15th Triennial World Congress, Barcelona, Spain

IDENTIFICATION OF A CLASS OF HAMMERSTEIN NONLINEAR MIMO SYSTEMS

Aldo Balestrino, Andrea Caiti, Luca Sani

DSEA – Dept. Electrical Systems & Automation
University of Pisa – Italy

Abstract: The identification of nonlinear MIMO systems of the Hammerstein kind is discussed in the paper. The procedure is based on the approximation of the linear system dynamics by Laguerre filter banks, and on the approximation of the static nonlinearities by neural networks. The parameters of the approximating structure are identified from input-output noisy data with the modulating functions method. Under some restrictions on the MIMO structure of the system the proposed identification procedure is shown to be convergent and robust to measurement noise. *Copyright © 2002 IFAC*

Keywords: Nonlinear systems, Identification, Neural networks, Laguerre filters.

1. INTRODUCTION

Hammerstein systems are models for nonlinear systems consisting in the series of a static nonlinear element followed by a linear dynamic system. The simplicity of the structure, and the ability to account for several nonlinear phenomena of scientific and engineering interest, has made Hammerstein models a powerful and popular tool, together with their symmetric counterparts, Wiener models, in which the static and dynamic elements are exchanged (Balestrino *et al.* 2001, and references therein, Al-Duwaish 2000, Sun et al. 1999, Bai 1998, Greblicki 1996, Boutayeb and Darouach 1994, Zi-Qiang 1994). One of the problems often encountered in the study of Hammerstein systems is the (parametric or non-parametric) identification of both the nonlinear and linear elements from the measurement of input-output data. Several approaches have been proposed to this aim in the literature (see references above); however, most of the work has been focused on the identification of SISO Hammerstein systems. The case of MIMO has been much less explored, since it has been treated only in the MISO case in Boutayeb and Darouach (1994), and within a subspace approach by Verhaegen and Westwick (1996). In the papers (Balestrino *et al*., 1999, Balestrino and Caiti, 2000), a specific approximation structure for Hammerstein SISO systems has been proposed. In particular, the nonlinear static element has been approximated by a neural network, and the linear dynamics have been approximated by orthogonal functions of the Laguerre kind. Note that with this approach two different aspects have to be discussed. The first is related to the convergence properties of the selected approximation scheme, and it requires an asymptotic analysis as the number of elementary elements in the approximating structure goes to infinity. The second deals with the more mundane problem of determining the set of parameter values that best fit the available data, once a truncated (i.e., with a finite number of parameters) approximation has been chosen. In the mentioned papers it has been shown that the Laguerre/neural structure is convergent to the true underlying Hammerstein system with assumptions that are milder with respect to those requested by several other Authors proposing different approximating structure. Specific algorithms for parameter identification of the mixed Laguerre/neural structure have also been reported, still in the SISO case (Balestrino et al., 2001). In this work, the extension to the MIMO Hammerstein case of the mixed Laguerre/neural structure firstly introduced in Balestrino et al. (1999) is proposed.

While the convergence of the approximation can be demonstrated in the MIMO case just as in the SISO case (and hence it will not be discussed in this paper), the parameter identification algorithms are not so easily extended from SISO to MIMO. In particular, an algorithmic procedure based on the modulating functions (Preising and Ripping 1993) is illustrated in this paper. Modulating functions have already been demonstrated successfull in the parametric identification of SISO Hammerstein systems (Balestrino et al. 2001). The results reported here shows that the extension of the procedure to the MIMO case leads to a correct identification under some restrictive assumptions on the nonlinearity and on the coupling structure of the linear dynamic part. In particular, it is required that the nonlinearity does not present coupling among the input channels (the same structure as in Boutayeb and Darouach (1994); moreover, it is required that the transfer matrix of the linear dynamic element be diagonally dominant at the lower frequencies (i.e., at constant stationary regime the MIMO system reduces to the parallel of several SISO systems). The assumptions on the linear structure are more restrictive of those of Verhaegen and Westwick (1996); however, with respect to that paper, where the nonlinearity is constrained to be or a known polynomial, or to have polynomial structure with a priori information (as dominantly quadratic), the assumptions made in this paper are milder. The paper is organized as follows: in the next section, the problem is formally stated; in section III, the proposed algorithm is described, and conditions for its convergence to the "true" parameter values are given; simulation results are presented in section IV. Finally, discussion and conclusions are given in section V.

2. PROBLEM STATEMENT

A general Hammerstein MIMO system is illustrated in Fig. 1.

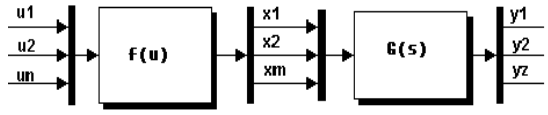


Fig. 1. Structure of a Hammerstein MIMO system.

Let $\mathbf{u} \in \Re^n$ be the input vector, $\mathbf{x} \in \Re^m$ the vector of output of the nonlinear vector function $\mathbf{f}$, $\mathbf{y} \in \Re^z$ the vector of system output. The following general relations hold:

$$\begin{aligned} x_1 &= f_1(u_1, u_2, ..., u_n) \\ &\vdots \\ x_m &= f_m(u_1, u_2, ..., u_n) \end{aligned} \tag{1}$$

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_z \end{bmatrix} = \mathrm{G}(s) \begin{bmatrix} X_1 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} g_{11}(s) & \cdots & g_{1m}(s) \\ \vdots & \vdots & \vdots \\ g_{z1}(s) & \cdots & g_{zm}(s) \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_m \end{bmatrix} \tag{2}$$

Being Y_i, $i = 1,...,z$ the Laplace transform of the i-th component of the vector $\mathbf{y}$, X_j, $j = 1,...,m$ the Laplace transform of the j-th component of the vector $\mathbf{x}$, G(s) the transfer matrix from $\mathbf{x}$ to $\mathbf{y}$. For such a system, the vectors $\mathbf{u}$ and $\mathbf{y}$ are measurable, while the vector $\mathbf{x}$ is not accessible.
The following simplifying assumptions are made:

A1: $z \geq n$, i.e., the system has at least as many output as input channels;

A2: the nonlinear vector function $\mathbf{f}$ is such that $m = n$ and $f_i(\mathbf{u}) = f_i(u_i), i = 1, \cdots, n$;

A3: the transfer functions $g_{ij}(s)$, $i,j = 1,...,n$ have all poles with negative real part;

A4: the functions $f_k, k = 1, \cdots, n$ are all Lipschitz continuous over their domain of definition, and the trivial case $f_k(u) = const.\ \forall u$ is excluded;

A5: the transfer matrix G is diagonally dominant at low frequencies; by this we mean the following:

a) $|g_{ii}(0)| = \bar{g}_{ii} > 0, \quad i = 1, \cdots, n$;

b) $\lim_{\omega \to 0} |g_{ij}(j\omega)| = 0, \ i \neq j$

Assumption *A1* has been made in Boutayeb and Darouach (1994), where the MISO case was treated. In this respect, assumptions *A1-A5* extend the structure of Boutayeb and Darouach (1994) to MIMO systems loosely coupled at low frequency. From a physical point of view, assumption *A5* requires that all the non-diagonal transfer functions in G have a high-pass behaviour. To approximate the Hammerstein system described, the following structure is proposed (Balestrino and Caiti, 2000):

- the nonlinear functions f_k, $k = 1,...,n$ are each approximated by a feedforward neural network with one hidden layer and sigmoidal activation function $\theta(\cdot)$:

$$\begin{aligned} &\hat{f}_k(x_k) = \sum_{j=1}^{n_k} (w^{(k)}{}_{3j} \theta(w^{(k)}{}_{2j} \theta(w^{(k)}{}_{1} x_k))), \\ &\theta(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \quad k = 1, \cdots n \end{aligned} \tag{3}$$

- the transfer functions $g_{ij}(s)$ are each approximated by a series of Laguerre filter with pole $-p_i$:

$$\hat{g}_{ij}(s) = \sum_{k=0}^{m_i} h_{ijk} \frac{(s - p_i)^k}{(s + p_i)^{k+1}} = \frac{\sum_{k=0}^{m_i} b_{ijk} s^k}{\sum_{k=0}^{m_i+1} a_{ik} s^k} \tag{4}$$

$$p_i > 0, \qquad i = 1, \cdots, z; \ j = 1, \cdots, n$$

Note that equation (4) implies that, for any given row i in the transfer matrix G, each transfer function g_{ij} is approximated with the same number m_i of Laguerre filters, all of them with the same pole $-p_i$. Note that the coefficients a_{ik} are known, being the coefficients of the binomial expansion of $(s + p_i)^{m_i+1}$, and that the coefficients b_{ijk} are a linear combination of the Laguerre filter coefficients h_{ijk}. The approximation

proposed for the linear part (equation (4)) implies that the following relation holds:

$$\hat{Y}_i(s)=\sum_{j=1}^{n}\hat{g}_{ij}(s)X_j(s)=\frac{1}{\sum_{k=0}^{m_i+1}a_{ik}s^k}\sum_{j=1}^{n}\left(\sum_{l=0}^{m_i}b_{ijl}s^l X_j(s)\right) \qquad (5)$$

$$i=1,\cdots,z$$

Assuming zero initial conditions, and taking the inverse Laplace transform of equation (5), one gets the following differential equation for the *i*-th output of the linear system:

$$\sum_{k=0}^{m_i+1}a_{ik}\frac{d^k}{dt^k}\hat{y}_i(t)=\sum_{j=1}^{n}\left(\sum_{l=0}^{m_i}b_{ijl}\frac{d^l}{dt^l}x_j(t)\right) \qquad (6)$$

$$i=1,\cdots,z$$

It is finally assumed, without loss of generality, that the transfer functions $g_{ii}(s)$ have unitary static gain for every *i*. This is due to the well-known fact that, for Hammerstein/Wiener systems, the problem of attributing the stationary gain to the nonlinearity or to the linear dynamics is undecidable from observation of input-output data only. The customary choice is to assign a unitary gain to the linear subsystem. In view of this choice, also the approximating linear system is chosen with unitary gain, imposing $b_{ii0} = a_{i0}$ for every *i*. Moreover, due to the high pass nature of the off-diagonal transfer functions in G (assumption A5), the approximating linear system has also to verify the relation $b_{ij0} = 0, j = 1,...,n, j \neq i$.

To summarize, once the number of neural units n_k, the number of Laguerre filters m_i, and the Laguerre poles $-p_i$ have been a priori fixed, the identification problem consists in the determination of the neural network weights $w_{ij}^{(k)}$ and of the Laguerre filter coefficients h_{ijk} (or equivalently b_{ijk}) that best match the available input-output data accordingly to some merit criterion.

3. THE IDENTIFICATION PROCEDURE

The identification algorithm is performed in three main steps. In the first step some values (two are sufficient) of the approximating nonlinear functions $\hat{f}_k$, for $k=1,\cdots,n$, are determined, exploiting the steady state properties of the system. In the second step, these values are used to generate Pseudo Random (Binary) Signals to allow identification of the coefficients of the approximating transfer functions $\hat{g}_{ij}$ one at a time, starting from those on the diagonal. In the third and final step, a set of input-output values of the nonlinearity is generated, to allow generation of network weights by standard learning techniques. In order to identify values of $\hat{f}_k$, a constant vector $\bar{\mathbf{u}}$ is input to the system, and the output is observed for a sufficient time to allow the system to reach steady state (such steady state must exist given the assumptions on the linear dynamics). Once at steady state, the following set of relations must hold:

$$\bar{y}_i=\begin{cases}g_{ii}(0)x_i=g_{ii}(0)f_i(\bar{u}_i)=f_i(\bar{u}_i) & i\le n\\ 0 & i>n\end{cases} \qquad (7)$$

Since in the approximating structure the assumption of unitary gain for each $\hat{g}_{ii}(s)$ has been made, the relation in equation (9) allows to equate the measured output steady-state with the value of the nonlinearity at the point.

The same procedure is then repeated with another constant input vector $\underline{\mathbf{u}}$, component-wise different from $\bar{\mathbf{u}}$; observation of the output at steady-state allows to obtain the relation:

$$\underline{y}_i=\begin{cases}g_{ii}(0)x_i=g_{ii}(0)f_i(\underline{u}_i)=f_i(\underline{u}_i) & i\le n\\ 0 & i>n\end{cases} \qquad (8)$$

After this preliminary procedure, for each nonlinear function $\hat{f}_i$ there are available two known values ($\bar{x}_i$ and $\underline{x}_i$) in correspondence of the input values $\bar{u}_i$ and $\underline{u}_i$ (note that, if any of these values should turn out to be zero, or if they should turn out to be equal, the procedure can be repeated until two values different from each other and from zero are determined). Note also that the reason why the class of MIMO Hammerstein systems has been restricted to those having more output channels than input channel is due to the fact that the above procedure, in the case $n > z$, will fail to establish values for the functions $\hat{f}_k$, $k=z+1,\cdots,n$. The identification of the coefficients $b_{ii\mathrm{k}}$ of the approximating transfer function $\hat{g}_{ii}$ can now be pursued. To this aim, let us now feed the system as *i*-th input a Pseudo Random Binary Signal (PSRBS) $u_i(t)$ that switches between the values $\bar{u}_i$ and $\underline{u}_i$ at every time interval Δt with probability $q = 0.5$; the other system input $u_j(t)$, $j \neq i$, are held constant. Due to the high pass nature of the coupling of the system, the recorded output at the *i*-th channel will only be affected by the transfer function $g_{ii}(s)$. Assuming for $g_{ii}(s)$ the Laguerre filter banks approximation of equation (4), the measured output $y_i(t)$ is obtained as the output of a SISO system described by the following differential equation:

$$\sum_{k=0}^{m_i+1}a_{ik}\frac{d^k}{dt^k}\hat{y}_i(t)=\sum_{l=0}^{m_i}b_{iil}\frac{d^l}{dt^l}x_i(t) \qquad (9)$$

Consider a known function $\Phi(t)$ ("modulating function") defined over the interval [0 T], (with $T = h\Delta T$, h integer), and having the following properties:

- $\Phi(t)$ is differentiable at least $n + 1$ times;
- $\Phi(t)$ and all its first n derivatives have zero values at both the extremes of the interval [0 T].

Multiplying both sides of equation (11) by Φ and integrating by parts on the interval [0 T], one gets:

$$\sum_{k=0}^{m_i+1}(-1)^k a_{ik}\int_0^T \Phi^{(k)}(t)y_i(t)dt = \\ = \sum_{l=0}^{m_i}(-1)^l b_{iil}\int_0^T \Phi^{(l)}(t)x_i(t)dt \tag{10}$$

Since the input signal and the $\Phi(t)$ (together with all its derivatives) are known, and the output signal is measurable, equation (10) establishes an algebraic relation among the coefficients a_i, b_i of the system described by equation (7). Since $x_i(t)$ is a known function switching between the values $\bar{x}_i$ and $\underline{x}_i$, the interval T can be parted in those subintervals in which $x_i(t)$ takes value $\bar{x}_i$, and in those subintervals in which $x_i(t)$ takes value $\underline{x}_i$, so that equation (12) can be re-written as:

$$\sum_{k=0}^{m_i+1}(-1)^k a_{ik}\int_0^T \Phi^{(k)}(t)y_i(t)dt = \\ = \sum_{l=0}^{m_i}(-1)^l b_{iil}\left(\bar{x}_i\sum_{\bar{h}}\int_{\bar{h}}\Phi^{(l)}dt + \underline{x}_i\sum_{\underline{h}}\int_{\underline{h}}\Phi^{(l)}dt\right) \tag{11}$$

where the total sum of the intervals with indexes $\bar{h}$ and $\underline{h}$ is equal to T. In equation (13), the left hand side is known, and the right hand side contains a known term (the one corresponding to the index $l = 0$, since $b_{ii0} = a_{i0}$). By defining the vector of unknown coefficients $\mathbf{z}_{ii} = \left[b_{iim_i}\cdots b_{ii1}\right]^t$, equation (13) can be rewritten as $\mathbf{c}^t\mathbf{z}_{ii} = d$; by repeating the same procedure, generating q different PSRBS $u_i(t)$ over intervals of length T (while always keeping constant the inputs u_j, $j \neq i$), $q > m_i$, one finally gets the following system of algebraic equations in the unknown z_{ii}:

$$\mathrm{C}\,\mathbf{z}_{ii} = \mathbf{d} \tag{12}$$

Equation (15) can then be solved in a least-square sense. The procedure described can be applied for every $i = 1,\cdots,\min(n,z)$, leading to the identification of the coefficients of the approximating transfer functions $\hat{g}_{ii}$. Some observations are in order to guarantee that equation (12) admits indeed solution. The matrix C will be of full rank m_i if at least m_i different PSRBS can be generated within the time interval T, and if, among the chosen PSRBS, at least m_i linearly independent vector of coefficients (see equation (12)):

$$\mathbf{c}_j^t = (-1)^{m_i}\left(\bar{x}_i\sum_{\bar{h}_j}\int_{\bar{h}_j}\Phi^{(m_i)}dt + \underline{x}_i\sum_{\underline{h}_j}\int_{\underline{h}_j}\Phi^{(m_i)}dt\right)\cdots \\ (-1)\left(\bar{x}_i\sum_{\bar{h}_j}\int_{\bar{h}_j}\Phi^{(1)}dt + \underline{x}_i\sum_{\underline{h}_j}\int_{\underline{h}_j}\Phi^{(1)}dt\right) \tag{13}$$

can be generated. The first requirement is met if the switching interval Δh satisfies the relation:

$$\Delta h < \frac{T}{\log_2 m_i} \tag{14}$$

General conditions for the fulfilment of the second requirement are more difficult to establish, depending also on the choice of the modulating function; however, this is a condition that, for a given choice of Φ and for a given sequence of PSRBS, can be checked a priori, to establish if the input choice allows for the identification of the system parameters. If this were not the case, the switching interval Δh (or even the modulating function) can be changed in a trial and error fashion until linear independence is obtained. Once the coefficients of the approximating transfer functions $\hat{g}_{ii}$ have been determined, the same procedure can be applied to the determination of the off-diagonal parameters in $\hat{\mathrm{G}}$: the system is fed as j-th input a Pseudo Random Binary Signal (PSRBS) $u_j(t)$ that switches between the (known) values $\bar{u}_j$ and $\underline{u}_j$, while the other system input $u_k(t)$, $k \neq j$, are held constant at known values. By recording the system outputs $y_i, i = 1,\cdots,z$, the coefficients of the approximating transfer functions $\hat{g}_{ij}$ can be determined with the modulating function method; however, in this case care must be taken to remove from each measured output $y_i(t)$ the constant term due to the steady state behaviour of the transfer function $\hat{g}_{ii}$ with constant input u_i; such term is known, at this stage of the identification process.

The final step in the identification procedure is more standard. In order to determine the coefficients $w_{ij}^{(k)}$ (the synaptic weights in the neural network jargon – see equation (4)) of the approximating functions $\hat{f}_k$, a set of input-output points of the map f is needed (the "training set"). The weights can then be determined by applying some learning procedure, as the backpropagation or one of its many modifications. In order to generate the training set, several solutions are possible. One is to exploit again the steady state properties of the system, feeding constant inputs to the system, and measuring the steady state output. For the assumptions made on the approximating structure, $y_i = \hat{f}_i(u_i), i = 1,\cdots,n$, being y_i the i-th system steady state output, and u_i the constant input on the i-th channel. This procedure allows also to mitigate the effect of zero-mean measurement noise, since it is possible to average the recorded output over arbitrary time intervals after stea1dy state is reached. A possible alternative is to use the approximating transfer matrix $\hat{G}$, which at this stage is known, to invert the linear dynamics; with this option a much wider training set can be collected in a shorter period of time, however this option is much more sensitive to measurement noise and to mismatch between the true system and the approximated one. In our implementation of the method, we have used the first approach (averaged outputs at steady state) for generation of the training set, and the backpropagation with Levenberg-Marquardt correction as the training algorithm for neural approximating structures. The results obtained are reported in the next section.

4. RESULTS

To illustrate the performance of the proposed identification procedure, in this section results are presented for the identification of a simple 2×2 Hammerstein system. The system nonlinearity is given by :

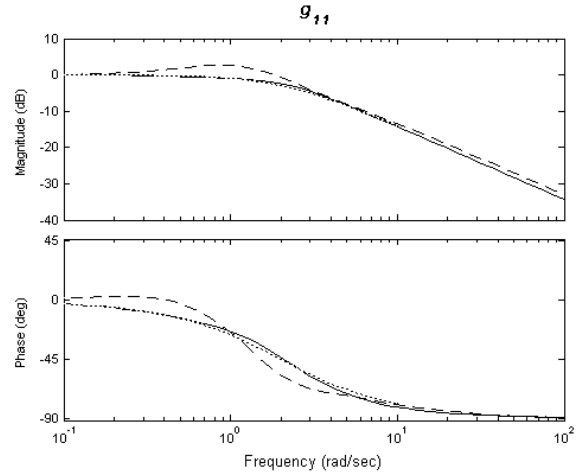
$$\begin{aligned} x_1 &= u_1^2 + 2 \\ x_2 &= u_2^3 + 5 \end{aligned} \tag{15}$$

defined over the interval [-4 4]. The transfer matrix of the linear subsystem is given by:

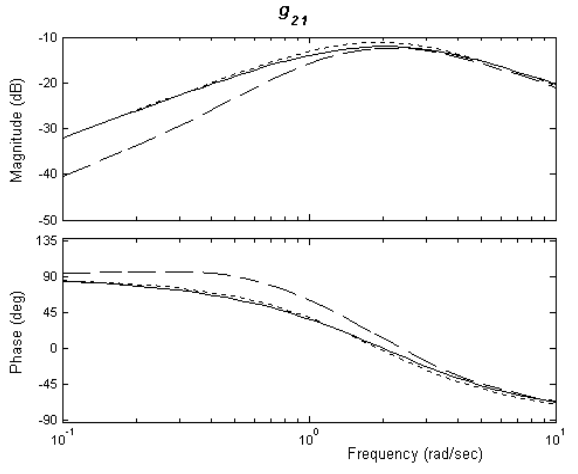
$$G(s) = \begin{bmatrix} \dfrac{2}{s+2} & \dfrac{s}{s^2+3s+4} \\ \dfrac{s}{s^2+4s+4} & \dfrac{s+3}{s^2+3s+3} \end{bmatrix} \tag{16}$$

The approximating structure of the linear subsystem has been chosen as a Laguerre filter banks with 4 elements (i.e., $m_i = 3,\ i = 1,2$), and pole – $p_i = -1.3$. The neural networks have been selected with 4 neurons in the hidden layer. The algorithm has been implemented with the choice of Malentinsky splines as modulating functions (Preising and Rippin 93). The identification procedure has been applied to the case in which the measured output is noise-free, in the case in which the measurement is corrupted with gaussian white noise with a signal-to-noise ratio (SNR) of 40 dB.

Figures 2-3 show the results of the identification of the linear part by comparing the frequency response of the true system and of its approximation. The identified system nonlinearities are shown in figure 4. The results shows a gentle degradation of the approximation performance as the measurement noise increase. Note that the approximation structure has not been changed within the reported test; this means that better results may be obtained by increasing the number of Laguerre filter banks and/or the number of neural units.

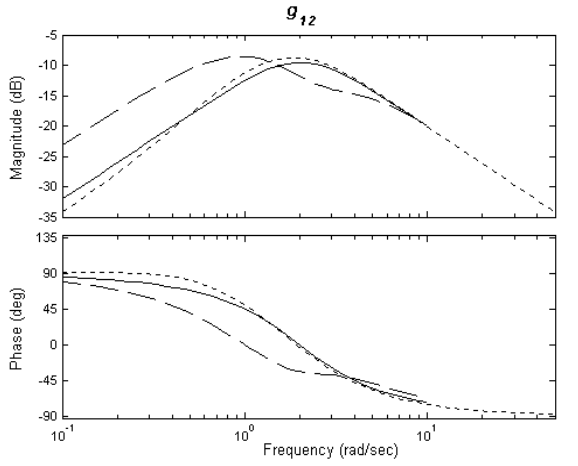


(a)

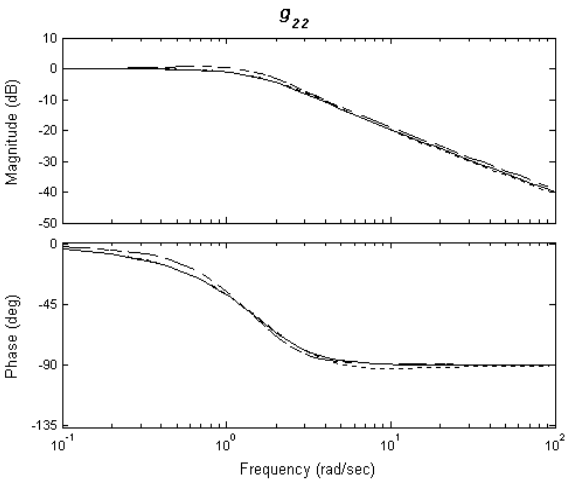


(b)

Fig. 2. Frequency responses. Comparison among true (solid line), approximated without noise (dotted line) and approximated with SNR 40 dB (dashed line). (a) g_{11}, (b) g_{21}.



(a)



(b)

Fig. 3. Frequency responses. Comparison among true (solid line), approximated without noise (dotted line) and approximated with SNR 40 dB (dashed line). (a) g_{12}, (b) g_{22}.

5. CONCLUSIONS

A procedure for the identification of a class of Hammerstein MIMO systems has been proposed. The procedure relies on the approximation of the Hammerstein system with neural networks and Laguerre filter banks. The parameters of the approximating structure can be identified with the method of the modulating function by injecting the system with Pseudo Random Binary Signals, if the class of Hammerstein MIMO systems is restricted to those whose linear dynamics are loosely coupled at low frequencies, and that have at least as many output channels as input channels. The algorithm has an excellent performance in the noise-free case, and has a gentle degradation with the addition of noise, with results that are still fairly acceptable at 20 dB SNR. Variation of the basic algorithm proposed here are of course possible, and numerous to mention. Among the various possibilities that will be subject of further studies, there are the use of generalized orthogonal functions as a substitute of the Laguerre filter banks, the use of different neural structures (for instance, Radial Basis Functions networks).

5. REFERENCES

Al-Duwaish, H.N. (2000). A genetic approach to the identification of linear dynamical systems with static nonlinearities, *Int. J. Sys. Sci.*, vol. 31, n. 3, pp. 307-313.

Bai, E.W. (1998). An optimal two-stage identification algorithm for Hammerstein /Wiener nonlinear systems, *Automatica* vol. 34, n. 3, pp. 333-338.

Balestrino, A., A. Landi, M. Ould-Zmirli and L. Sani (2001). Automatic Nonlinear Auto-Tuning Method for Hammerstein Modeling of Electrical Drives, *IEEE Trans. Ind. Elect.*, vol. 48, pp. 645-655.

Balestrino A. and A. Caiti (2000). Approximation of Hammerstein/Wiener dynamic models, *Proc. IEEE Int. J. Conf. Neural Networks,* Como, Italy.

Boutayeb, M. and M. Darouach (1994), Recursive identification method for the Hammerstein model: extension to nonlinear systems, *Control Theory and Advanced Technology*, vol. 10, pp. 57-72.

Chen, J.-K. and C.C. Yu (1997). Optimal input design of generalized sequences, *Automatica*, vol. 33, n. 11, pp. 2081-2084.

Greblicki W. (1996). Nonlinearity estimation in Hammerstein systems based on ordered observation, *IEEE Trans. Sig. Proc.*, vol. 44, pp. 1224-1233, 1996.

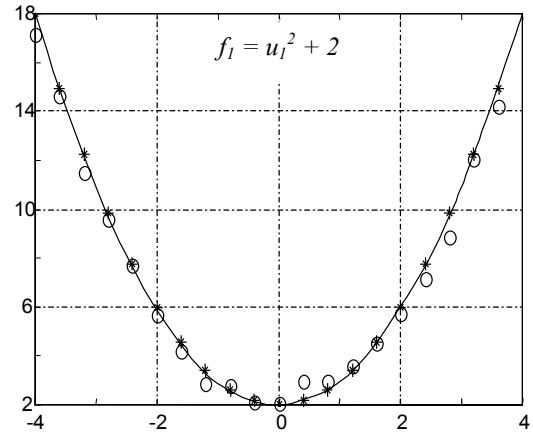
Preising H.A. and D.W.T Rippin (1993). Theory and application of the modulating function method-I. Review and theory of the method and tehory of the spline-type modulating functions, *Computers Chem. Eng.*, vol. 17, n. 1, pp. 1-16.

Sun L., W. Liu and A. Sano (1999). Identification a dynamical system with inout nonlinearity, *IEE Proc. Contr. Th. Appl.*, vol. 146, n. 1, pp. 41-51.

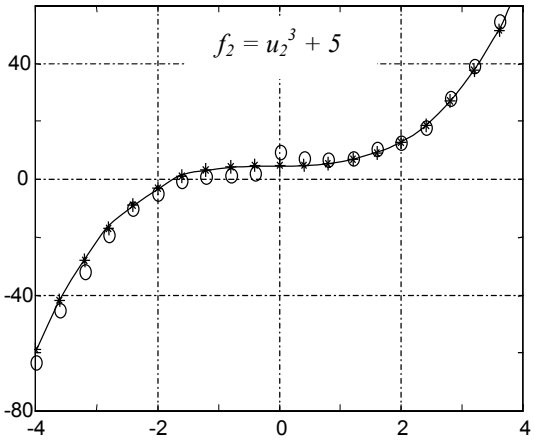
Verhaegen, M. and D. Westwick (1996) Identifying MIMO Hammerstein systems in the context of subspace model identification methods, *Int. J. Contr.*, vol. 63, n. 2, pp. 331-349.

Wahlberg, B. and P.M. Makila (1996). On approximation of stable linear dynamical systems using Laguerre and Kautz functions, *Automatica*, vol. 32, n. 5, pp. 693-708.

Zi-Qiang L. (1994). On the identification of the controlled plants described by the Hammerstein system, *IEEE Trans. Autom. Contr.*, vol. 39, pp. 569-574.



(a)



(b)

Fig. 3. Nonlinear functions. Comparison among true (solid line), approximated without noise (star) and approximated with SNR 40 dB (circle).
(a) $f_1 = u_1^2 + 2$;
(b) $f_2 = u_2^3 + 5$.