0 \\ w_0 \\ \vdots \\ w_0 \\ w_1 \\ w_1 \\ \vdots \\ w_1 \\ \vdots \\ w_{N_2-1} \\ \vdots \\ w_{N_2-1} \end{bmatrix}, \tag{9}$$

where N_1 and N_2 are the number of levels of $u_1(t)$ and $u_2(t)$ respectively. Also, let us define the vectors $\tilde{\boldsymbol{y}}_1 \in \mathbb{R}^{N_1 N_2}$, and $\tilde{\boldsymbol{y}}_2 \in \mathbb{R}^{N_1 N_2}$ where the samples of the outputs $y_1(t)$ and $y_2(t)$ taken during $kT_{C,1} + T_S < t < (k+1)T_{C,1}$ are averaged for each k in order to minimize the effect of the measurement noise during each step

$$\begin{aligned} \tilde{\boldsymbol{y}}_{1,k} &= \frac{1}{N_{\Delta_T}} \sum_{t=kT_{C,1}+T_S}^{kT_{C,1}+T_S+\Delta_T} y_1(t), \\ \tilde{\boldsymbol{y}}_{2,k} &= \frac{1}{N_{\Delta_T}} \sum_{t=kT_{C,1}+T_S}^{kT_{C,1}+T_S+\Delta_T} y_2(t), \end{aligned} \tag{10}$$

for $k = 1, \ldots, N_1 N_2$ and with N_{Δ_T} the number of samples taken during Δ_T.

Using $\tilde{\boldsymbol{u}}_1$ and $\tilde{\boldsymbol{u}}_2$ as inputs and $\tilde{\boldsymbol{y}}_1$ as an output, an LS-SVM model can be trained to approximate the first nonlinearity $\hat{f}_1(\cdot)$ of the system. In a similar fashion, using $\tilde{\boldsymbol{u}}_1$ and $\tilde{\boldsymbol{u}}_2$ as inputs and $\tilde{\boldsymbol{y}}_2$ as an output, another LS-SVM model can be trained to approximate the second nonlinearity of the system $\hat{f}_2(\cdot)$ (See Fig. 4).

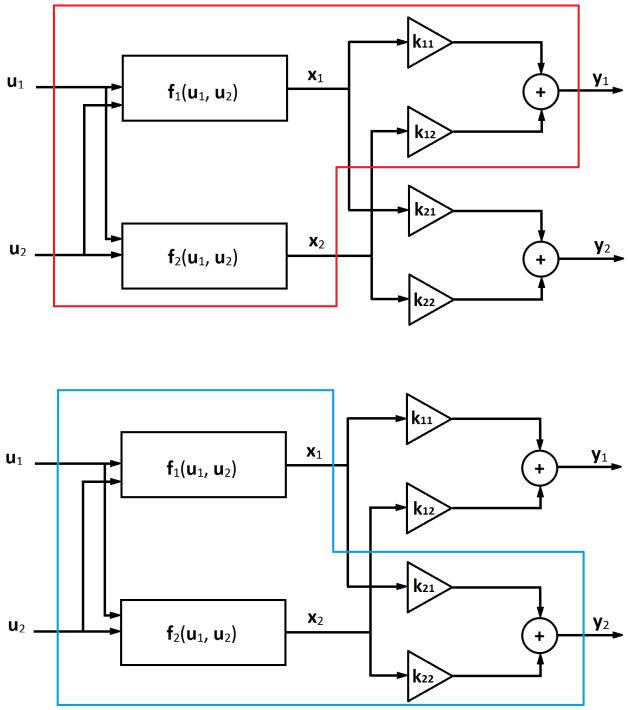

Fig. 4. Modeling of the nonlinear block of a system with two inputs and two outputs. (Red) Nonlinearity corresponding to the output $y_1(t)$. (Blue) Nonlinearity corresponding to the output $y_2(t)$.

Notice that

$$\begin{aligned} \hat{f}_1(\cdot) &= k_{11} f_1(\cdot) + k_{12} f_2(\cdot) \\ \hat{f}_2(\cdot) &= k_{21} f_1(\cdot) + k_{22} f_2(\cdot), \end{aligned} \tag{11}$$

where k_{11}, k_{12}, k_{21} and k_{22} are the steady state gains of $G_{11}(q)$, $G_{12}(q)$, $G_{21}(q)$ and $G_{22}(q)$ respectively.

With models corresponding to the nonlinear part available (i.e. $\hat{f}_1(\cdot)$ and $\hat{f}_2(\cdot)$), the second stage of the method can take place. On this stage, an independent experiment is performed where the system is fed with inputs $u_{l,1}(t)$ and $u_{l,2}(t)$ and the corresponding outputs $y_{l,1}(t)$ and $y_{l,2}(t)$ are measured. Then, the intermediate variables $\hat{x}_{l,1}(t) = \hat{f}_1(u_{l,1}(t), u_{l,2}(t))$ and $\hat{x}_{l,2}(t) = \hat{f}_2(u_{l,1}(t), u_{l,2}(t))$ are computed. With $\hat{x}_{l,1}(t)$ and $\hat{x}_{l,1}(t)$ and the known outputs $y_{l,1}(t)$ and $y_{l,2}(t)$, subspace methods can be used to obtain a model of the linear block. Note that the number of intermediate variables estimated in this way will be equal to the number of outputs. For non-square systems this implies that the estimated model will have a different internal structure than the actual one.

In Fig. 5, a summary of the method is presented.

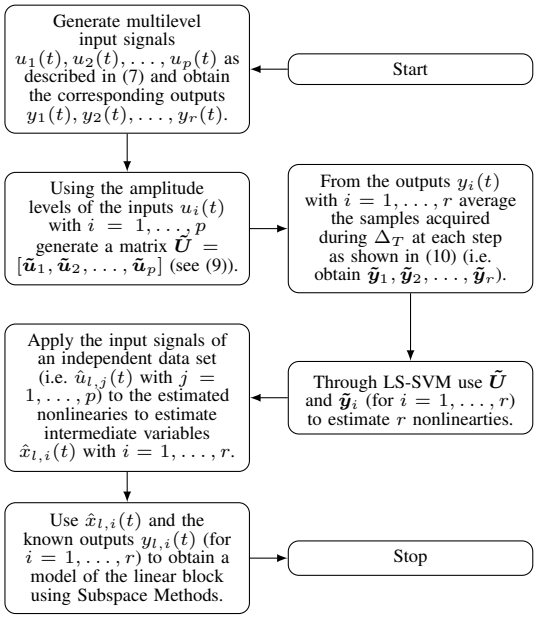

Fig. 5. Summary of the method for a system with p inputs and r outputs.

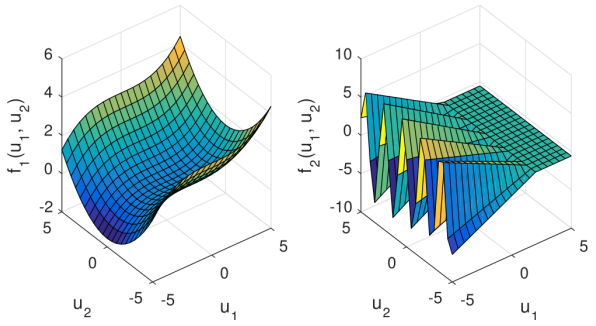

Fig. 6. Nonlinear functions for Example 1. (Left) $f_1(u_1(t), u_2(t))$ and (Right) $f_2(u_1(t), u_2(t))$.

IV. Experimental Results and Comparisons

The proposed method is applied to two examples. Each example has two inputs and two outputs and consists of two nonlinear functions and four Linear Time Invariant (LTI) blocks as illustrated in Fig. 2. The corresponding nonlinear functions of Example 1 are given in (12) and plotted in Fig. 6:

$$f_1(u_1, u_2) = \frac{u_1^3}{80} + \frac{0.9u_2^2}{8} \tag{12a}$$

$$f_2(u_1, u_2) = \begin{cases} (|u_1 - 2| \sin(3(u_2 - 2))) + d, & \text{for } u_1 \leq 2 \\ d, & \text{for } u_1 > 2, \end{cases} \tag{12b}$$

with $d = -0.5588$.

The transfer functions of Example 1 are presented in (13) and the magnitude of their frequency response is shown in Fig. 7:

$$G_{11}(q) = \frac{1.813}{q - 0.8187} \tag{13a}$$

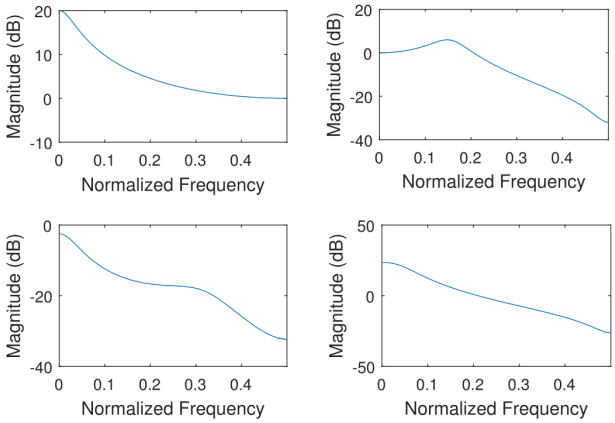

Fig. 7. Magnitude of the frequency response of the LTI blocks for Example 1. (Up and left) $G_{11}(q)$, (Up and right) $G_{12}(q)$, (Down and left) $G_{21}(q)$ and (Down and right) $G_{22}(q)$.

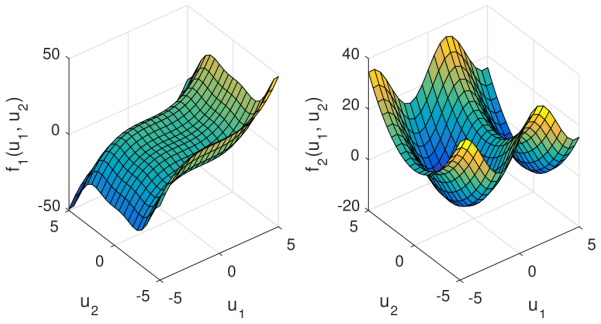

Fig. 8. Nonlinear functions for Example 2. (Left) $f_1(u_1(t), u_2(t))$ and (Right) $f_2(u_1(t), u_2(t))$.

$$G_{12}(q) = \frac{0.3929q + 0.3308}{q^2 - 0.8828q + 0.6065} \tag{13b}$$

$$G_{21}(q) = \frac{-0.045244(q+1.668)(q-1.646)(q+0.2122)}{(q-0.7408)^2(q^2+0.5048q+0.3679)} \tag{13c}$$

$$G_{22}(q) = \frac{0.79928(q + 0.8185)}{(q^2 - 1.452q + 0.5488)}. \tag{13d}$$

For Example 2 the corresponding nonlinear functions are presented in (14) and plotted in Fig. 8:

$$f_1(u_1, u_2) = \frac{u_1^3}{5} + \sin(u_2)u_2^2 \tag{14a}$$

$$f_2(u_1, u_2) = 10 \sin(u_1) + u_2^2. \tag{14b}$$

The transfer functions of Example 2 are given in (15) and the magnitude of their frequency response is shown in Fig. 9:

$$G_{11}(q) = \frac{100q^3 + 300q^2 + 300q + 100}{q^3 - 2.458q^2 + 2.262q - 0.7654} \tag{15a}$$

$$G_{12}(q) = \frac{18000q^2 - 32400q + 14400}{q^2 - 1.5q + 0.7225} \tag{15b}$$

$$G_{21}(q) = 1000\frac{q^4 - 1.884q^3 + 2.506q^2 - 1.884q + 1}{q^4 - 2.485q^3 + 2.528q^2 - 1.184q + 0.2245} \tag{15c}$$

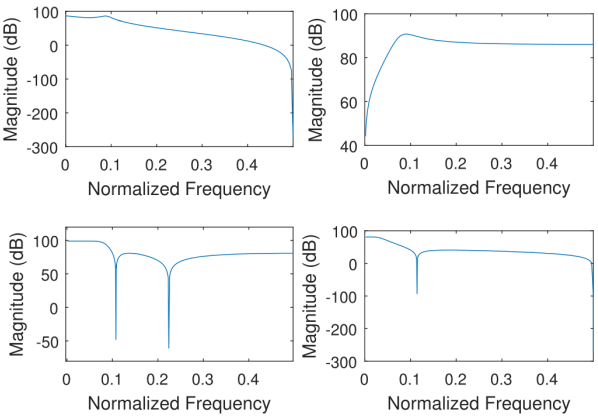

Fig. 9. Magnitude of the frequency response of the LTI blocks for Example 2. (Up and left) $G_{11}(q)$, (Up and right) $G_{12}(q)$, (Down and left) $G_{21}(q)$ and (Down and right) $G_{22}(q)$.

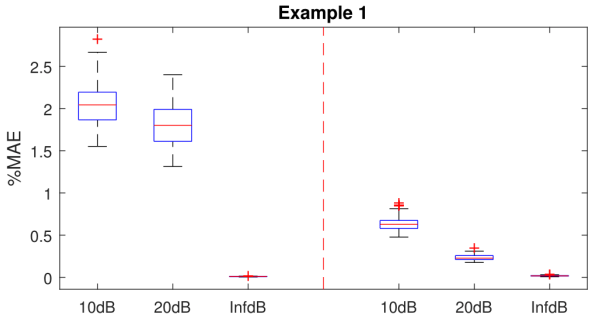

Fig. 10. Results of 100 Monte Carlo simulations of the proposed method in Example 1. (Left) Output 1. (Right) Output 2.

$$G_{22}(q) = \frac{100q^3 - 50.64q^2 - 50.64q + 100}{q^3 - 2.564q^2 + 2.218q - 0.6456}. \tag{15d}$$

In order to be able to make a comparison between the results, let us have the Normalized MAE defined as shown in (16) for a signal with N measurements. Note that the Normalized MAE uses the noise free signal $y_{test}(t)$ (i.e. the actual output) and the estimated output $\hat{y}_{test}(t)$:

$$\%\text{MAE} = \frac{100}{N} \frac{\sum_{t=0}^{N-1} |y_{test}(t) - \hat{y}_{test}(t)|}{|\max(y_{test}(t)) - \min(y_{test}(t))|}. \tag{16}$$

The results of 100 Monte Carlo simulations of the proposed method for different Signal to Noise Ratios (SNR) are presented in Figs. 10 and 11 for Examples 1 and 2 respectively.

The proposed method, from now on referred to as MIMO-H-STST, is compared with 3 other state of the art methods, namely:

- NARX LS-SVM [20].
- The method in [14] where an approximation to the impulse response of the system is obtained and with it and the known outputs an estimation of the intermediate variables is found. Using this approximation and the known inputs, a mapping of the nonlinear block is done through the fitting of multivariate polynomials. From now on, this method will be referred to as IR H-MIMO.
- Using orthonormal bases for the identification of block oriented nonlinear systems is proposed in [13]. This method will be referred to as ONBF.

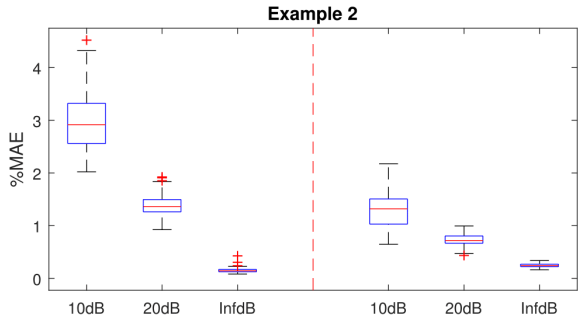

Fig. 11. Results of 100 Monte Carlo simulations of the proposed method in Example 2. (Left) Output 1. (Right) Output 2.

The results of 100 Monte Carlo simulations are summarized in Table I where for each of the methods mentioned above, the median is presented for different Signal to Noise Ratios.

In the proposed method for estimating the nonlinear part 900 points were used. To obtain those points, the length of the steps with the shortest duration for Example 1 was set to 50 samples, meaning that the whole time series used consisted of 45000 samples. For Example 2 the length of the steps with the shortest duration was fixed to 90 samples, thus the time series consisted of 81000 samples. In both examples Δ_T was set to 10 samples. The linear part was identified from a dataset with 4500 samples generated by applying Pseudo Random Multilevel Signals to the system and using the subspace method N4SID [19]. The model order was selected by looking at the plot of the singular values of the Hankel matrices of the impulse response for different orders (from 1 to 10).

For the NARX LS-SVM approach a training set was generated using the combination of the amplitudes in the input signals used for the proposed method (i.e. $\tilde{\boldsymbol{U}}$). This means that 900 points were used for training the model. For the parameter tuning, Coupled Simulating Annealing followed by a Simplex approach was used under a 10 fold cross validation scheme. 10 lags of input and 10 lags of output were employed.

Pseudo Random Binary Signals (PRBS) of 800 samples were created in order to identify the linear part when using the IR H-MIMO method. In a first stage $u_1(t)$ was a PRBS and $u_2(t)$ was kept at 0. Then, in a second stage $u_1(t)$ was 0 and $u_2(t)$ was a PRBS. After the impulse responses were estimated, the nonlinear part was modeled. To do this, signals of 980 points were used of which the last 80 where included to make the corresponding linear system overdetermined. The initial 900 points where generated guaranteeing that all combinations of 30 points drawn from a uniform distribu-

TABLE I
$\%MAE$ COMPARISON FOR THE DIFFERENT METHODS TESTED. MEDIANS ARE OFFERED FOR 100 MONTE CARLO SIMULATIONS FOR EACH CASE.

		Example 1		Example 2	
		$\boldsymbol{y}_1$	$\boldsymbol{y}_2$	$\boldsymbol{y}_1$	$\boldsymbol{y}_2$
SNR 10dB	**MIMO-H-STST**	**2.0431**	**0.62828**	**2.9171**	**1.3187**
	NARX LS-SVM	9.2493	3.4636	14.0744	5.9171
	IR H-MIMO	13.1221	20.7751	15.5643	20.2475
	ONBF	12.2418	12.4902	2.954	6.8069
SNR 20dB	**MIMO-H-STST**	**1.8017**	**0.22772**	**1.3635**	**0.71816**
	NARX LS-SVM	5.6842	1.7901	13.6958	3.3421
	IR H-MIMO	10.4575	16.9498	3.1664	7.2392
	ONBF	10.9409	12.363	1.5856	4.3417
SNR InfdB	**MIMO-H-STST**	**0.008942**	**0.017892**	**0.1428**	**0.24264**
	NARX LS-SVM	4.1052	0.9849	13.6734	2.5123
	IR H-MIMO	9.6007	13.8155	0.23985	0.98531
	ONBF	10.7441	12.3495	1.3339	3.8984

tion between -5 and 5 were included. With these signals the nonlinearities were estimated by fitting two-dimensional polynomials with degrees 3 and 7 for Examples 1 and 2 respectively.

Polynomial basis functions were used for identifying the nonlinearity for the ONBF method. For Example 1 until degree 3 and for Example 2 until degree 5. It was found empirically that the use of simpler basis functions yielded better results for the modelling of the linear part, consequently q^{-n} was used. The number of bases used for Example 1 was 10 while for Example 2 was 40. These values were set by trial and error and were the ones that offered a good trade-off between complexity and accuracy. The number of data points used was 1600 for the first example and 3600 for the second one.

For the examples presented, the proposed method clearly outperforms the other methods considered. It is important to highlight that the nonlinearities in the examples used are very difficult to model using polynomial basis functions as they do not belong to the problem class. For the proposed method, which does not require previous knowledge about the problem class, this is not a problem at all.

It can be seen that the proposed method is robust against the type of noise employed, as the results remain good even when adding high levels of noise.

V. CONCLUSIONS

A new methodology for identifying MIMO Hammerstein Systems is presented which exploits the steady-state behavior of the system in order to approximate the nonlinear part. To do this the method profits from the good generalization capabilities of LS-SVM which allow it to deal with hard nonlinearities.

The proposed method is very flexible with respect to the number of inputs and outputs it can handle. However, for non-square systems the estimated model will have a different internal structure than the actual one.

The used examples illustrate that the method has very good generalization capabilities and can work with different problem classes including systems with hard nonlinearities. This constitutes a nice advantage when the class of problem is unknown or is difficult to model with certain basis functions.

It is shown that the proposed method is robust against the type of noise employed as even in the presence of high levels of noise it has a good performance. In fact, for the examples presented it performed better than the other state of the art methods compared in this paper.

REFERENCES

[1] S. Billings and S. Fakhouri, "Identification of systems containing linear dynamic and static nonlinear elements," *Automatica*, vol. 18, no. 1, pp. 15–26, 1982.

[2] R. Haber and H. Unbehauen, "Structure identification of nonlinear dynamic systems – A survey on input/output approaches," *Automatica*, vol. 26, no. 4, pp. 651–677, 1990.

[3] A. Hammerstein, "Nichtlineare Integralgleichungen nebst Anwendungen," *Acta Mathematica*, vol. 54, pp. 117–176, 1930.

[4] F. Giri and E.-W. Bai, Eds., *Block-oriented nonlinear system identification*. Springer, 2010, vol. 1.

[5] E. W. Bai, "Frequency domain identification of Hammerstein models," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 530–542, 2003.

[6] R. Haber and L. Keviczky, *Nonlinear System Identification – Input-Output Modeling Approach*. Springer The Netherlands, 1999, vol. 1.

[7] A. Janczak, *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models: A Block-Oriented Approach*. Springer-Verlag Berlin Heidelberg, 2005, vol. 310.

[8] K. Fruzzetti, A. Palazoglu, and K. McDonald, "Nolinear model predictive control using Hammerstein models," *Journal of process control*, vol. 7, no. 1, pp. 31–41, 1997.

[9] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electronics Letters*, vol. 37, no. 23, pp. 1417 – 1418, 2001.

[10] J. Stapleton and S. Bass, "Adaptive noise cancellation for a class of nonlinear, dynamic reference channels," *IEEE Transactions on Circuits and Systems*, vol. 32, no. 2, pp. 143–150, 1985.

[11] E. Eskinat, S. H. Johnson, and W. L. Luyben, "Use of Hammerstein models in identification of nonlinear systems," *AIChE Journal*, vol. 37, no. 2, pp. 255–268, 1991.

[12] I. Hunter and M. Korenberg, "The identification of nonlinear biological systems: Wiener and hammerstein cascade models," *Biological cybernetics*, vol. 55, no. 2-3, pp. 135–144, 1986.

[13] J. C. Gomez and E. Baeyens, "Identification of block-oriented nonlinear systems using orthonormal bases," *Journal of Process Control*, vol. 14, no. 6, pp. 685–697, 2004.

[14] J.-C. Jeng and H.-P. Huang, "Nonparametric identification for control of mimo hammerstein systems," *Industrial & Engineering Chemistry Research*, vol. 47, no. 17, pp. 6640–6647, 2008.

[15] I. Goethals, K. Pelckmans, J. A. K. Suykens, and B. De Moor, "Identification of MIMO Hammerstein models using Least-Squares Support Vector Machines," *Automatica*, vol. 41, no. 7, pp. 1263–1272, 2005.
[16] M. Verhaegen and D. Westwick, "Identifying MIMO Hammerstein systems in the context of subspace model identification methods," *International Journal of Control*, vol. 63, no. 2, pp. 331–349, 1996.
[17] Y. J. Lee, S. W. Sung, S. Park, and S. Park, "Input test signal design and parameter estimation method for the hammerstein-wiener processes," *Industrial & engineering chemistry research*, vol. 43, no. 23, pp. 7521–7530, 2004.
[18] H. Ll-Duwaish and M. N. Karim, "A new method for the identification of hammerstein model," *Automatica*, vol. 33, no. 10, pp. 1871–1875, 1997.
[19] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory-Implementation-Applications*. Springer Science & Business Media, 1996.
[20] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. World Scientific, 2002.
[21] J. Mercer, "Functions of positive and negative type, and their connection with the theory of integral equations," *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, pp. 415–446, 1909.
[22] R. Castro-Garcia, O. M. Agudelo, K. Tiels, and J. A. K. Suykens, "Hammerstein system identification using LS-SVM and steady state time response," in *Proc. of the 15th European Control Conference*, 2016, pp. 1063–1068.