

Computational burden reduction in Set-membership Hammerstein system identification

Vito Cerone* Dario Piga* Diego Regruto*

* *Dipartimento di Automatica e Informatica, Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 Torino, Italy.
(emails: vito.cerone@polito.it, dario.piga@polito.it,
diego.regruto@polito.it)*

Abstract: Hammerstein system identification from measurements affected by bounded noise is considered in the paper. First, we show that the computation of tight parameter bounds requires the solution to nonconvex optimization problems where the number of decision variables increases with the length of the experimental data sequence. Then, in order to reduce the computational burden of the identification problem, we propose a procedure to relax the previously formulated problem to a set of polynomial optimization problems where the number of variables does not depend on the measurements sequence size. Advantages of the presented approach with respect to previously published results are discussed.

Keywords: Hammerstein systems, Convex relaxation, Parameters bounds, Set-membership identification.

1. INTRODUCTION

Nonlinear system identification has attracted the attention of many authors in the last decades. In spite of the rich literature (see, e.g., the survey paper by Sjöberg et al. (1995)), modeling and identification of nonlinear dynamic systems still remains a difficult task as highlighted in some recent papers by Ljung (2006, 2008). One of the major challenge is the search for simple and flexible model structure able to cover the most relevant nonlinear phenomena encountered in practice. This problem has stimulated a number of contributions about the identification of block-structured nonlinear systems, which are obtained through interconnection of memoryless nonlinear gains and linear dynamic subsystems (see the recent book edited by Bai and Giri (2010) for an up-to-date collection of results and algorithms). The configuration we are dealing with in this note, commonly referred to as a Hammerstein model, is shown in Fig. 1; it consists of a static nonlinear part \mathcal{N} followed by a linear dynamic system. The identification of such a model relies solely on input-output measurements, while the inner signal x_t , i.e. the output of the nonlinear block, is not assumed to be available. A number of algorithm have been proposed in the literature to address such a problem. Among others we mention the over-parametrization method (Chang and Luus (1971); Hsia (1976); Bai (1998)), the subspace identification Verhaegen and Westwick (1996), the blind approach Bai and Fu (2002), the iterative method (Narendra and Gallman (1966); Bai and Li (2004)), the nonparametric approach (Greblicki and Pawlak (1989); Krzyżak (1993)), the frequency domain method (Bai (2003); Krzyżak (1996)) and the algorithms based on the Bussgang's theorem (see, e.g., Hunter and Korenberg (1986)). In all the papers mentioned above, the authors assume that the measurement

error η_t is statistically described. A worthwhile alternative to the stochastic description of measurement errors is the bounded-error, or set-membership, characterization, where uncertainties are assumed to belong to a given set (see, e.g., Milanese and Vicino (1991); Walter and Piet-Lahanier (1990)). In this paper we consider the identification of single-input single-output (SISO) Hammerstein models when the nonlinear block can be modeled by a linear combination of a finite and known number of nonlinear static functions, the linear dynamic part is described by an output error model with bounded output measurement errors. As shown in Sznaier (2009), the Hammerstein system identification problem in presence of bounded noise is NP-hard in the size of the experimental data sequence. In a previous paper by Cerone and Regruto (2003) a two-stage identification procedure is presented. First, parameters of the nonlinear block are tightly bounded using input-output data collected from the steady-state response of the system to a set of step inputs with different amplitudes. Then, through a dynamic experiment, for all u_t belonging to a given input transient sequence $\{u_t\}$, tight bounds on the inner signal are computed which, together with noisy output measurements are used for bounding the parameters of the linear part. The main limit of the procedure proposed in Cerone and Regruto (2003) is that it requires two different experiments where two specific input signals have to be used. In Cerone et al. (2011) a one-stage procedure without particular constraints on the input signal, which requires the solution to polynomial optimization problems,

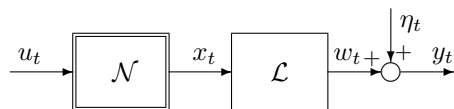


Fig. 1. Hammerstein system.

is proposed. In this paper we present an alternative one-stage approach (referred to as semi-static relaxation) that significantly reduces the computational complexity of the algorithm proposed in Cerone et al. (2011). The paper is organized as follows. Section 2 is devoted to the formulation of the identification problem. In Section 3 we show that computation of tight parameters bounds requires the solution to nonconvex optimization problems where the number of decision variables increases with the number of measured data. A relaxation procedure to reduce the number of decision variables for such optimization problems is presented in Section 4 and an analysis of its properties is reported in Section 5. The simulated example in Section 6 shows the effectiveness of the proposed approach.

2. PROBLEM SETTING

Consider the SISO discrete-time Hammerstein model depicted in Fig. 1. The nonlinear block maps the input signal u_t into the unmeasurable inner variable x_t through the nonlinear function

$$x_t = \sum_{k=1}^{n_\gamma} \gamma_k \varphi_k(u_t), \quad t = 1, \dots, N, \quad (1)$$

where $(\varphi_1, \dots, \varphi_{n_\gamma})$ is a known basis of nonlinear functions and N is the length of data sequence. The linear dynamic block is supposed to be stable and it is modeled by a discrete-time system transforming x_t into the noise-free output w_t according to the equation

$$w_t = - \sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} b_j x_{t-j}. \quad (2)$$

Let y_t be the noise-corrupted output:

$$y_t = w_t + \eta_t. \quad (3)$$

Measurement uncertainty η_t is known to range within given bounds $\Delta\eta_t$, i.e.,

$$|\eta_t| \leq \Delta\eta_t. \quad (4)$$

Unknown parameter vectors $\gamma \in \mathbb{R}^{n_\gamma}$ and $\theta \in \mathbb{R}^{n_\theta}$ are defined, respectively, as

$\gamma = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_{n_\gamma}]^T$, $\theta = [a_1 \ \dots \ a_{na} \ b_0 \ b_1 \ \dots \ b_{nb}]^T$ with $n_\theta = na + nb + 1$. It must be pointed out that the parametrization of the structure in Fig. 1 is not unique. As a matter of fact, any parameters set $\tilde{b}_j = \alpha^{-1} b_j$, $j = 0, 1, \dots, nb$, and $\tilde{\gamma}_k = \alpha \gamma_k$, $k = 1, 2, \dots, n_\gamma$, for some nonzero and finite constant α , provides the same input-output behaviour. Thus, any identification procedure cannot perceive the difference between parameters $\{b_j, \gamma_k\}$ and $\{\alpha^{-1} b_j, \alpha \gamma_k\}$. To get a unique parametrization, we assume, without loss of generality, that the steady-state gain of the linear part be one, i.e.

$$1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j. \quad (5)$$

In this paper we address the problem of deriving uncertainty intervals on the unknown parameters γ and θ .

3. EVALUATION OF TIGHT PARAMETERS BOUNDS

In this section we show that the evaluation of tight parameters uncertainty intervals requires the computation of global optimum solutions to a set of nonconvex optimization problems with $n_\gamma + n_\theta + N$ variables.

The input-output mapping for the Hammerstein model in Fig. 1 can be obtained by substituting (1) in (2), leading to relation

$$w_t = - \sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \varphi_k(u_{t-j}). \quad (6)$$

Therefore, from (3) and (6), we get the following mapping between input signal and the output measurement:

$$y_t = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \varphi_k(u_{t-j}) + \eta_t. \quad (7)$$

Indeed, the set \mathcal{D} of the Hammerstein system parameters γ and θ and noise samples η_t consistent with the measurement data sequence, the assumed model structure and the error bounds is described by (4), (5) and (7), i.e.

$$\mathcal{D} = \left\{ (\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \right.$$

$$y_t = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \varphi_k(u_{t-j}) + \eta_t,$$

$$|\eta_r| \leq \Delta\eta_r, \quad 1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j,$$

$$t = na + 1, \dots, N; \quad r = 1, \dots, N \left. \right\}, \quad (8)$$

with $\eta = [\eta_1, \dots, \eta_N]^T$. Therefore, for each $k = 1, \dots, n_\gamma$ and $j = 1, \dots, n_\theta$, tight bounds on the parameters γ_k and θ_j can be computed by solving the constrained optimization problems

$$\underline{\gamma}_k = \min_{(\gamma, \theta, \eta) \in \mathcal{D}} \gamma_k, \quad \bar{\gamma}_k = \max_{(\gamma, \theta, \eta) \in \mathcal{D}} \gamma_k, \quad (9)$$

$$\underline{\theta}_j = \min_{(\gamma, \theta, \eta) \in \mathcal{D}} \theta_j, \quad \bar{\theta}_j = \max_{(\gamma, \theta, \eta) \in \mathcal{D}} \theta_j. \quad (10)$$

The main features enjoyed by the identification problems (9) and (10) are reported in the following property.

Property 1. Problems (9) and (10) are semialgebraic optimization problems enjoying the following features:

- P 1.1.** the number of optimization variables is $n_\gamma + n_\theta + N$,
- P 1.2.** the objective function is linear,
- P 1.3.** the feasible region \mathcal{D} is defined by $N - na$ bilinear equalities and $2N + 1$ linear constraints. ■

Property 1.3 holds because the constraints in (7) defining \mathcal{D} are bilinear equalities since there is the product between the variable a_i and the noise η_{t-i} , as well as the product between the unknown parameters b_j and γ_k . Because of such bilinear constraints, problems (9) and (10) are nonconvex. Therefore, standard nonlinear optimization tools (gradient method, Newton method, etc.) cannot be used since they can trap in local minima/maxima. Then, the computed uncertainty intervals are not guaranteed to

contain the true parameters, key requirement of any set-membership identification method. One possible solution to overcome such a problem is to relax (9) and (10) into convex optimization problems, in order to numerically compute lower bounds of $\underline{\gamma}_k$ and $\underline{\theta}_j$ as well as upper bounds of $\bar{\gamma}_k$ and $\bar{\theta}_j$. As discussed in Cerone et al. (2011), problems (9) and (10) are semialgebraic with an inherent structured sparsity and they can be relaxed by exploiting the LMI-relaxation techniques proposed in Waki et al. (2006) and Lasserre (2006). In this way, system parameter bounds can be evaluated by solving SDP problems with:

- (i) $(N - na) \binom{n_\gamma + n_\theta + na + 1 + 2\delta}{2\delta} + (N - na - 1) \binom{n_\gamma + n_\theta + na + 2\delta}{2\delta}$ decision variables,
- (ii) $N - na$ LMIs, each one of size $\binom{n_\gamma + n_\theta + na + 1 + \delta}{\delta}$ and $2(N - na) + 2N + 1$ LMIs, each one of size $\binom{n_\gamma + n_\theta + na + \delta}{\delta - 1}$, defining the feasible region of the relaxed problem.

In this paper we present an alternative approach (referred to as semi-static relaxation) to relax (9) and (10) to polynomial optimization problems where only the parameters γ and θ are considered as variables, thus allowing computation of parameter bounds for large values of N .

4. SEMI-STATIC RELAXATION

For the sake of clarity a general overview of the proposed method is first presented in Section 4.1. Then, detailed technical results are given in Section 4.2.

4.1 Overview of the method

The main idea of the semi-static procedure is to relax (9) and (10) by partly neglecting the correlation among consecutive uncertain regressors defining the feasible parameter set \mathcal{D} in (8), leading to an outer approximation of \mathcal{D} . We first provide an alternative description of \mathcal{D} as intersection of $N - na$ sets S_z , each one obtained as follows:

- (i) given a fixed integer n , we consider the equality constraints, defining \mathcal{D} in (8), corresponding to n consecutive measurements;
- (ii) we substitute the first constraint in the second one, leading to a new equality retaining correlations between the two; then, the new obtained equation is substituted in the third constraint in order to retain correlation among the first three equalities selected in (i). The procedure is repeated until all the newly obtained equations are substituted into the n -th considered constraint.

Then, an outer approximation $S_z^{(n)}$ of each set S_z is constructed by assuming that the noise variables η_t appearing in the n equations defining S_z are independent with each other. Thus, a relaxed set $\mathcal{D}^{(n)}$ of \mathcal{D} is obtained as intersection of the $N - na$ sets $S_z^{(n)}$. In such a way, the correlation between n consecutive measurements is not completely lost since it is kept by the nested substitutions in (ii). Thanks to the structure of $\mathcal{D}^{(n)}$, parameter bounds evaluation will be formulated in terms of polynomial optimization problems with $n_\gamma + n_\theta$ variables, instead of the $n_\gamma + n_\theta + N$ variables involved in (9) and (10).

4.2 Technical results

Technical details of the proposed method are here discussed. Proofs of propositions, results and properties reported in the paper can be found in Cerone et al. (2010b). As stated in the previous section, we first rewrite the feasible parameter set \mathcal{D} in (8) in an alternative way, as described by the next proposition.

Proposition 1. Given an integer $n \in [1, N - na]$, let us define the set $S_z^{(n)}$ as

$$S_z^{(n)} = \{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \begin{aligned} y_t &= - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \varphi_k(u_{t-j}) + \eta_t, \\ |\eta_r| &\leq \Delta \eta_r, \quad 1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j, \\ t &= na + z, na + z + 1, \dots, \min\{na + n, N\}; \\ r &= z, z + 1, \dots, \min\{na + n, N\} \}. \end{aligned} \quad (11)$$

Then, \mathcal{D} can be written as

$$\mathcal{D} = \bigcap_{z=1}^{N-na} S_z^{(n)} \quad (12)$$

It is worth noting that the set \mathcal{D} in (12) is written as the intersection of $N - na$ sets $S_z^{(n)}$, each one defined by the constraints describing \mathcal{D} in (8) obtained by at most n consecutive measurements. For instance, $S_1^{(n)}$ is only defined by the constraints obtained by the measurements from time $t = na + 1$ up to time $t = na + n$, $S_2^{(n)}$ is defined by the constraints obtained by the measurements from time $t = na + 2$ up to time $t = na + n + 1$, and so on, up to $S_{N-na}^{(n)}$ that will be defined just by the constraint obtained by the measurement at time N . In this work we refer to n as the dynamic-horizon. We point out that the generic constraint $y_t = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \varphi_k(u_{t-j}) + \eta_t$ can be written in the matrix form $\underline{y}_t = F(\underline{y}_{t-1} - \underline{\eta}_{t-1}) + G\underline{\varphi}_t + \underline{\eta}_t$, where

$$\begin{aligned} \underline{y}_t &\in \mathbb{R}^{na} : \underline{y}_t^T = [y_t, y_{t-1}, \dots, y_{t-na+1}], \\ \underline{\varphi}_t &\in \mathbb{R}^{(nb+1) \cdot n_\gamma} : \underline{\varphi}_t^T = [\varphi_1(u_t), \varphi_1(u_{t-1}), \dots, \varphi_1(u_{t-nb}), \\ &\quad \varphi_2(u_t), \dots, \varphi_{n_\gamma}(u_{t-nb})], \\ \underline{\eta}_t &\in \mathbb{R}^{na} : \underline{\eta}_t^T = [\eta_t, \eta_{t-1}, \dots, \eta_{t-na+1}], \\ F &\in \mathbb{R}^{na, na} : F = \begin{bmatrix} -a_1 & -a_2 & \dots & -a_{na-1} & -a_{na} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \cdot & \cdot & \cdot & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \\ G &\in \mathbb{R}^{na, (nb+1) \cdot n_\gamma} : G = \begin{bmatrix} \gamma_1 b_0 & \dots & \gamma_1 b_{nb} & \gamma_2 b_0 & \dots & \gamma_{n_\gamma} b_{nb} \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \cdot & \cdot & \cdot & \cdot & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}. \end{aligned}$$

An alternative description of sets $S_z^{(n)}$, obtained by nested substitution process described in (ii), is provided by the following proposition in terms of entries of F and G .

Proposition 2. The set $\mathcal{S}_z^{(n)}$ in (11) can be written as

$$\begin{aligned} \mathcal{S}_z^{(n)} &= \{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\ y_{na+z+s-1} &= \sum_{i=1}^{na} a_i^{(s)} (y_{na+z-i} - \eta_{na+z-i}) + \\ &+ \sum_{j=1}^s \langle \underline{g}_j^{(s)}, \underline{\varphi}_{na+z+j-1} \rangle + \eta_{na+z+s-1}, \\ |\eta_r| &\leq \Delta\eta_r, \quad 1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j, \\ s &= 1, 2, \dots, \min\{n, N - z + 1\} \\ r &= z, z + 1, \dots, \min\{na + n, N\} \} \end{aligned} \quad (13)$$

where coefficients $a_i^{(s)}$ as well as the components of the vector $\underline{g}_j^{(s)}$, with $s = 1, 2, \dots, \min\{n, N - k + 1\}$, are polynomial functions of the unknown parameters γ and θ . In particular $a_i^{(s)} = [F^s](1, i)$ and $\underline{g}_j^{(s)} = [F^{s-j}G](1, :)$, with $[R](1, i)$ denoting the entry of the first row and the i -th column of a generic matrix R , $[R](1, :)$ is the first row of R , while $\langle \underline{g}_j^{(s)}, \underline{\varphi}_{na+z+j-1} \rangle$ denotes the inner product between the vectors $\underline{g}_j^{(s)}$ and $\underline{\varphi}_{na+z+j-1}$. ■

Remark 1. Let $q^{(n)} \in \mathbb{R}^{na \cdot n}$ be the collection of the polynomials $a_i^{(s)}$ for $i = 1, \dots, na$ and $s = 1, \dots, n$, i.e.

$$q^{(n)} = [a_1^{(1)}, \dots, a_{na}^{(1)}, a_1^{(2)}, \dots, a_{na}^{(2)}, \dots, a_{na}^{(n)}]^T. \quad (14)$$

Indeed, the components of the vectors $q^{(n)}$ and $\underline{g}_j^{(s)}$, with $j = 1, \dots, s$ and $s = 1, \dots, n$, are polynomial functions in the variables γ and θ of maximum degree equal to $n + 1$. ■

Outer-bounds $\mathcal{S}_z^{ss(n)}$ and $\mathcal{D}^{ss(n)}$ of $\mathcal{S}_z^{(n)}$ and \mathcal{D} , respectively, are constructed as described by the next results.

Result 1. Construction of an outer bound of $\mathcal{S}_z^{(n)}$
Let us define the set $\mathcal{S}_z^{ss(n)}$ as

$$\begin{aligned} \mathcal{S}_z^{ss(n)} &= \{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\ y_{na+z+s-1} &- \sum_{i=1}^{na} a_i^{(s)} y_{na+z-i} - \sum_{j=1}^s \langle \underline{g}_j^{(s)}, \underline{\varphi}_{na+z+j-1} \rangle \\ &\leq \Delta\eta_{na+z+s-1} + \sum_{i=1}^{na} a_i^{(s)} \text{sign}(a_i^{(s)}) \Delta\eta_{na+z-i}; \\ y_{na+z+s-1} &- \sum_{i=1}^{na} a_i^{(s)} y_{na+z-i} - \sum_{j=1}^s \langle \underline{g}_j^{(s)}, \underline{\varphi}_{na+z+j-1} \rangle \\ &\geq -\Delta\eta_{na+z+s-1} - \sum_{i=1}^{na} a_i^{(s)} \text{sign}(a_i^{(s)}) \Delta\eta_{na+z-i}; \\ 1 + \sum_{i=1}^{na} a_i &= \sum_{j=0}^{nb} b_j, \\ s &= 1, 2, \dots, \min\{n, N - z + 1\} \} \end{aligned} \quad (15)$$

Then, for every $z = 1, \dots, N - na$, the set $\mathcal{S}_z^{ss(n)}$ is an outer approximation of $\mathcal{S}_z^{(n)}$, i.e. $\mathcal{S}_z^{(n)} \subseteq \mathcal{S}_z^{ss(n)}$ and the set $\mathcal{D}^{ss(n)}$, defined as $\mathcal{D}^{ss(n)} = \bigcap_{z=1}^{N-na} \mathcal{S}_z^{ss(n)}$ is an outer approximation of the FPS \mathcal{D} , i.e. $\mathcal{D} \subseteq \mathcal{D}^{ss(n)}$. ■

For a given dynamic horizon n , and for every $k = 1, \dots, n_\gamma$ and $j = 1, \dots, n_\theta$, the relaxed parameter uncertainty intervals on γ and θ , defined as

$$PUI_{\gamma_k}^{ss(n)} = [\underline{\gamma}_k^{ss(n)}; \bar{\gamma}_k^{ss(n)}] \quad (16)$$

$$PUI_{\theta_j}^{ss(n)} = [\underline{\theta}_j^{ss(n)}; \bar{\theta}_j^{ss(n)}] \quad (17)$$

can be evaluated by solving the optimization problems

$$\underline{\gamma}_k^{ss(n)} = \min_{\gamma, \theta \in \mathcal{D}^{ss(n)}} \gamma_k; \quad \bar{\gamma}_k^{ss(n)} = \max_{\gamma, \theta \in \mathcal{D}^{ss(n)}} \gamma_k, \quad (18)$$

$$\underline{\theta}_j^{ss(n)} = \min_{\gamma, \theta \in \mathcal{D}^{ss(n)}} \theta_j; \quad \bar{\theta}_j^{ss(n)} = \max_{\gamma, \theta \in \mathcal{D}^{ss(n)}} \theta_j. \quad (19)$$

Remark 2. Only the unknown Hammerstein system parameters γ and θ are optimization variables for problems (18) and (19); on the contrary in the original identification problems (9) and (10) also the noise samples η are treated as variables. ■

Evaluation of the intervals $PUI_{\gamma_k}^{ss(n)}$ and $PUI_{\theta_j}^{ss(n)}$ requires to solve problems (18) and (19) over the nonconvex feasible set $\mathcal{D}^{ss(n)}$. In the following we describe how to solve numerically such problems by exploiting the particular structure of $\mathcal{D}^{ss(n)}$. In order to analyze topological features of $\mathcal{D}^{ss(n)}$, the following notation is introduced.

Let Γ be the set of all and only vectors with $n \cdot na$ components, each one equal to ± 1 . This means that $\Gamma = \{\alpha_1, \dots, \alpha_l, \dots, \alpha_L\}$, where $L = 2^{n \cdot na}$ and α_l is a vector with $n \cdot na$ components, each one equal to ± 1 and such that $\alpha_l \neq \alpha_i$ if $l \neq i$, for all $l, i = 1, \dots, L$. For any $\alpha_l \in \Gamma$, let us define the set $\mathcal{O}(\alpha_l)$ as

$$\mathcal{O}(\alpha_l) = \{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \alpha_{lj} q_j^{(n)}(\theta) \geq 0, j = 1, \dots, n \cdot na\}, \quad (20)$$

where α_{lj} and $q_j^{(n)}$ are the j -th element of the vectors α_l and $q^{(n)}$, respectively. Topological features of $\mathcal{D}^{ss(n)}$ are now highlighted by Property 2.

Property 2. The set $\mathcal{D}^{ss(n)}$ is the union of at most L sets $\mathcal{D}_l^{ss(n)}$ in $\mathbb{R}^{n_\gamma + n_\theta + N}$, that is $\mathcal{D}^{ss(n)} = \bigcup_{l=1}^L \mathcal{D}_l^{ss(n)}$, where $\mathcal{D}_l^{ss(n)} = \mathcal{D}^{ss(n)} \cap \mathcal{O}(\alpha_l)$. Furthermore, the set $\mathcal{D}_l^{ss(n)}$, if not empty, is a semialgebraic region in $\mathbb{R}^{n_\gamma + n_\theta + N}$. ■

Remark 3. From Remark 1 the maximum degree over $s = 1, \dots, n$ of $a_i^{(s)}$ and of the components of the vector $\underline{g}_j^{(s)}$ is less than or equal to $n + 1$. This means that, for every $l = 1, \dots, L$, the constraints describing $\mathcal{D}_l^{ss(n)}$ are polynomial inequalities with maximum degree $n + 1$. ■

Thanks to the structure of $\mathcal{D}^{ss(n)}$ highlighted by Property 2, (18) and (19) can be decomposed into the collection of the following polynomial optimization problems:

$$\underline{\gamma}_k^{ss(n)} = \min_{l=1, \dots, L} \underline{\gamma}_{kl}^{ss(n)}; \quad \bar{\gamma}_k^{ss(n)} = \max_{l=1, \dots, L} \bar{\gamma}_{kl}^{ss(n)}, \quad (21)$$

$$\underline{\theta}_j^{ss(n)} = \min_{l=1, \dots, L} \underline{\theta}_{jl}^{ss(n)}; \quad \bar{\theta}_j^{ss(n)} = \max_{l=1, \dots, L} \bar{\theta}_{jl}^{ss(n)}, \quad (22)$$

where

$$\underline{\gamma}_{kl}^{ss(n)} = \min_{\gamma, \theta \in \mathcal{D}_l^{ss(n)}} \gamma_k; \quad \overline{\gamma}_{kl}^{ss(n)} = \max_{\gamma, \theta \in \mathcal{D}_l^{ss(n)}} \gamma_k \quad (23)$$

$$\underline{\theta}_{jl}^{ss(n)} = \max_{\gamma, \theta \in \mathcal{D}_l^{ss(n)}} \theta_j; \quad \overline{\theta}_{jl}^{ss(n)} = \max_{\gamma, \theta \in \mathcal{D}_l^{ss(n)}} \theta_j. \quad (24)$$

Since the number of optimization variables involved in (23) and (24) is small, being equal to the number of unknown parameters γ and θ , computation of relaxed solution to such polynomial problems by means of LMI relaxation techniques is computationally tractable. In particular, for a given relaxation order δ , application of the theory-of-moments relaxation to problems (23) and (24) leads to approximated solutions $\underline{\gamma}_{kl}^{ss(n,\delta)}$, $\overline{\gamma}_{kl}^{ss(n,\delta)}$ and $\underline{\theta}_{jl}^{ss(n,\delta)}$, $\overline{\theta}_{jl}^{ss(n,\delta)}$ computed by solving the convex SDP problems

$$\underline{\gamma}_{kl}^{ss(n,\delta)} = \min_{p \in \mathcal{D}_l^{ss(n,\delta)}} f_k(p); \quad \overline{\gamma}_{kl}^{ss(n,\delta)} = \max_{p \in \mathcal{D}_l^{ss(n,\delta)}} f_k(p) \quad (25)$$

$$\underline{\theta}_{jl}^{ss(n,\delta)} = \min_{p \in \mathcal{D}_l^{ss(n,\delta)}} h_j(p); \quad \overline{\theta}_{jl}^{ss(n,\delta)} = \max_{p \in \mathcal{D}_l^{ss(n,\delta)}} h_j(p) \quad (26)$$

where p is the decision variable vector of dimension $\begin{pmatrix} n_\gamma + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$. The functionals $f_k(p)$, $h_j(p)$ in (25)-(26) are linear in the variables p and the feasible region $\mathcal{D}_l^{ss(n,\delta)}$ is a convex set defined by one LMI of size $\begin{pmatrix} n_\gamma + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$ and $2n(N - na) + 1$ LMIs whose maximum size is equal to $\begin{pmatrix} n_\gamma + n_\theta + \delta - 1 \\ \delta - 1 \end{pmatrix}$. The reader is referred to Lasserre (2001) for technical details on the relaxation of polynomial optimization problems through the theory of moments.

Remark 4. The minimum allowed value $\underline{\delta}$ of the LMI relaxation order, so that (25) and (26) are well-defined, is $\lceil \frac{\rho(\mathcal{D}_l^{ss(n)})}{2} \rceil$, where $\lceil \cdot \rceil$ is the ceiling operator and $\rho(\mathcal{D}_l^{ss(n)})$ denotes the maximum order of the polynomial constraints defining $\mathcal{D}_l^{ss(n)}$. From Remark 3 the maximum degree of the polynomial constraints describing $\mathcal{D}_l^{ss(n)}$ is equal to $n + 1$, therefore $\underline{\delta} = \lceil \frac{n+1}{2} \rceil$. ■

Remark 5. Since the linear block is known to be stable, stability constraints on the parameters a_1, \dots, a_{na} can be enforced in problems (23) and (24) with the method presented in Cerone et al. (2010a) in order to improve the parameter bounds evaluation. ■

For a given dynamic horizon $n \geq 1$ and a relaxation order $\delta \geq \underline{\delta}$, let us define the parameter uncertainty intervals $PUI_{\gamma_k}^{ss(n,\delta)}$ and $PUI_{\theta_j}^{ss(n,\delta)}$ as

$$PUI_{\gamma_k}^{ss(n,\delta)} = \left[\underline{\gamma}_k^{ss(n,\delta)}; \overline{\gamma}_k^{ss(n,\delta)} \right] \quad (27)$$

$$PUI_{\theta_j}^{ss(n,\delta)} = \left[\underline{\theta}_j^{ss(n,\delta)}; \overline{\theta}_j^{ss(n,\delta)} \right] \quad (28)$$

where

$$\underline{\gamma}_k^{ss(n,\delta)} = \min_{l=1,\dots,L} \underline{\gamma}_{kl}^{ss(n,\delta)}; \quad \overline{\gamma}_k^{ss(n,\delta)} = \max_{l=1,\dots,L} \overline{\gamma}_{kl}^{ss(n,\delta)} \quad (29)$$

$$\underline{\theta}_j^{ss(n,\delta)} = \min_{l=1,\dots,L} \underline{\theta}_{jl}^{ss(n,\delta)}; \quad \overline{\theta}_j^{ss(n,\delta)} = \max_{l=1,\dots,L} \overline{\theta}_{jl}^{ss(n,\delta)} \quad (30)$$

In the next section the main features enjoyed by the computed intervals $PUI_{\gamma_k}^{ss(n,\delta)}$ and $PUI_{\theta_j}^{ss(n,\delta)}$ are described.

5. PROPERTIES OF THE COMPUTED PARAMETER UNCERTAINTY INTERVALS

For all relaxation order $\delta \geq \underline{\delta} = \lceil \frac{n+1}{2} \rceil$ intervals $PUI_{\gamma_k}^{ss(n,\delta)}$ enjoy the features given by next properties. Similar results hold for $PUI_{\theta_j}^{ss(n,\delta)}$.

Property 3. Guaranteed relaxed uncertainty intervals.

For every dynamic horizon $n \in [1, N - na]$ the interval $PUI_{\gamma_k}^{ss(n,\delta)}$ is guaranteed to contain the true non-linear block parameter γ_k to be estimated, i.e. $\gamma_k \in PUI_{\gamma_k}^{ss(n,\delta)}$ for every $k = 1, \dots, n_\gamma$. ■

Property 4. Monotone convergence to parameter uncertainty intervals $PUI_{\gamma_k}^{ss(n)}$

For every dynamic horizon $n \in [1, N - na]$ the uncertainty interval $PUI_{\gamma_k}^{ss(n,\delta)}$ becomes tighter as the relaxation order δ increases, that is $PUI_{\gamma_k}^{ss(n,\delta+1)} \subseteq PUI_{\gamma_k}^{ss(n,\delta)}$. Furthermore, the computed interval $PUI_{\gamma_k}^{ss(n,\delta)}$ converges to $PUI_{\gamma_k}^{ss(n)}$ as the relaxation order δ goes to infinity. ■

Property 5. Increasing accuracy in uncertainty intervals evaluation

For every dynamic horizon $n \in [1, N - na - 1]$ the uncertainty interval $PUI_{\gamma_k}^{ss(n,\delta)}$ becomes tighter as the dynamic horizon n increases, i.e. $PUI_{\gamma_k}^{ss(n+1,\delta)} \subseteq PUI_{\gamma_k}^{ss(n,\delta)}$. ■

6. NUMERICAL EXAMPLE

In this section a simulated example is presented in order to show the effectiveness of the proposed approach. The numerical computation is carried out on a 2.40-GHz Intel Pentium IV with 3 GB of RAM. The nonlinear block of the Hammerstein system considered here is modeled by the polynomial function $x_t = 0.3u_t + 0.1u_t^2 - u_t^3$, i.e. $\gamma^T = [\gamma_1 \ \gamma_2 \ \gamma_3] = [0.3 \ 0.1 \ -1]$, while the linear part is a strictly-proper second order system with parameters $\theta^T = [a_1 \ a_2 \ b_1 \ b_2] = [0.7 \ 0.8 \ 1.4 \ 1.1]$. The output data sequence is corrupted by random additive noises η_t , uniformly distributed in $[-\Delta\eta_t, +\Delta\eta_t]$. The chosen error bounds $\Delta\eta_t$ is such that the signal to noise ratio

$$SNR_w = 10 \log \left\{ \frac{\sum_{t=1}^N w_t^2}{\sum_{t=1}^N \eta_t^2} \right\}$$

is equal to 27 db. The length of the data sequence is $N = 1000$. The chosen value of the dynamic horizon n is equal to 2. Parameter bounds are evaluated by solving (25) and (26) for a relaxation order $\delta = 2$. The open source software Gloptipoly by Henrion and Lasserre (2003), has been used to convert the original identification problems (23) and (24) into the corresponding LMI relaxed problems (25) and (26), which are numerically solved by the solver SeDuMi in Matlab. Results on the evaluation of nonlinear and the linear block parameters are reported in Table 1 and 2, respectively, which show the obtained parameters

Table 1. Nonlinear block. Parameter bounds ($\underline{\gamma}_k^{ss(n,\delta)}$, $\overline{\gamma}_k^{ss(n,\delta)}$) and parameter uncertainties $\Delta\gamma_k^{(n,\delta)}$ for relaxation order $\delta = 2$ and dynamic horizon $n = 2$.

n	Parameter	$\underline{\gamma}_k^{ss(n,\delta)}$	True value	$\overline{\gamma}_k^{ss(n,\delta)}$	$\Delta\gamma_k^{(n,\delta)}$
2	γ_1	0.269	0.300	0.332	0.031
	γ_2	0.072	0.100	0.126	0.027
	γ_3	-1.164	-1.000	-0.836	0.164

Table 2. Linear block. Parameter bounds ($\underline{\theta}_j^{ss(n,\delta)}$, $\overline{\theta}_j^{ss(n,\delta)}$) and parameter uncertainties $\Delta\theta_j^{(n,\delta)}$ for relaxation order $\delta = 2$ and dynamic horizon $n = 2$.

n	Parameter	$\underline{\theta}_j^{ss(n,\delta)}$	True value	$\overline{\theta}_j^{ss(n,\delta)}$	$\Delta\theta_j^{(n,\delta)}$
2	a_1	0.632	0.700	0.769	0.069
	a_2	0.675	0.800	0.958	0.141
	b_1	1.298	1.400	1.497	0.099
	b_2	1.023	1.100	1.179	0.078

bounds, as well as the parameter uncertainties $\Delta\gamma_k^{(n,\delta)}$ and $\Delta\theta_j^{(n,\delta)}$ defined as

$$\Delta\gamma_k^{(n,\delta)} = \frac{\overline{\gamma}_k^{ss(n,\delta)} - \underline{\gamma}_k^{ss(n,\delta)}}{2}, \quad \Delta\theta_j^{(n,\delta)} = \frac{\overline{\theta}_j^{ss(n,\delta)} - \underline{\theta}_j^{ss(n,\delta)}}{2}$$

The CPU elapsed time to compute a single parameter bound is between 115 s and 132 s. The reported results show that the true value of the parameters is always included in the computed uncertainty intervals, as expected. We point out that the presented procedure provides satisfactory parameter uncertainty intervals, both on the nonlinear block and on the linear one, also for a small value of dynamic horizon n and relaxation order δ .

7. CONCLUSION

A single-stage procedure to evaluate parameter bounds of Hammerstein systems in the presence of bounded errors is presented. Parameter bounds evaluation is formulated as a collection of constrained polynomial optimization problems, with a number of variables that increases with the number of measurements. Therefore, because of high computational complexity, LMI relaxation techniques proposed in literature to find approximate solutions of such polynomial problems can be exploited only for small/medium number of experimental data. In order to reduce the computational complexity of the identification problem, an outer bound of the feasible parameter set is sought. Since this outer-bound is the union of a finite number of semialgebraic sets over the parameter space, bounds are evaluated by solving suitable polynomial optimization problems involving a small number of variables, i.e. only the unknown system parameters. Then, LMI relaxation techniques are used to approximate global optima, computing uncertainty intervals that are proven to contain the true system parameters to be estimated. The simulated example shows that a large number of measurements can

be handled by the presented identification procedure, providing satisfactory parameter bounds also for low values of dynamic horizon and LMI-relaxation order.

REFERENCES

- Bai, E. (1998). An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 34(3), 333–338.
- Bai, E. (2003). Frequency domain identification of Hammerstein models. *IEEE Trans. Automatic Control*, 48(4), 530–542.
- Bai, E. and Fu, M. (2002). A blind approach to Hammerstein model identification. *IEEE Trans. Signal Processing*, 50(7), 1610–1619.
- Bai, E. and Giri, F. (2010). *Block-oriented nonlinear system identification*. Lecture notes in Control and Information sciences. Springer, Berlin.
- Bai, E. and Li, D. (2004). Convergence of the iterative Hammerstein system identification algorithm. *IEEE Trans. Automatic Control*, 49(11), 1929–1940.
- Cerone, V., Piga, D., and Regruto, D. (2010a). Bounding the parameters of linear systems with stability constraints. In *Proc. of the American Control Conference 2010*, 2152–2157.
- Cerone, V., Piga, D., and Regruto, D. (2010b). Computational complexity reduction in set-membership hammerstein system identification. *Internal Report*, DAUIN(2010/10/13).
- Cerone, V., Piga, D., and Regruto, D. (2011). Hammerstein systems identification through convex relaxation techniques. In *Proc. of American Control Conference, San Francisco, California, USA*.
- Cerone, V. and Regruto, D. (2003). Parameter bounds for discrete-time Hammerstein models with bounded output errors. *IEEE Trans. Automatic Control*, 48(10), 1855–1860.
- Chang, F. and Luus, R. (1971). A noniterative method for identification using Hammerstein model. *IEEE Trans. Automatic Control*, AC-16, 464–468.
- Greblicki, W. and Pawlak, M. (1989). Nonparametric identification of Hammerstein systems. *IEEE Trans. Automatic Control*, 35(2), 409–418.
- Henrion, D. and Lasserre, J.B. (2003). Gloptipoly: Global optimization over polynomials with Matlab and SeDuMi. *ACM Transactions Math. Soft*, 29, 165–194.
- Hsia, T. (1976). A multi stage least squares method for identifying Hammerstein model nonlinear systems. In *Proc. of IEEE Conference on Decision and Control*, 934–938.
- Hunter, I. and Korenberg, M. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biolog. Cybernet.*, 55, 135–144.
- Krzyżak, A. (1993). Identification of nonlinear block-oriented systems by the recursive kernel estimate. *Int. J. Franklin Inst.*, 330(3), 605–627.
- Krzyżak, A. (1996). On nonparametric estimation of nonlinear dynamic systems by the Fourier series estimate. *Signal Processing*, 52, 299–321.
- Lasserre, J.B. (2001). Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11, 796–817.
- Lasserre, J.B. (2006). Convergent semidefinite relaxations in polynomial optimization with sparsity. *SIAM Journal*

- on *Optimization*, 17(1), 822–843.
- Ljung, L. (2006). Some aspects on nonlinear system identification. In *Proc 14th IFAC Symposium on System Identification*. Newcastle, Australia.
- Ljung, L. (2008). Perspectives on system identification. In *Proc. 17th IFAC World Congress*. Seoul, Korea.
- Milanese, M. and Vicino, A. (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*, 27(6), 997–1009.
- Narendra, K. and Gallman, P. (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Trans. Automatic Control*, AC-11, 546–550.
- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P.Y., Hjalmarsson, H., and Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31, 1691–1724.
- Sznaier, M. (2009). Computational complexity analysis of set membership identification of Hammerstein and Wiener systems. *Automatica*, 45(3), 701–705.
- Verhaegen, M. and Westwick, D. (1996). Identifying MIMO Hammerstein systems in the context of subspace model identification methods. *Int. J. Control*, 63(2), 331–349.
- Waki, H., Kim, S., Kojima, M., and Muramatsu, M. (2006). Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1), 218–242.
- Walter, E. and Piet-Lahanier, H. (1990). Estimation of parameter bounds from bounded-error data: a survey. *Mathematics and Computers in simulation*, 32, 449–468.