

# Hammerstein systems parameters bounding through sparse polynomial optimization

V. Cerone, D. Piga, D. Regruto

**Abstract**—A single-stage procedure for the evaluation of tight bounds on the parameters of Hammerstein systems from output measurements affected by bounded errors is presented. The identification problem is formulated in terms of polynomial optimization, and relaxation techniques based on linear matrix inequalities are proposed to evaluate parameters bounds by means of convex optimization. The structured sparsity of the identification problem is exploited to reduce the computational complexity of the convex relaxed problem. Convergence properties, complexity analysis and advantages of the proposed technique with respect to previously published ones are discussed.

**Index Terms**—Bounded error identification, Hammerstein systems, Sparse LMI relaxation, Parameters bounds.

## I. INTRODUCTION

Identification of block-structured nonlinear systems, modeled by interconnected memoryless nonlinear gains and linear dynamic subsystems, has attracted the attention of the system identification research community in the last decades. An up-to-date collection of results and algorithms can be found in the recent book [1]. Block-oriented models are successfully used in many engineering fields, thanks to their ability to embed prior process structure knowledge like, e.g., the presence of nonlinearity either in the actuator or in the measurement equipment. 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  $z_t$ , i.e. the output of the nonlinear block, is not assumed to be available. A number of algorithm can be found in the literature to address such a problem. Among others, we mention the over-parametrization method [2], the subspace identification [3], the blind approach [4], the iterative method [5], the nonparametric approach [6] and the frequency domain method [7]. In all the papers mentioned above, the authors assume that the measurement error  $\eta_t$  is statistically described. A worthwhile alternative to the stochastic description of measurement errors is the bounded-errors, or set-membership, characterization, where uncertainties are assumed to belong to a given set. The reader can find further details on this approach in the survey papers [8], [9] and in the book [10]. As far as set-membership identification of Hammerstein systems is concerned, Sznaier has recently shown in [11] that the problem is NP-hard in

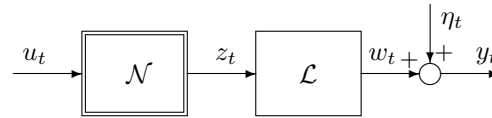


Fig. 1. Hammerstein system.

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, the linear dynamic part is described by an output error model and the output measurement errors are bounded. In a previous paper by the authors [12], 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 drawback of the procedure proposed in [12] is that it requires two different experiments where two specific input signals have to be used. On the contrary, when the input signal cannot be arbitrarily chosen, a one-step procedure without particular constraints on the input signal is required. In this paper a one-stage algorithm based on LMI-relaxation is proposed for the computation of bounds on the parameters of both the nonlinear and the linear subsystems without constraints on the class of input signals. The paper is organized as follows. Background results on the relaxation of semialgebraic optimization problems through the theory of moments is presented in Section II. Section III is devoted to the problem formulation. In Section IV we show that computation of tight parameters bounds requires the solution to nonconvex optimization problems. The proposed LMI relaxation-based technique, together with a detailed analysis of its properties, is described in Section V. A simulated example is reported in Section VI.

## II. NOTATION AND BACKGROUND RESULTS ON CONSTRAINED POLYNOMIAL OPTIMIZATION

In this section we briefly review some preliminary results on the relaxation of sparse polynomial optimization problems

The authors are with the Dipartimento di Automatica e Informatica, Politecnico di Torino, corso Duca degli Abruzzi 24, 10129 Torino, Italy; e-mail: vito.cerone@polito.it, dario.piga@polito.it, diego.regruto@polito.it; Tel: +39-011-564 7064; Fax: +39-011-564 7198

through a hierarchy of semidefinite programming (SDP) problems of increasing dimension. The reader is referred to [13] and the references therein for further details.

### A. Polynomial representation and theory of moments

Let us denote with  $\mathcal{P}_m^n[x]$  the space of real-valued polynomials of degree at most  $m$  in the variable  $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$  and let  $h$  be the canonical basis of  $\mathcal{P}_m^n[x]$ , i.e.  $h = [1 \ x_1 \ x_2 \ \dots \ x_n \ x_1^2 \ x_1 x_2 \ \dots \ x_1 x_n \ x_2^2 \ x_2 x_3 \ \dots \ x_n^3 \ \dots \ x_1^m]^T$ . Let us define the set  $\mathcal{A}_m$  as  $\mathcal{A}_m = \{\alpha \in \mathbb{N}_0^n : \sum_i \alpha_i \leq m\}$ , where  $\alpha_i$  is the  $i$ -th component of the vector  $\alpha$  and  $\mathbb{N}_0^n$  denotes the set of  $n$ -dimensional nonnegative integer vectors. Then, the basis  $h$  of the space  $\mathcal{P}_m^n[x]$  can be written as  $h = \{x^\alpha\}_{\alpha \in \mathcal{A}_m}$ , where  $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$ .

Let  $f$  and  $g_s$  be in  $\mathcal{P}_m^n[x]$ . We denote the sequence  $\mathbf{f} = \{f_\alpha\}_{\alpha \in \mathcal{A}_m}$  and  $\mathbf{g}_s = \{g_{s\alpha}\}_{\alpha \in \mathcal{A}_m}$  as the coefficients of the polynomials  $f$  and  $g_s$ , respectively, on the basis  $h_m$ , i.e.  $f(x) = \sum_{\alpha \in \mathcal{A}_m} f_\alpha x^\alpha$ ,  $g_s(x) = \sum_{\alpha \in \mathcal{A}_m} g_{s\alpha} x^\alpha$ .

Let  $p = \{p_\alpha\}_{\alpha \in \mathcal{A}_m}$  be the sequence of moments (up to order  $m$ ) of a probability measure  $\mu$  on  $\mathbb{R}^n$ , i.e.  $p_\alpha = \int x^\alpha \mu(dx)$  and  $M_m(p)$  be the truncated moment matrix associated with the distribution  $\mu$ , i.e.  $M_m(p) = \int h h^T \mu(dx)$ . Let us denote with  $M_m(g_k p)$  the localizing matrix associated with the sequence of moments  $p$  and with the polynomial  $g_k(x)$ . The reader is referred to [14] for details on the construction of the localizing matrix associated with a polynomial.

### B. LMI-relaxation for polynomial optimization problems

The LMI-relaxation technique based on the theory of moments and proposed in [14] is briefly reviewed here. Let us consider the constrained optimization problem

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

where  $f \in \mathcal{P}_m^n[x]$  and  $\mathcal{S} \subseteq \mathbb{R}^n$  is a compact semialgebraic set defined as

$$\mathcal{S} = \{x \in \mathbb{R}^n : g_s(x) \geq 0, s = 1, \dots, \Xi\}, \quad (2)$$

where  $g_s$  is a real-valued polynomial in the variable  $x \in \mathbb{R}^n$  of degree  $d_s = \deg(g_s)$ , i.e.  $g_s \in \mathcal{P}_{d_s}^n[x]$ . Let  $\delta \in \mathbb{N}$  be such that  $2\delta \geq \max\{m, \max_s d_s\}$  and  $h = \{x^\alpha\}_{\alpha \in \mathcal{A}_{2\delta}}$  be the canonical basis of the space  $\mathcal{P}_{2\delta}^n[x]$ . Indeed,  $f$  and  $g_s$  belong to  $\mathcal{P}_{2\delta}^n[x]$ .

Now, let us consider the SDP problem

$$\begin{aligned} f^\delta &= \min_p \sum_{\alpha \in \mathcal{A}_{2\delta}^n} f_\alpha p_\alpha \\ \text{s.t. } &M_\delta(p) \succeq 0, M_{\delta-\tilde{d}_s}(g_s p) \succeq 0, s = 1, \dots, \Xi \end{aligned} \quad (3)$$

where  $\tilde{d}_s = \lceil \frac{d_s}{2} \rceil$ ,  $p = \{p_\alpha\}_{\alpha \in \mathcal{A}_{2\delta}}$  is the sequence of moments up to order  $2\delta$  of some probability measure  $\mu$  with support on  $\mathcal{S}$ , while  $M_\delta(p)$  is the moment matrix associated with the moments  $p$  and  $M_{\delta-\tilde{d}_s}(g_s p)$  is the localizing matrix associated with the polynomial  $g_s$ . Problem (3) is referred to as LMI-relaxed problem of order  $\delta$  of the original polynomial problem (1). The solution  $f^\delta$  to the convex problem (3) is

a lower bound of the global optimum  $f^*$  of the nonconvex problem (1). Besides, under mild conditions,  $f^\delta$  converges to  $f^*$  as the relaxation order  $\delta$  goes to infinity. Although the convergence properties are guaranteed as the relaxation order goes to infinity, exact global optimum  $f^*$  can be obtained in practice with a reasonably low relaxation order (see [15] for a collection of problems solved with relaxation order less than 5). Unfortunately, due to high computational complexity, the discussed LMI-relaxation is restricted to polynomial problems with a small number of variables, in general not greater than 10 for relaxation order smaller than 4. Several efforts on the reduction of LMI relaxation complexity, by exploiting the structured sparsity of the original polynomial problems, have been carried out in recent years (see, e.g., [16], [17], [18]). Roughly speaking, an optimization problem has a structured sparsity when the functional and each constraint defining the feasible region involve only a small subset of variables. In the next section we describe the relaxation technique presented in [19] in the spirit of the work of Waki *et al* [17]. Such a technique exploits the sparsity in the original polynomial problems to formulate a sparse version of the SDP-relaxation previously described, in order to extend the applicability of such a methodology to medium and large scale problems.

### C. Sparse LMI-relaxation for polynomial problems

Let us consider the optimization problem (1) with  $\mathcal{S}$  as in (2). Let  $\mathcal{I}_0 = \{1, \dots, n\}$  be the union of a collection of  $R$  sets  $\mathcal{I}_r \subset \{1, \dots, n\}$ , that is  $\{1, \dots, n\} = \bigcup_{r=1}^R \mathcal{I}_r$ . Further,

let us partition the index set  $\mathcal{S}_0 = \{1, \dots, \Xi\}$  into  $R$  disjoint sets  $\mathcal{S}_r$ ,  $r = 1, \dots, R$ . Let  $h(\mathcal{I}_r)$  be the canonical basis of the polynomial  $\mathcal{P}_m^{n_r}[x(\mathcal{I}_r)]$ , where  $x(\mathcal{I}_r) = \{x_i | i \in \mathcal{I}_r\}$ . Let us construct the partial moment matrix  $M_m(p, \mathcal{I}_r)$  (respectively the partial localizing matrix  $M_m(g_s p, \mathcal{I}_r)$ ) by retaining only those rows and columns of the moment matrix  $M_m(p)$  (respectively of the localizing matrix  $M_m(g_s p)$ ), where the variables  $p_\alpha$  are such that  $\text{supp}(\alpha) \in \mathcal{I}_r$ , with  $\text{supp}(\alpha)$  denoting the support of the vector  $\alpha$ .

For a given  $\delta \in \mathbb{N}$  such that  $2\delta \geq \max\{m, \max_s d_s\}$ , let us define the SDP problem

$$\begin{aligned} f_{sp}^\delta &= \min_p \sum_{\alpha \in \mathcal{A}_{2\delta}} f_\alpha p_\alpha \\ \text{s.t. } &M_\delta(p, \mathcal{I}_r) \succeq 0 \\ &M_{\delta-\tilde{d}_s}(g_s p, \mathcal{I}_r) \succeq 0, s \in \mathcal{S}_r, r = 1, \dots, R. \end{aligned} \quad (4)$$

Let us consider the following assumptions.

*Assumption 1:* For all  $r = 1, \dots, R$  and for all  $s \in \mathcal{S}_r$ , the constraint  $g_s(x) \geq 0$  defining  $\mathcal{S}$  in (2), depends only on the variables  $x(\mathcal{I}_r) = \{x_i | i \in \mathcal{I}_r\}$ .

*Assumption 2:* The objective function  $f$  can be written as

$$f = \sum_{r=1}^R f_r, \text{ with } f_r \in \mathcal{P}_m^n[x(\mathcal{I}_r)], \text{ for every } r = 1, \dots, R.$$

*Assumption 3:* There exists a value  $G > 0$  such that  $\|x\|_\infty \leq G$  for all  $x \in \mathcal{S}$ .

*Assumption 4:* For all  $r = 1, \dots, R - 1$ ,

$$\mathcal{I}_{r+1} \cap \bigcup_{j=1}^r \mathcal{I}_j \subseteq \mathcal{I}_q, \text{ for some } q \leq r.$$

According to [19], the following result holds.

**Theorem 1:** Under Assumptions 1 and 2 we have:  $f_{sp}^\delta \leq f_{sp}^{\delta+1} \leq f^*$ . Furthermore, if also Assumptions 3 and 4 are satisfied, then  $\lim_{\delta \rightarrow \infty} f_{sp}^\delta = f^*$ .

An implementation of the discussed sparse LMI-relaxation can be found in the Matlab package SparsePOP [20], which exploits the SDP solvers SeDuMi [21] and SDPA [22].

### III. 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 following nonlinear function

$$z_t = \sum_{k=1}^{n_\gamma} \gamma_k \psi_k(u_t), \quad t = 1, \dots, N; \quad (5)$$

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

$$w_t = - \sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} b_j z_{t-j}. \quad (6)$$

Let  $y_t$  be the noise-corrupted output

$$y_t = w_t + \eta_t. \quad (7)$$

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

$$|\eta_t| \leq \Delta\eta_t. \quad (8)$$

Unknown parameter vectors  $\gamma \in \mathbb{R}^{n_\gamma}$  and  $\theta \in \mathbb{R}^{n_\theta}$  are defined, respectively, as  $\gamma = [\gamma_1 \ \gamma_2 \ \dots \ \gamma_{n_\gamma}]^\top$  and  $\theta = [a_1 \ \dots \ a_{na} \ b_0 \ b_1 \ \dots \ b_{nb}]^\top$ , with  $n_\theta = n_a + n_b + 1$ . It must be pointed out that the parametrization of the structure of 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  $\alpha$ , 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.

$$\sum_{i=1}^{na} a_i = 1 + \sum_{j=0}^{nb} b_j. \quad (9)$$

In this paper we address the problem of deriving bounds on the parameters  $\gamma$  and  $\theta$  consistently with given measurements, error bounds and the assumed model structure.

### IV. EVALUATION OF TIGHT PARAMETER UNCERTAINTY INTERVALS

The mapping between the input signal  $u_t$  and the noise-free output  $w_t$  for the Hammerstein model in Fig. 1 can be obtained by substituting (5) in (6), which leads to the relation

$$w_t = - \sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}). \quad (10)$$

Therefore, from (7) and (10), we get the following mapping between the input signal and the output measurement:

$$y_t = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t. \quad (11)$$

Indeed, the set  $\mathcal{D}_{\gamma\theta\eta}$  of all the Hammerstein system parameters  $(\gamma, \theta)$  and the noise samples  $\eta_t$  consistent with the measurement data sequence, the assumed model structure and the error bounds is described by (8), (9) and (11), i.e.

$$\begin{aligned} \mathcal{D}_{\gamma\theta\eta} = \{ & (\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\ & y_t = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \\ & + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t, \quad (12) \\ & |\eta_r| \leq \Delta\eta_r, \quad \sum_{i=1}^{na} a_i = 1 + \sum_{j=0}^{nb} b_j, \\ & t = na + 1, \dots, N; \quad r = 1, \dots, N \}, \end{aligned}$$

which is rewritten as

$$\begin{aligned} \mathcal{D}_{\gamma\theta\eta} = \{ & (\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : \\ & g_t(\gamma, \theta, \eta) = - \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \\ & + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) + \eta_t - y_t \geq 0, \\ & g_{t+N}(\gamma, \theta, \eta) = \sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) - \\ & - \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k(u_{t-j}) - \eta_t + y_t \geq 0, \quad (13) \\ & g_{r+2N}(\gamma, \theta, \eta) = \Delta\eta_r - \eta_r \geq 0, \\ & g_{r+3N}(\gamma, \theta, \eta) = \Delta\eta_r + \eta_r \geq 0, \\ & g_{4N+1}(\gamma, \theta, \eta) = \sum_{i=1}^{na} a_i - 1 - \sum_{j=0}^{nb} b_j \geq 0, \\ & g_{4N+2}(\gamma, \theta, \eta) = - \sum_{i=1}^{na} a_i + 1 + \sum_{j=0}^{nb} b_j \geq 0, \\ & t = na + 1, \dots, N; \quad r = 1, \dots, N \}. \end{aligned}$$

with  $\eta = [\eta_1, \dots, \eta_N]^\top$ . Therefore, for all  $k = 1, \dots, n_\gamma$  and  $j = 1, \dots, n_\theta$ , tight bounds on the parameters  $\gamma_k$  and  $\theta_j$  can be computed by solving the optimization problems:

$$\underline{\gamma}_k = \min_{(\gamma, \theta, \eta) \in \mathcal{D}_{\gamma\theta\eta}} \gamma_k, \quad \bar{\gamma}_k = \max_{(\gamma, \theta, \eta) \in \mathcal{D}_{\gamma\theta\eta}} \gamma_k, \quad (14)$$

$$\underline{\theta}_j = \min_{(\gamma, \theta, \eta) \in \mathcal{D}_{\gamma\theta\eta}} \theta_j, \quad \bar{\theta}_j = \max_{(\gamma, \theta, \eta) \in \mathcal{D}_{\gamma\theta\eta}} \theta_j. \quad (15)$$

Thus, parameter uncertainty intervals on  $\gamma_k$  and  $\theta_j$  are implicitly defined as  $PUI_{\gamma_k} = [\underline{\gamma}_k; \bar{\gamma}_k]$  and  $PUI_{\theta_j} = [\underline{\theta}_j; \bar{\theta}_j]$ . Note that the identification problems (14) and (15) are semialgebraic optimization problems. In fact, the objective function is linear and the feasible region  $\mathcal{D}_{\gamma\theta\eta}$  is semialgebraic, since the constraints  $g_t(\gamma, \theta, \eta) \geq 0$  and  $g_{t+N}(\gamma, \theta, \eta) \geq 0$  defining  $\mathcal{D}_{\gamma\theta\eta}$  in (13) are bilinear inequalities because of the product between the variable  $a_i$  and the noise  $\eta_{t-i}$  as well as the product between the unknown parameters  $b_j$  and  $\gamma_k$ . Because of bilinear constraints  $g_t(\gamma, \theta, \eta) \geq 0$  and  $g_{t+N}(\gamma, \theta, \eta) \geq 0$  defining the feasible region  $\mathcal{D}_{\gamma\theta\eta}$ , problems (14) and (15) are nonconvex. Therefore, standard nonlinear optimization tools (gradient method, Newton method, etc.) cannot be used since they can trap in local minima, which may prevent the computed uncertainty intervals from containing the true parameters, key requirement of any set-membership identification method. One possible solution to overcome such a problem is to relax identification problems (14) and (15) to convex optimization problems in order to numerically compute guaranteed parameter bounds.

#### V. EVALUATION OF PARAMETERS BOUNDS THROUGH CONVEX RELAXATION TECHNIQUES

Since (14) and (15) are semialgebraic optimization problems, they can be relaxed through a direct implementation of the *dense* LMI-relaxation technique described in Section II-B, which guarantees monotone converge to the exact parameters bounds defined in (14) and (15). In particular, for a given relaxation order  $\delta$ , relaxation of (14) and (15) through dense LMI-relaxation leads to SDP problems where the number of variables is  $O(N^{2\delta})$  and the size of the largest LMI defining the feasible region is  $O(N^\delta)$ . Thus, the use of the dense LMI-relaxation technique is limited to Hammerstein system identification problems with a small number  $N$  of measurements (in general not greater than 5) because of high computational burden. In order to handle a larger number of measurements, the particular structure of the identification problems (14) and (15) has been analyzed to apply the *sparse* LMI-relaxation presented in Section II-C. The following result shows that problems (14) and (15) have inherent structured sparsity.

**Property 1:** Problems (14) (resp. (15)) enjoy the following features:

**P 1.1:** The functional involves only the variable  $\gamma_k$  (resp.  $\theta_j$ ).

**P 1.2:** For all  $t = na + 1, \dots, N$ , the bilinear constraints  $g_t \geq 0$  and  $g_{t+N} \geq 0$  depend only on  $n_\gamma + n_\theta + na + 1$  variables, i.e.  $n_\gamma$  nonlinear block parameters  $\gamma$ ;  $n_\theta$  linear block parameters  $\theta$  and  $na + 1$  noise samples  $\eta_{t-i}$ , for  $i = 0, 1, \dots, na$ .

**P 1.3:** For all  $r = 1, \dots, N$ , constraints  $g_{r+2N} \geq 0$  and  $g_{r+3N} \geq 0$  depend only on the noise variable  $\eta_r$ .

**P 1.4:** The linear constraints  $g_{4N+1} \geq 0$  and  $g_{4N+2} \geq 0$  depend only on the system parameters  $\theta$ . ■

A sparsity pattern in the identification problems (14) and (15) has been detected by exploiting results provided in Property 1. This allows the formulation of sparse SDP-relaxed problems for (14) and (15) as described in the following.

Let  $X \in \mathbb{R}^{n_\gamma + n_\theta + N}$  be the collection of the optimization variables for the identification problems (14) and (15), i.e.  $X = [\gamma^T \theta^T \eta^T]^T$  and let  $X_i$  the  $i$ -th component of the vector  $X$ . In such a way, the first  $n_\gamma$  components of  $X$  are the nonlinear block parameters  $\gamma$ , the components from position  $n_\gamma + 1$  to  $n_\gamma + n_\theta$  are the linear block parameters  $\theta$ , while the components from position  $n_\gamma + n_\theta + 1$  to  $n_\gamma + n_\theta + N$  are the noise variables  $\eta$ . Let us define the index sets  $\mathcal{I}_r \subset \{1, 2, \dots, n_\gamma + n_\theta + N\}$  and  $\mathcal{S}_r \subset \{na + 1, \dots, N, N + na + 1, \dots, 2N + 1, \dots, 4N + 2\}$  as

$$\begin{aligned} \mathcal{I}_r = \{ & 1, 2, \dots, n_\gamma + n_\theta, \\ & n_\gamma + n_\theta + r, n_\gamma + n_\theta + r + 1, \dots, n_\gamma + n_\theta + r + na \} \\ & \text{for } r = 1, \dots, N - na \end{aligned} \quad (16)$$

$$\begin{aligned} \mathcal{S}_1 = \{ & na + 1, N + na + 1, \\ & 2N + 1, 2N + 2, \dots, 2N + na + 1, \\ & 3N + 1, 3N + 2, \dots, 3N + na + 1, 4N + 1, 4N + 2 \} \end{aligned} \quad (17)$$

$$\begin{aligned} \mathcal{S}_r = \{ & na + r, N + na + r, 2N + na + r, 3N + na + r \}, \\ & \text{for } r = 2, \dots, N - na \end{aligned} \quad (18)$$

The index sets  $\mathcal{I}_r$  and  $\mathcal{S}_r$  defined in (16)-(18) enjoy the following features.

**Property 2:** For all  $r = 1, \dots, N - na$ , the index sets  $\mathcal{I}_r$  and  $\mathcal{S}_r$  are such that:

**P 2.1:** The set of the variables indexes  $\mathcal{I}_0 = \{1, 2, \dots, n_\gamma + n_\theta + N\}$  is the union of the sets  $\mathcal{I}_r$ , that is  $\mathcal{I}_0 = \bigcup_{r=1}^{N-na} \mathcal{I}_r$ .

**P 2.2:** The set of the constraints indexes  $\mathcal{S}_0 = \{na + 1, \dots, N, N + na + 1, \dots, 2N, 2N + 1, \dots, 4N + 2\}$  defining  $\mathcal{D}_{\gamma\theta\eta}$  is the union of the sets  $\mathcal{S}_r$ , that is  $\mathcal{S}_0 = \bigcup_{r=1}^{N-na} \mathcal{S}_r$ .

**P 2.3:** The sets  $\mathcal{S}_r$  are mutually disjoint.

**P 2.4:** For every  $s \in \mathcal{S}_r$ , the polynomial constraint  $g_s(\gamma, \theta, \eta) \geq 0$  defining  $\mathcal{D}_{\gamma\theta\eta}$  depends only on the variables  $X(\mathcal{I}_r) = \{X_i : i \in \mathcal{I}_r\}$ .

**P 2.5:** The functional of identification problems (14) and (15) depends only on the variables  $X(\mathcal{I}_r) = \{X_i : i \in \mathcal{I}_r\}$ .

**P 2.6:** For every  $r = 1, \dots, N - na - 1$ ,

$$\mathcal{I}_{r+1} \cap \bigcup_{j=1}^r \mathcal{I}_j \subseteq \mathcal{I}_r. \quad \blacksquare$$

Now, for a given relaxation order  $\delta \geq 1$ , let us consider the SDP problems

$$\underline{\gamma}_k^\delta = \min_{p \in \mathcal{D}_{\gamma\theta\eta}^\delta} \sum_{\alpha \in \mathcal{A}_{2\delta}} \Gamma_{k\alpha} p_\alpha, \quad \bar{\gamma}_k^\delta = \max_{p \in \mathcal{D}_{\gamma\theta\eta}^\delta} \sum_{\alpha \in \mathcal{A}_{2\delta}} \Gamma_{k\alpha} p_\alpha \quad (19)$$

$$\underline{\theta}_j^\delta = \min_{p \in \mathcal{D}_{\gamma\theta\eta}^\delta} \sum_{\alpha \in \mathcal{A}_{2\delta}} \Theta_{j\alpha} p_\alpha, \quad \bar{\theta}_j^\delta = \max_{p \in \mathcal{D}_{\gamma\theta\eta}^\delta} \sum_{\alpha \in \mathcal{A}_{2\delta}} \Theta_{j\alpha} p_\alpha \quad (20)$$

where  $\Gamma_k = \{\Gamma_{k\alpha}\}_{\alpha \in \mathcal{A}_{2\delta}}$  (resp.  $\Theta_j = \{\Theta_{j\alpha}\}_{\alpha \in \mathcal{A}_{2\delta}}$ ) is the coefficient vector of the function  $\gamma_k$  (resp.  $\theta_j$ ) in the basis  $h = \{X^\alpha\}_{\alpha \in \mathcal{A}_{2\delta}}$ , which is the canonical basis of the real-valued polynomials of degree  $2\delta$  in the variables vector  $X$ . The feasible region  $\mathcal{D}_{\gamma\theta\eta}^\delta$  is a convex set defined as

$$\mathcal{D}_{\gamma\theta\eta}^\delta = \{p : M_\delta(p, \mathcal{I}_r) \succeq 0, r = 1, \dots, N - na \\ M_{\delta-1}(g_s p, \mathcal{I}_r) \succeq 0, s \in \mathcal{S}_r, r = 1, \dots, N - na\} \quad (21)$$

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

**Remark 1:** Since the linear block is known to be stable, stability constraints on the parameters  $a_1, \dots, a_{na}$  can be enforced in problems (14) and (15) with the method presented in [23] in order to tighten parameter bounds. ■

The  $\delta$ -relaxed uncertainty intervals, defined as  $PUI_{\gamma_k}^\delta = [\underline{\gamma}_k^\delta; \bar{\gamma}_k^\delta]$  and  $PUI_{\theta_j}^\delta = [\underline{\theta}_j^\delta; \bar{\theta}_j^\delta]$ , enjoy the following properties.

**Property 3:** For all  $k = 1, \dots, n_\gamma$  and relaxation order  $\delta \geq 1$ , the  $\delta$ -relaxed uncertainty interval  $PUI_{\gamma_k}^\delta$  satisfy the following properties.

**P 3.1: Guaranteed relaxed uncertainty intervals.**

The interval  $PUI_{\gamma_k}^\delta$  is guaranteed to contain the true parameter  $\gamma_k$  to be estimated, i.e.  $\gamma_k \in PUI_{\gamma_k}^\delta$ .

**P 3.2: Monotone convergence to tight uncertainty intervals.** The interval  $PUI_{\gamma_k}^\delta$  becomes tighter as the relaxation order  $\delta$  increases, that is  $PUI_{\gamma_k}^{\delta+1} \subseteq PUI_{\gamma_k}^\delta$ . Besides,  $PUI_{\gamma_k}^\delta$  converges to the tight interval  $PUI_{\gamma_k}$  as the LMI relaxation order goes to infinity, that is:

$$\lim_{\delta \rightarrow \infty} \underline{\gamma}_k^\delta = \underline{\gamma}_k, \quad \lim_{\delta \rightarrow \infty} \bar{\gamma}_k^\delta = \bar{\gamma}_k. \quad (22) \quad \blacksquare$$

The proof of Properties P3.1 and P3.2 (see [24] for details) follows from the properties of the index sets  $\mathcal{I}_r$  and  $\mathcal{S}_r$  highlighted in Property 2 and direct application of Theorem 1 to problems (14), (15) and the corresponding SDP-relaxed problems (19), (20). Similar results to Property 3 hold for the relaxed intervals  $PUI_{\theta_j}^\delta$ .

**Property 4: Computational complexity of the SDP-problems (19) and (20)**

- (i) The number of free optimization 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  $\mathcal{D}_{\gamma\theta\eta}^\delta$  is described by:
  - $N - na$  moment matrixes, each one of size  $\binom{n_\gamma + n_\theta + na + 1 + \delta}{\delta}$ ,
  - $2(N - na) + 2N + 2$  localizing matrixes, each one of size  $\binom{n_\gamma + n_\theta + na + \delta}{\delta - 1}$ . ■

Technical details on the computation of the number of optimization variables  $p$  and dimension of the LMIs describing  $\mathcal{D}_{\gamma\theta\eta}^\delta$  in (19) and (20) are reported in [24].

## VI. A SIMULATED 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 2.40-GHz Intel Pentium IV with 3 GB RAM. The nonlinear block of the Hammerstein system considered here is modeled by the function  $z_t = 0.3u_t + 0.2u_t^2 - u_t^3$ , i.e.  $\gamma^T = [\gamma_1 \ \gamma_2 \ \gamma_3] = [0.3 \ 0.2 \ -1]$ , while the linear part is a second order system with parameters  $\theta^T = [a_1 \ a_2 \ b_1 \ b_2] = [1.8 \ 0.9 \ 1.6 \ 2.1]$ . The input is a random sequence uniformly distributed between  $[-1, +1]$ . Two different numerical experiments are performed. In the first one, only  $N = 50$  number of measurements are exploited to compute the parameters bounds, while in the second experiment  $N = 250$  data are used. The noise-free output  $w_t$  is corrupted by 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  $SNR_w = 10 \log \left\{ \frac{\sum_{t=1}^N w_t^2}{\sum_{t=1}^N \eta_t^2} \right\}$  is equal to 27 db. Bounds on the parameters are evaluated by solving (19) and (20) for a relaxation order  $\delta = 2$ . Stability constraints on the linear block are enforced through the method described in [23]. Note that, in the considered example with  $N = 50$ , the number of optimization variable in (19) and (20) is 22,275, while the feasible region is defined by 98 moment matrixes of size 55 and 396 localizing matrixes of size 10. On the other hand, if the sparsity was not taken into account, the number of variable of the SDP relaxed problems would be about 6 million and the feasible region would be described by a moment matrix of size 5,886 and 396 localizing matrixes of size 108, leading to an untractable optimization problem in the employed workstation. Results about the nonlinear and the linear block are reported in Table I and II, respectively, which show the obtained parameter bounds, the central estimates  $\gamma_k^c$  and  $\theta_j^c$  as well as the parameter uncertainties  $\Delta\gamma_k$

Table I: Nonlinear block. Parameter central estimates ( $\gamma_k^c$ ), parameter bounds ( $\underline{\gamma}_k, \bar{\gamma}_k$ ) and parameter uncertainties  $\Delta\gamma_k$  for  $N = 50$  and  $N = 250$  measurements.

N	Parameter	True Value	$\underline{\gamma}_k$	$\gamma_k^c$	$\bar{\gamma}_k$	$\Delta\gamma_k$
50	$\gamma_1$	0.300	0.206	0.295	0.384	0.089
	$\gamma_2$	0.200	0.121	0.190	0.259	0.069
	$\gamma_3$	-1.000	-1.211	-1.000	-0.789	0.211
250	$\gamma_1$	0.300	0.259	0.299	0.340	0.041
	$\gamma_2$	0.200	0.173	0.203	0.232	0.030
	$\gamma_3$	-1.000	-1.103	-1.000	-0.898	0.103

Table II: Linear block. Parameter central estimates ( $\theta_j^c$ ), parameter bounds ( $\underline{\theta}_j, \bar{\theta}_j$ ) and parameter uncertainties  $\Delta\theta_j$  for  $N = 50$  and  $N = 250$  measurements.

N	Parameter	True Value	$\underline{\theta}_j$	$\theta_j^c$	$\bar{\theta}_j$	$\Delta\theta_j$
50	$a_1$	1.800	1.758	1.800	1.842	0.042
	$a_2$	0.900	0.856	0.899	0.941	0.043
	$b_1$	1.600	1.418	1.626	1.834	0.208
	$b_2$	2.100	1.882	2.087	2.292	0.205
250	$a_1$	1.800	1.788	1.800	1.812	0.012
	$a_2$	0.900	0.888	0.900	0.911	0.012
	$b_1$	1.600	1.579	1.600	1.620	0.021
	$b_2$	2.100	2.071	2.099	2.126	0.028

and  $\Delta\theta_j$  defined as  $\gamma_k^c = \frac{\bar{\gamma}_k + \underline{\gamma}_k}{2}$ ,  $\theta_j^c = \frac{\bar{\theta}_j + \underline{\theta}_j}{2}$ ,  $\Delta\gamma_k = \frac{\bar{\gamma}_k - \underline{\gamma}_k}{2}$  and  $\Delta\theta_j = \frac{\bar{\theta}_j - \underline{\theta}_j}{2}$ . The CPU time taken by SeDuMi to solve a single problem (19) and (20) is between 223 s and 263 s when the number of measurements  $N$  is equal to 50, and between 2442 s and 2578 s when  $N = 250$ . The reported results show that as the number of data increases (from  $N = 50$  to  $N = 250$ ), the width of uncertainty intervals  $\Delta\gamma_k$  and  $\Delta\theta_j$  decreases, as expected. Furthermore, we point out that the presented procedure provides satisfactory parameter uncertainty intervals, also for a small set of data ( $N = 50$ ).

## VII. CONCLUSIONS

A single-stage procedure for the evaluation of parameter uncertainty intervals for Hammerstein systems is presented. The parameters bounds evaluation is formulated in terms of constrained polynomial optimization problems, whose approximate solutions can be computed by a hierarchy of SDP problems, which guarantees monotone convergence to global optima as the relaxation order increases. The particular structure of the original optimization problems allows a reduction of the computational complexity of the SDP relaxed problems, preserving convergence to tight parameter bounds. The presented identification procedure can be also applied to the case of noise corrupted input sequence. As a matter of fact, in such a case, the identification problem can be still formulated in terms of sparse polynomial optimization problem by considering the noise samples on the input signal as variables.

## REFERENCES

[1] E. Bai and F. Giri, *Block-oriented nonlinear system identification*, ser. Lecture notes in Control and Information sciences. Berlin: Springer.

[2] F. Chang and R. Luus, "A noniterative method for identification using Hammerstein model," *IEEE Trans. Automatic Control*, vol. AC-16, pp. 464–468, 1971.

[3] M. Verhaegen and D. Westwick, "Identifying MIMO Hammerstein systems in the context of subspace model identification methods," *Int. J. Control*, vol. 63, no. 2, pp. 331–349, 1996.

[4] E. Bai and M. Fu, "A blind approach to Hammerstein model identification," *IEEE Trans. Signal Processing*, vol. 50, no. 7, pp. 1610–1619, 2002.

[5] E. Bai and D. Li, "Convergence of the iterative Hammerstein system identification algorithm," *IEEE Trans. Automatic Control*, vol. 49, no. 11, pp. 1929–1940, 2004.

[6] W. Greblicki and M. Pawlak, "Nonparametric identification of Hammerstein systems," *IEEE Trans. Automatic Control*, vol. 35, no. 2, pp. 409–418, 1989.

[7] A. Krzyżak, "On nonparametric estimation of nonlinear dynamic systems by the Fourier series estimate," *Signal Processing*, vol. 52, pp. 299–321, 1996.

[8] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty: an overview," *Automatica*, vol. 27(6), pp. 997–1009, 1991.

[9] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded-error data: a survey," *Mathematics and Computers in simulation*, vol. 32, pp. 449–468, 1990.

[10] M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, Eds., *Bounding approaches to system identification*. New York: Plenum Press, 1996.

[11] M. Sznaier, "Computational complexity analysis of set membership identification of Hammerstein and Wiener systems," *Automatica*, vol. 45, no. 3, pp. 701–705, 2009.

[12] V. Cerone and D. Regruto, "Parameter bounds for discrete-time Hammerstein models with bounded output errors," *IEEE Trans. Automatic Control*, vol. 48, no. 10, pp. 1855–1860, 2003.

[13] M. Laurent, "Sums of squares, moment matrices and optimization over polynomials," *Emerging Applications of Algebraic Geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications*, M. Putinar and S. Sullivant (eds.), pp. 157–270, 2009.

[14] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optim.*, vol. 11, pp. 796–817, 2001.

[15] D. Henrion and J. B. Lasserre, "Solving nonconvex optimization problems," *IEEE Contr. Systems Magazine*, vol. 24, no. 3, pp. 72–83, 2004.

[16] M. Kojima, S. Kim, and H. Waki, "Sparsity in sums of squares of polynomials," *Mathematical Programming*, vol. 103, pp. 45–62, 2005.

[17] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, "Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity," *SIAM Journal on Optim.*, vol. 17, no. 1, pp. 218–242, 2006.

[18] P. Parrillo, "Exploiting algebraic structure in sum of squares programs," *Positive Polynomials in Control*, D. Henrion and A. Garulli (eds.), pp. 181–194, 2005.

[19] J. Lasserre, "Convergent semidefinite relaxations in polynomial optimization with sparsity," *SIAM Journal on Optim.*, vol. 17, no. 1, pp. 822–843, 2006.

[20] H. Waki, S. Kim, M. Kojima, M. Muramatsu, and H. Sugimoto, "SparsePOP: a sparse semidefinite programming relaxation of polynomial optimization problems," *ACM Trans. on Mathematical Software*, vol. 35, no. 2, 2008.

[21] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB Toolbox for optimization over symmetric cones," *Optim. Methods Software*, vol. 11, no. 12, pp. 625–653, 1999.

[22] K. Fujisawa, M. Fukuda, K. Kobayashi, M. Kojima, K. Nakata, M. Nakata, and M. Yamashita, "SDPA (semidefinite programming algorithm) users manual version 7.0.5," *Research Report B-448, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology*, 2008.

[23] V. Cerone, D. Piga, and D. Regruto, "Bounding the parameters of linear systems with stability constraints," in *Proc. of the American Control Conference 2010*, 2010, pp. 2152–2157.

[24] —, "Hammerstein systems identification through convex relaxation techniques," *Internal Report*, vol. DAUIN, no. 2010/09/24, 2010.