Bounded error identification of Hammerstein systems through sparse polynomial optimization

V. Cerone [a,§], D. Piga [b,♯], D. Regruto [a]

[a] *Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: vito.cerone@polito.it, diego.regruto@polito.it*

[b] *Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, Delft, 2628 CD, The Netherlands, e-mail: D.Piga@tudelft.nl*

Abstract

In this paper we present a procedure for the evaluation of bounds on the parameters of Hammerstein systems, from output measurements affected by bounded errors. The identification problem is formulated in terms of polynomial optimization, and relaxation techniques, based on linear matrix inequalities, are proposed to evaluate parameter bounds by means of convex optimization. The structured sparsity of the formulated identification problem is exploited to reduce the computational complexity of the convex relaxed problem. Analysis of convergence properties and computational complexity is reported.

Key words: Bounded error identification, Hammerstein systems, Parameter bounds, Sparse LMI relaxation.

1 Introduction

Identification of block-structured nonlinear systems, modeled by interconnected memoryless nonlinear gains and linear dynamic subsystems, has attracted the attention of many authors in the last decades. Early works are surveyed in the papers (Billings, 1980), (Haber and Unbehauen, 1990) while an up-to-date collection of results and algorithms can be found in the recent book (Bai and Giri, 2010). These models are successfully employed in many engineering fields, because they can embed prior information on the process structure like, e.g., the presence of nonlinearity either in the actuator or in the measurement equipment. The configuration we are dealing with in this paper, 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 $\mathcal{L}$. The identification of such a model relies solely on input-output measurements, while the inner signal z_t, i.e. the output of the nonlinear block, is not assumed to be available. A number of algorithms addressing such a problem can be found in the literature. 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 (Narenda and Gallman, 1966), the nonparametric approach (Greblicki and Pawlak, 1989; Krzyżak, 1993), the frequency domain method (Krzyżak, 1996) and the algorithms based on the Bussgang's theorem (Hunter and Korenberg, 1986). As for Hammerstein system identification in set-membership context, in (Sznaier, 2009) it is shown that the problem is NP-hard in the size of the experimental data sequence, pointing out the need of computationally tractable relaxations. 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, while the linear dynamic part is described by an IIR output error model with bounded output measurement uncertainties. To the authors' best knowledge, the only contribution in the literature addressing such

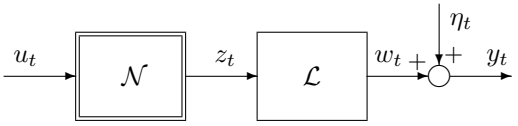

Fig. 1. Hammerstein system.

⋆ § Corresponding author V. Cerone. Tel. +39-(0)11-5647064 Fax +39-(0)11-5647099

♯ This research was developed while Dr. D. Piga was a Ph.D student at the Politecnico di Torino.

a problem is a previous work by some of the authors (Cerone and Regruto, 2003), where a two-stage identification procedure is presented. The main drawback of the procedure proposed in (Cerone and Regruto, 2003) is that it requires two different experiments where two specific input signals are exploited. In this paper an LMI-relaxation-based one-stage algorithm is presented to compute bounds on parameters of both the nonlinear and the linear subsystems. The proposed solution is an improvement over the one presented in (Cerone and Regruto, 2003) since: (i) the new approach provides parameter bounds of both nonlinear and linear subsystems by performing a single dynamic experiment without constraints on the class of input signals; (ii) the computed parameter bounds are guaranteed to monotonically converge to the exact ones as the order of relaxation increases, while the parameter uncertainty intervals computed in (Cerone and Regruto, 2003) are in general not tight and their degree of conservativeness is not systematically evaluated. The paper is organized as follows. Main results on relaxation of polynomial optimization problems are reviewed in Section 2 for self-consistency of the paper. Section 3 is devoted to the problem formulation. In Section 4 we show that computation of tight parameter bounds requires the solution to nonconvex polynomial optimization problems. In Section 5 we show that the formulated identification problem can be efficiently solved by means of LMI-relaxation techniques. The effectiveness of the presented identification procedure is shown in Section 6 through a simulation example.

2 Notation and background results on constrained polynomial optimization

Preliminary results on the relaxation of sparse polynomial optimization problems proposed by Lasserre in (Lasserre, 2006), in the spirit of the work (Waki *et al.*, 2006), are reviewed here.

2.1 Moment matrix and localizing matrixes

Let us consider the constrained optimization problem

$$f^\star = \min_{x\in\mathcal{S}} f(x), \tag{1}$$

where $\mathcal{S} \subseteq \mathbb{R}^n$ is a semialgebraic set defined as $\mathcal{S} = \{x \in \mathbb{R}^n : g_s(x) \geq 0,\ s = 1,\dots,\Xi\}$ with g_s a real-valued polynomial in the variable $x = [x_1, x_2, \dots, x_n]^{\mathrm{T}} \in \mathbb{R}^n$ of degree $d_s = deg(g_s)$, and $f \in \mathcal{P}_m^n[x]$, with $\mathcal{P}_m^n[x]$ denoting the space of real-valued polynomials with degree at most m in the variable $x \in \mathbb{R}^n$. By defining the set $\mathcal{A}_m^n = \{\alpha \in \mathbb{N}_0^n : \sum_i^n \alpha_i \leq m\}$, where α_i is the i-th component of the vector α and $\mathbb{N}_0^n$ denotes the set of n-dimensional nonnegative integers vectors, the canonical basis $h_m^n = \begin{bmatrix}1\ x_1\ x_2\ \cdots\ x_n\ x_1^2\ x_1x_2\ \cdots\ x_1x_n\ x_2^2\ x_2x_3\ \cdots\ x_n^2\ \cdots\ x_1^3\ \cdots x_n^m\end{bmatrix}^{\mathrm{T}}$ of the space $\mathcal{P}_m^n[x]$ can be written as $h_m^n = \{x^\alpha\}_{\alpha\in\mathcal{A}_m^n}$, where $x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$.
Let $p = \{p_\alpha\}_{\alpha\in\mathcal{A}_m^n}$ be the sequence of moments (up to order m) of a probability measure μ on $\mathbb{R}^n$, i.e. $p_\alpha = \int x^\alpha \mu(\mathrm{d}x)$ and $\mathcal{I}_0 = \{1,\dots,n\}$ be the union of a collection of R sets $\mathcal{I}_r \subset \{1,\dots,n\}$, that is $\{1,\dots,n\} = \bigcup_{r=1}^{R} \mathcal{I}_r$. Furthermore, let us partition the index set $\mathcal{S}_0 = \{1,\dots,\Xi\}$ into R disjoint sets S_r, $r = 1,\dots,R$. Let $h_m^{n_r}(\mathcal{I}_r)$ be the canonical basis of the polynomial $\mathcal{P}_m^{n_r}[x(\mathcal{I}_r)]$, where n_r is the cardinality of the set $\mathcal{I}_r$ and $x(\mathcal{I}_r) = \{x_i | i \in \mathcal{I}_r\}$. The truncated moment matrix $M_m(p,\mathcal{I}_r)$ associated with the moments sequence p and the variables $x(\mathcal{I}_r)$ is defined as $M_m(p,\mathcal{I}_r) = \int h_m^{n_r}(\mathcal{I}_r)h_m^{n_r}(\mathcal{I}_r)^{\mathrm{T}}\mu(\mathrm{d}x)$. By denoting with $p_{\beta(i,j)(\mathcal{I}_r)}$ the entry (i,j) of the matrix $M_m(p,\mathcal{I}_r)$, the localizing matrix $M_m(g_sp,\mathcal{I}_r)$ associated with the moments sequence p and the polynomial $g_s(x)$ is defined as $M_m(g_sp,\mathcal{I}_r)(i,j) = \sum_{\alpha\in\mathcal{A}_m^{n_r}} g_{s\alpha}p_{\{\beta(i,j)(\mathcal{I}_r)+\alpha\}}$, where $g_{s\alpha}$ is the coefficient of the term x^α in the polynomial $g_s(x)$ and $M_m(g_sp,\mathcal{I}_r)(i,j)$ denotes the entry (i,j) of the matrix $M_m(g_sp,\mathcal{I}_r)$. The reader is referred to (Lasserre, 2006) for an illustrative example about construction of moment and localizing matrixes.

2.2 Sparse LMI-relaxation for polynomial optimization problems

For a given integer δ such that $2\delta \geq \max\{m, \max_s d_s\}$, let us consider the SDP problem

$$\begin{aligned} f^\delta = &\min_p \sum_{\alpha\in\mathcal{A}_{2\delta}^n} f_\alpha p_\alpha \\ &s.t.\quad M_\delta(p,\mathcal{I}_r) \succeq 0,\quad M_{\delta-\tilde{d}_s}(g_sp,\mathcal{I}_r) \succeq 0, \\ &s \in \mathcal{S}_r,\quad r = 1,\dots,R \end{aligned}$$

where $\tilde{d}_s$ is the smallest integer greater than or equal to $\frac{d_s}{2}$ and $\mathbf{f} = \{f_\alpha\}_{\alpha\in\mathcal{A}_{2\delta}^n}$ is the sequence of coefficients of the polynomial f in the canonical basis $h_{2\delta}^n = \{x^\alpha\}_{\alpha\in\mathcal{A}_{2\delta}^n}$ of the space $\mathcal{P}_{2\delta}^n[x]$, i.e. $f(x) = \sum_{\alpha\in A_{2\delta}^n} f_\alpha x^\alpha$.

Proposition 1 If (i) constraints $g_s(x) \geq 0$ defining the feasible set $\mathcal{S}$ in problem (1) depend only on the variables $x(\mathcal{I}_r) = \{x_i | i \in \mathcal{I}_r\}$ for all $r = 1,\dots,R$ and for all $s \in \mathcal{S}_r$ and (ii) the objective function f can be written as $f = \sum_{r=1}^R f_r$, with $f_r \in \mathcal{P}_m^{n_r}[x(\mathcal{I}_r)]$, for all $r = 1,\dots,R$, then $f^\delta \leq f^{\delta+1} \leq f^*$. Furthermore, if (iii) there exists a finite value $G > 0$ such that $\|x\|_\infty \leq G$ for all $x \in \mathcal{S}$ and (iv) for all $r = 1,\dots,R-1$, $\mathcal{I}_{r+1}$ is such that: $\mathcal{I}_{r+1} \cap \bigcup_{j=1}^{r} \mathcal{I}_j \subseteq \mathcal{I}_q$, for some $q \leq r$, then then $\lim_{\delta\to\infty} f^\delta = f^*$.

3 Problem statement

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 z_t through the nonlinear function

$$z_t = \sum_{k=1}^{n_\gamma} \gamma_k \psi_k(u_t), \quad t = 1, \ldots, N; \tag{2}$$

where $(\psi_1,.....,\psi_{n_\gamma})$ is a known basis of nonlinear functions and N is the length of the data sequence. The linear dynamic part $\mathcal{L}$ is modeled by a stable discrete-time system which transforms z_t into the noise-free output w_t according to the linear difference equation

$$w_t = -\sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} b_j z_{t-j}, \tag{3}$$

with na and nb known constant. The measurement y_t of the noise-free output signal w_t is corrupted by additive noise η_t, i.e.

$$y_t = w_t + \eta_t, \tag{4}$$

where η_t is assumed to range within given bounds $\Delta\eta_t$, i.e.

$$\mid \eta_t \mid \leq \Delta\eta_t. \tag{5}$$

The unknown system parameters are collected into the vectors $\gamma = \left[\gamma_1, \gamma_2, \ldots, \gamma_{n_\gamma}\right]^{\mathrm{T}} \in \mathbb{R}^{n_\gamma}$ and $\theta = [a_1, \ldots, a_{na}, b_0, \ldots, b_{nb}]^{\mathrm{T}} \in \mathbb{R}^{n_\theta}$, where $n_\theta = n_a + n_b + 1$. It must be pointed out that the parametrization of the structure of Fig. 1 is not unique. In order to get a unique parametrization, we assume, without loss of generality, that the steady-state gain of the linear part is equal to one, i.e.

$$1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j. \tag{6}$$

In this paper we address the problem of deriving bounds on parameters γ and θ consistently with given measurements, error bounds and assumed model structure.

4 Evaluation of tight parameter uncertainty intervals

In this section we show how evaluation of parameter uncertainty intervals can be reduced to the computation of global optimum solutions to a set of semialgebraic optimization problems.

The mapping between the input signal u_t and the noise-free output w_t for the Hammerstein model considered in the paper can be obtained by substituting (2) into (3), leading to the relation

$$w_t = -\sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}). \tag{7}$$

Thus, from (4) and (7), the mapping between input signal u_t and output measurement y_t is given by

$$y_t = -\sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t. \tag{8}$$

Equations (5), (6) and (8) provide the following implicit description of set $\mathcal{D}_{\gamma\theta\eta}$ of all Hammerstein system parameters (γ, θ) and noise samples η_t consistent with given measurement data sequence, assumed model structure and error bounds, i.e.

$$\begin{aligned}
&\mathcal{D}_{\gamma\theta\eta} = \Big\{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\
&y_t = -\sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t, \\
&|\eta_r| \leq \Delta\eta_r, \quad 1 + \sum_{i=1}^{na} a_i = \sum_{j=0}^{nb} b_j, \\
&t = na + 1, \ldots, N;\ r = 1, \ldots, N\Big\},
\end{aligned}$$

which can be rewritten as a set of nonnegative polynomial constraints:

$$\begin{aligned}
&\mathcal{D}_{\gamma\theta\eta} = \Big\{(\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\
&g_t(\gamma, \theta, \eta) = -\sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \\
&\qquad + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t - y_t \geq 0, \\
&g_{t+N}(\gamma, \theta, \eta) = \sum_{i=1}^{na} a_i (y_{t-i} + \eta_{t-i}) + \\
&\qquad - \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) - \eta_t + y_t \geq 0, \\
&g_{r+2N}(\gamma, \theta, \eta) = \Delta\eta_r - \eta_r \geq 0, \\
&g_{r+3N}(\gamma, \theta, \eta) = \Delta\eta_r + \eta_r \geq 0,
\end{aligned} \tag{9}$$

$$\begin{aligned}
&g_{4N+1}(\gamma, \theta, \eta) = \sum_{i=1}^{na} a_i - 1 - \sum_{j=0}^{nb} b_j \geq 0, \\
&g_{4N+2}(\gamma, \theta, \eta) = -\sum_{i=1}^{na} a_i + 1 + \sum_{j=0}^{nb} b_j \geq 0, \\
&t = na + 1, \ldots, N;\ r = 1, \ldots, N\Big\},
\end{aligned} \tag{10}$$

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

$$\underline{\gamma}_k = \min_{(\gamma,\theta,\eta)\in\mathcal{D}_{\gamma\theta\eta}} \gamma_k, \quad \overline{\gamma}_k = \max_{(\gamma,\theta,\eta)\in\mathcal{D}_{\gamma\theta\eta}} \gamma_k, \tag{11}$$

$$\underline{\theta}_j = \min_{(\gamma,\theta,\eta)\in\mathcal{D}_{\gamma\theta\eta}} \theta_j, \quad \overline{\theta}_j = \max_{(\gamma,\theta,\eta)\in\mathcal{D}_{\gamma\theta\eta}} \theta_j. \qquad (12)$$

Uncertainty intervals on γ_k and θ_j are defined

$$PUI_{\gamma_k} = \left[\underline{\gamma}_k;\ \overline{\gamma}_k\right], \quad PUI_{\theta_j} = \left[\underline{\theta}_j;\ \overline{\theta}_j\right]. \qquad (13)$$

Remark 1 Intervals PUI_{γ_k} and PUI_{θ_j} in (13) are referred to as tight uncertainty intervals, in the sense that they are, by definition, the tightest intervals on the parameters γ_k and θ_j consistent with measurement data, model structure and error bounds. $\underline{\gamma}_k$, $\overline{\gamma}_k$, $\underline{\theta}_j$ and $\overline{\theta}_j$ are referred to as tight parameter bounds. ■

Assumption 1 In order to guarantee well-posedness of identification problems (11) and (12), we assume that $\underline{\gamma}_k$, $\overline{\gamma}_k$, $\underline{\theta}_j$ and $\overline{\theta}_j$ are bounded. ■

Property 1 The set $\mathcal{D}_{\gamma\theta\eta}$ described in (9) is: (i) bounded; (ii) basic closed semialgebraic (in particular, the set is described by $2(N-na)$ bilinear inequalities and $2N+2$ linear inequalities); (iii) compact.

Proof Because of Assumption 1 and eq. (5), $\mathcal{D}_{\gamma\theta\eta}$ is a bounded set. Besides, constraints $g_t \geq 0$ and $g_{t+N} \geq 0$ (with $t = na+1, \ldots, N$) defining $\mathcal{D}_{\gamma\theta\eta}$ in (9) are bilinear inequalities because of the product between the variable a_i and the noise η_{t-i} as well as the product between unknown parameters b_j and γ_k, while $g_{r+2N} \geq 0$, $g_{r+3N} \geq 0$ (with $r = 1, \ldots, N$), $g_{4N+1} \geq 0$ and $g_{4N+2} \geq 0$ are linear constraints. Thus, $\mathcal{D}_{\gamma\theta\eta}$ is a basic closed semialgebraic set, i.e., it is defined as the set of solutions of a finite system of nonnegative polynomial inequalities. $\mathcal{D}_{\gamma\theta\eta}$ is compact since it is closed and bounded. ■

Because of bilinear constraints $g_t(\gamma,\theta,\eta) \geq 0$ and $g_{t+N}(\gamma,\theta,\eta) \geq 0$, $\mathcal{D}_{\gamma\theta\eta}$ is, in general, a nonconvex set. Therefore, problems (11) and (12) can not be solved by means of standard optimization tools (gradient method, Newton method, etc.) since such tools can trap in local minima, which may prevent the computed uncertainty intervals from containing the true system parameters, key requirement in the context of set-membership identification. A possible solution to overcome such a problem is to relax identification problems (11) and (12) 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 $\overline{\gamma}_k$ and $\overline{\theta}_j$.

5 Evaluation of parameter bounds through convex relaxation techniques

Problems (11) and (12) are polynomial optimization problems since the functional is linear and the feasible set $\mathcal{D}_{\gamma\theta\eta}$ is semialgebraic. Therefore, (11) and (12) can be relaxed through a direct implementation of the *dense* LMI-relaxation technique proposed by Lasserre in (Lasserre, 2001). Such a procedure is based on the idea of relaxing a polynomial optimization problem into a sequence of semidefinite programming (SDP) problems with increasing dimension, whose optima are guaranteed to converge monotonically to the global optimum of the original polynomial problem. In particular, for a given relaxation order $\delta \geq 1$, application of the dense LMI-relaxation approach to identification problems (11) and (12), leads to semidefinite programming problems where the number of optimization variables is $\binom{n_\gamma+n_\theta+N+2\delta}{2\delta}$ and the feasible region is described by a moment matrix of size $\binom{n_\gamma+n_\theta+N+\delta}{\delta}$ and $2(N-na)+2N+2$ localizing matrixes, each one of size $\binom{n_\gamma+n_\theta+N+\delta-1}{\delta-1}$. Unfortunately, because of high computational burden and memory storage requirements, a direct use of the dense LMI-relaxation technique is limited to Hammerstein system identification problems with a small number N of measurements (roughly, N not greater than 5). In order to deal with identification problems with a larger number of measurements, the peculiar structured sparsity of identification problems (11) and (12) has been analyzed to apply the *sparse* LMI-relaxation approach discussed in Section 2.2. The inherent structured sparsity of problems (11) and (12) is highlighted by the following property.

Property 2 Problems (11) (resp. (12)) enjoy the following features: (i) the objective function depends only on the parameter γ_k (resp. θ_j); (ii) bilinear constraints $g_t \geq 0$ and $g_{t+N} \geq 0$ defining the feasible set $\mathcal{D}_{\gamma\theta\eta}$ depend only on Hammerstein system parameters γ and θ and noise samples η_{t-i}, with $i = 0, 1, \ldots, na$; linear constraints $g_{r+2N} \geq 0$ and $g_{r+3N} \geq 0$ depend only on the noise variable η_r; linear constraints $g_{4N+1} \geq 0$ and $g_{4N+2} \geq 0$ depend only on linear block parameters θ. ■

Thanks to the inherent structured sparsity of identification problems (11) and (12) described in Property 2, sparse SDP-relaxed problems for (11) and (12) can be formulated as described in the following.
Let $X \in \mathbb{R}^{n_\gamma+n_\theta+N}$ be the collection of the optimization variables for identification problems (11) and (12), i.e. $X = [\gamma^{\mathrm{T}}, \theta^{\mathrm{T}}, \eta^{\mathrm{T}}]^{\mathrm{T}}$ and X_i be the i-th component of the vector X. In such a way, the first n_γ components of X are the nonlinear block parameters γ, components from position $n_\gamma+1$ to $n_\gamma+n_\theta$ are the linear block parameters θ, while components from position $n_\gamma+n_\theta+1$ to $n_\gamma+n_\theta+N$ are the noise variables $\eta_1, \ldots, \eta_N$. Let us define the index sets $\mathcal{I}_r \subset \{1, 2, \ldots, n_\gamma+n_\theta+N\}$ and

$\mathcal{S}_r \subset \{na+1,\ldots,N, N+na+1,\ldots,4N+2\}$ as

$$\begin{aligned}\mathcal{I}_r = &\{1,2,\ldots,n_\gamma+n_\theta,\\ &n_\gamma+n_\theta+r, n_\gamma+n_\theta+r+1,\ldots,n_\gamma+n_\theta+r+na\}\\ &\text{for } r=1,\ldots,N-na\end{aligned} \tag{14}$$

$$\begin{aligned}\mathcal{S}_1 = &\{na+1, N+na+1,\\ &2N+1,2N+2,\ldots,2N+na+1,\\ &3N+1,3N+2,\ldots,3N+na+1,4N+1,4N+2\},\end{aligned} \tag{15}$$

$$\begin{aligned}\mathcal{S}_r = &\{na+r, N+na+r, 2N+na+r, 3N+na+r\},\\ &\text{for } r=2,\ldots,N-na.\end{aligned} \tag{16}$$

The index sets $\mathcal{I}_r$ and $\mathcal{S}_r$ are constructed on the basis of the sparse structure of identification problems (11) and (12) highlighted by Property 2. More precisely, the sets $\mathcal{I}_r$ and $\mathcal{S}_r$ are defined such that, for all $s \in \mathcal{S}_r$, all polynomial constraints $g_s \geq 0$ in the definition of $\mathcal{D}_{\gamma\theta\eta}$ depend only on variables X_i, with $i \in \mathcal{I}_r$.
For a given relaxation order $\delta \geq 1$, application of the sparse LMI-relaxation approach to problems (11) and (12) leads to the following SDP problems:

$$\underline{\gamma}_k^\delta = \min_{p\in\mathcal{D}^\delta_{\gamma\theta\eta}} \sum_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}} \Gamma_{k\alpha}p_\alpha, \quad \overline{\gamma}_k^\delta = \max_{p\in\mathcal{D}^\delta_{\gamma\theta\eta}} \sum_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}} \Gamma_{k\alpha}p_\alpha, \tag{17}$$

$$\underline{\theta}_j^\delta = \min_{p\in\mathcal{D}^\delta_{\gamma\theta\eta}} \sum_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}} \Theta_{j\alpha}p_\alpha, \quad \overline{\theta}_j^\delta = \max_{p\in\mathcal{D}^\delta_{\gamma\theta\eta}} \sum_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}} \Theta_{j\alpha}p_\alpha, \tag{18}$$

where $\Gamma_k = \{\Gamma_{k\alpha}\}_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}}$ and $\Theta_j = \{\Theta_{j\alpha}\}_{\alpha\in\mathcal{A}_{2\delta}^{n_\gamma+n_\theta+N}}$ are, respectively, the vectors of the coefficients of γ_k and θ_j in the canonical basis of polynomials of degree 2δ in the variables X. The feasible region $\mathcal{D}^\delta_{\gamma\theta\eta}$ of problems (17) and (18) is a convex set defined as

$$\begin{aligned}\mathcal{D}^\delta_{\gamma\theta\eta} = \{p:\ &M_\delta(p,\mathcal{I}_r)\succeq 0,\ M_{\delta-1}(g_s p,\mathcal{I}_r)\succeq 0,\\ &r=1,\ldots,N-na,\ \ s\in\mathcal{S}_r\ \},\end{aligned} \tag{19}$$

where $M_\delta(p,\mathcal{I}_r)$ is the moment matrix of order δ associated with variables $X(\mathcal{I}_r)$ and $M_{\delta-1}(g_s p,\mathcal{I}_r)$ is the localizing matrix associated with variables $X(\mathcal{I}_r)$ taking into account the constraint $g_s \geq 0$ defining the original semialgebraic feasible region $\mathcal{D}_{\gamma\theta\eta}$.

Property 3 For a given relaxation order $\delta \geq 1$, let us define the δ-relaxed uncertainty intervals on the nonlinear block parameters as $PUI^\delta_{\gamma_k} = \left[\underline{\gamma}_k^\delta;\ \overline{\gamma}_k^\delta\right]$. For all $k = 1,\ldots,n_\gamma$, intervals $PUI^\delta_{\gamma_k}$ satisfy the following properties.

P 3.1 ***Guaranteed relaxed uncertainty intervals.***
For any $\delta \geq 1$, the interval $PUI^\delta_{\gamma_k}$ is guaranteed to contain the true parameter γ_k^0, i.e. $\gamma_k^0 \in PUI^\delta_{\gamma_k}$.

P 3.2 ***Increasing accuracy in relaxed uncertainty intervals evaluation.***
For any $\delta \geq 1$, the interval $PUI^\delta_{\gamma_k}$ becomes tighter as the relaxation order δ increases, that is $PUI^{\delta+1}_{\gamma_k} \subseteq PUI^\delta_{\gamma_k}$.

P 3.3 ***Convergence to tight uncertainty intervals.***
The interval $PUI^\delta_{\gamma_k}$ converges to the tight interval PUI_{γ_k} as δ goes to infinity, that is $\lim_{\delta\to\infty} \underline{\gamma}_k^\delta = \underline{\gamma}_k$, $\lim_{\delta\to\infty} \overline{\gamma}_k^\delta = \overline{\gamma}_k$.

Proof Index sets $\mathcal{I}_r$ and $\mathcal{S}_r$ defined in (14) and (16) were carefully constructed in such a way that the assumptions of Proposition 1 are satisfied. Furthermore, from conditions (5) and Assumption 1, $\|X\|_\infty$ is bounded. Therefore, by applying the first part of Proposition 1 to identification problems (11) and (12) and to corresponding SDP-relaxed problems (17) and (18) we get:

$$\underline{\gamma}_k^\delta \leq \underline{\gamma}_k^{\delta+1} \leq \underline{\gamma}_k; \qquad \overline{\gamma}_k^\delta \geq \overline{\gamma}_k^{\delta+1} \geq \overline{\gamma}_k. \tag{20}$$

Then, from the definition of the intervals PUI_{γ_k} and $PUI^\delta_{\gamma_k}$ and conditions (20), we get:

$$\gamma_k^0 \in PUI_{\gamma_k} \subseteq PUI^{\delta+1}_{\gamma_k} \subseteq PUI^\delta_{\gamma_k}, \tag{21}$$

as stated in Properties **P**. 3.1 and **P**. 3.2. Besides, from the second part of Proposition 1, convergence conditions given by Property **P**. 3.3 follow. ■
Results similar to Property 3 hold for δ-relaxed intervals on the linear block parameters, defined as $PUI^\delta_{\theta_j} = \left[\underline{\theta}_j^\delta;\ \overline{\theta}_j^\delta\right]$. As to the computational complexity, the evaluation of intervals $PUI^\delta_{\gamma_k}$ and $PUI^\delta_{\theta_j}$ requires the solution to SDP problems (17) and (18), whose size is described by the following property.

Property 4 ***Computational complexity of SDP-problems*** (17) ***and*** (18)
Optimization problems (17) and (18) enjoy the following features:

(i) The number of free decision variables p is

$$(N-na)\binom{n_\gamma+n_\theta+na+1+2\delta}{2\delta} - (N-na-1)\binom{n_\gamma+n_\theta+na+2\delta}{2\delta}$$

(ii) The feasible region $\mathcal{D}^\delta_{\gamma\theta\eta}$ is described by $N-na$ moment matrixes, each one of size $\binom{n_\gamma+n_\theta+na+1+\delta}{\delta}$

and $2(N-na)+2N+2$ localizing matrixes, each one of size $\binom{n_\gamma + n_\theta + na + \delta}{\delta - 1}$. ■

Due to lack of space, the reader is referred to the technical report (Cerone *et al.*, 2011) for the proof of Property 4.

6 A simulation example

In this section we show the effectiveness of the presented parameter bounding procedure through a numerical example. The numerical computation is carried out on a single-thread 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 $z_t = 0.3u_t + 0.4u_t^2 - 0.9u_t^3$, thus the true nonlinear parameter vector is $\gamma^0 = \left[\gamma_1^0,\ \gamma_2^0,\ \gamma_3^0\right]^{\mathrm{T}} = [0.3,\ 0.4, 0.95]^{\mathrm{T}}$. The linear dynamic part is described by (3) with true parameter vector $\theta^0 = \left[a_1^0, a_2^0, b_1^0,\ b_2^0\right]^{\mathrm{T}} = [0.95,\ 0.85,\ 1.3,\ 1.5]^{\mathrm{T}}$. Parameter bounds are evaluated for three simulated data sets with different length N, i.e. $N = 50$, $N = 250$ and $N = 750$. The system is excited by a random input sequence uniformly distributed between $[-2,\ +2]$. The noise-free output w_t is corrupted by a random additive noise, uniformly distributed between $[-\Delta\eta_t,\ +\Delta\eta_t]$ and the chosen error bounds $\Delta\eta_t$ are such that the signal to noise ratio SNR_w on the output, defined as $SNR_w = 10\log\left\{\sum_{t=1}^{N} w_t^2 \Big/ \sum_{t=1}^{N} \eta_t^2\right\}$, is 27 db. Bounds on the parameters are evaluated by solving (17) and (18) for a relaxation order $\delta = 2$. The Matlab package SparsePOP (Waki *et al.*, 2008) has been used to convert the original identification problems (11) and (12) into their corresponding LMI relaxed problems (17) and (18), which are numerically solved by the SDP solver SeDuMi. In the case of identification of the Hammerstein system considered in this example, the complexity of the SDP-problems (17) and (18), in terms of number of decision variables, number and size of the moment matrixes and localizing matrixes defining the feasible region, is reported in Table 1. In the same table we also report the size of the SDP-problems, in terms of number of variables and constraints, that would be obtained when relaxing identification problems (11) and (12) through a direct application of the dense LMI-relaxation approach in (Lasserre, 2001), without taking into account the structured sparsity of such identification problems. Results in Table 1 show that a significant computational burden reduction is obtained by exploiting sparsity of problems (11) and (12). For instance, in the case $N = 750$, the number of optimization variables in (17) and (18) is 214,643, while the feasible region is defined by 748 moment matrixes of size 66 and $2,998$ localizing matrixes of size 11. On the other hand, if the sparsity was not taken into account, the number of optimization variables of the SDP relaxed problems would be about 14 billion and the feasible region would be described by a moment matrix of size $287,661$ and $2,998$ localizing matrixes of size 758, leading to an untractable optimization problem. Results about the nonlinear and the linear block parameter estimates are reported in Table 2 and 3, respectively, which show the obtained parameter bounds $\underline{\gamma}_k^\delta$, $\overline{\gamma}_k^\delta$, $\underline{\theta}_j^\delta$ and $\overline{\theta}_j^\delta$; the central estimates $\gamma_k^c = \frac{\overline{\gamma}_k^\delta + \underline{\gamma}_k^\delta}{2}$ and $\theta_j^c = \frac{\overline{\theta}_j^\delta + \underline{\theta}_j^\delta}{2}$ when $\delta = 2$. The CPU time taken by the SeDuMi solver to compute the solution of a single problem (17) and (18) is between 201 s and 253 s when the number of measurements N is equal to 50, between 2372 s and 2453 s when $N = 250$ and between 4568 s and 4796 s when $N = 750$. The reported results show that, as the number of observations increases (from $N = 50$ to $N = 750$), the width of the parameter uncertainty intervals $\Delta\gamma_k^\delta$ and $\Delta\theta_j^\delta$ decreases. Furthermore, true parameter values are included in the uncertainty intervals $PUI_{\gamma_k}^\delta$ and $PUI_{\theta_j}^\delta$, as expected.

Table 2
Nonlinear block. Parameter central estimates (γ_k^c), parameter bounds ($\underline{\gamma}_k^\delta$, $\overline{\gamma}_k^\delta$) for $N = 50$, $N = 250$, $N = 750$ and $\delta = 2$

N	Parameter	True Value	$\underline{\gamma}_k^\delta$	γ_k^c	$\overline{\gamma}_k^\delta$
50	γ_1	0.300	0.127	0.298	0.469
	γ_2	0.400	0.214	0.406	0.598
	γ_3	-0.900	-1.095	-0.933	-0.772
250	γ_1	0.300	0.235	0.304	0.372
	γ_2	0.400	0.289	0.402	0.515
	γ_3	-0.900	-1.020	-0.928	-0.837
750	γ_1	0.300	0.264	0.311	0.357
	γ_2	0.400	0.327	0.418	0.508
	γ_3	-0.900	-0.944	-0.911	-0.879

Table 3
Linear block. Parameter central estimates (θ_j^c), parameter bounds ($\underline{\theta}_j^\delta$, $\overline{\theta}_j^\delta$)for $N = 50$, $N = 250$, $N = 750$ and $\delta = 2$.

N	Parameter	True Value	$\underline{\theta}_j^\delta$	θ_j^c	$\overline{\theta}_j^\delta$
50	a_1	0.950	0.895	0.941	0.988
	a_2	0.850	0.805	0.850	0.895
	b_1	1.300	1.026	1.242	1.458
	b_2	1.500	1.266	1.477	1.689
250	a_1	0.950	0.947	0.950	0.954
	a_2	0.850	0.846	0.851	0.856
	b_1	1.300	1.197	1.293	1.389
	b_2	1.500	1.445	1.492	1.539
750	a_1	0.950	0.948	0.950	0.952
	a_2	0.850	0.848	0.849	0.851
	b_1	1.300	1.223	1.297	1.372
	b_2	1.500	1.462	1.482	1.517

7 Conclusions

A procedure for the evaluation of parameter uncertainty intervals for Hammerstein systems is presented. Param-

Table 1
Size of the SDP-problems obtained by relaxing identification problems (11) and (12) through sparse LMI-relaxation and dense LMI-relaxation.

N	LMI-Relaxation technique	Number of optimization variables	Number of moment matrixes defining the feasible region	Size of each moment matrix	Number of localizing matrixes defining the feasible region	Size of each localizing matrix
50	Sparse version	14,443	48	66	198	11
50	Dense version	521,855	1	1,711	198	58
250	Sparse version	71,643	248	66	998	11
250	Dense version	$189 \cdot 10^6$	1	33,411	998	258
750	Sparse version	214,643	748	66	2,998	11
750	Dense version	$14 \cdot 10^9$	1	287,661	2,998	758

eter bounds evaluation is formulated in terms of a set of polynomial optimization problems, whose approximate solutions can be computed by solving relaxed semidefinite programming problems. Unfortunately, because of high computational complexity, identification problems are practically intractable when more than 5 measurements are considered. In order to overcome this significant limitation, the peculiar structured sparsity of the identification problem is exploited, making it possible to reduce the computational complexity of the formulated relaxed problems. The computed uncertainty intervals are guaranteed to contain the true system parameters and to monotonically converge to the tight uncertainty parameters as the order of relaxation increases. The presented method can be efficiently applied to compute bounds on the parameters of linear output-error models, which are a subclass of the Hammerstein models considered in the paper. Such a method can be straightforwardly extended, at least theoretically, to the identification of nonlinear output-error models with polynomial input-output dependence. However, its direct application for parameter bounding of generic polynomial output-error structures could lead to relaxed SDP problems that can not be solved on commercial workstations because of high computational burden. Development of "ad hoc" relaxation strategies to evaluate parameter bounds of specific polynomial output-error models, such as Wiener systems with polynomial invertible nonlinearity, is currently under investigation.

References

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

Bai, E.W. and F. Giri (2010). *Block-oriented nonlinear system identification*. Lecture notes in Control and Information sciencess. Springer. Berlin.

Bai, E.W. and M. Fu (2002). A blind approach to Hammerstein model identification. *IEEE Trans. Signal Processing* **50**(7), 1610–1619.

Billings, S.A. (1980). Identification of nonlinear systems — a survey. *IEE Proc. Part D* **127**(6), 272–285.

Cerone, V. and D. Regruto (2003). Parameter bounds for discrete-time Hammerstein models with bounded output errors. *IEEE Trans. Automatic Control* **48**(10), 1855–1860.

Cerone, V. D. Piga and D. Regruto (2011). A sparse convex relaxation approach to hammerstein systems identification. *Internal Report*.

Chang, F.H.I. and R. Luus (1971). A noniterative method for identification using Hammerstein model. *IEEE Trans. Automatic Control* **AC-16**, 464–468.

Greblicki, W. and M. Pawlak (1989). Nonparametric identification of Hammerstein systems. *IEEE Trans. Automatic Control* **35**(2), 409–418.

Haber, R. and H. Unbehauen (1990). Structure identification of nonlinear dynamic systems – a survey on input/output approaches. *Automatica* **26**(4), 651–677.

Hsia, T.C. (1976). A multi stage least squares method for identifying Hammerstein model nonlinear systems. In: *Proc. of IEEE Conference on Decision and Control*. pp. 934–938.

Hunter, I.W. and M.J. Korenberg (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**(3), 822–843.

Narenda, K.S. and P.G. Gallman (1966). An iterative method for the identification of nonlinear systems using a Hammerstein model. *IEEE Trans. Automatic Control* **AC-11**, 546–550.

Sznaier, M. (2009). Computational complexity analysis of set membership identification of Hammerstein and Wiener systems. *Automatica* **45**(3), 701–705.

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

Waki, H. S. Kim, M. Kojima and M. Muramatsu (2006). Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization* **17**(1), 218–242.

Waki, H. S. Kim, M. Kojima, M. Muramatsu and H. Sugimoto (2008). SparsePOP: a sparse semidefinite programming relaxation of polynomial optimization problems. *ACM Transaction on Mathematical Software*.