Set-Membership identification of Hammerstein-Wiener systems

Vito Cerone, Dario Piga, Diego Regruto

Fig. 1.   Hammerstein-Wiener system.

Abstract— Set-membership identification of Hammerstein-Wiener models is addressed in the paper. First, it is shown that computation of tight parameter bounds requires the solutions to a number of nonconvex constrained polynomial optimization problems where the number of decision variables increases with the length of the experimental data sequence. Then, a suitable convex relaxation procedure is presented to significantly reduce the computational burden of the identification problem. A detailed discussion of the identification algorithm properties is reported. Finally, a simulated example is used to show the effectiveness and the computational tractability of the proposed approach.

I. INTRODUCTION

Identification of nonlinear systems has been an active research area in the last decades. Despite the rich literature available on the subject (see, e.g., the survey paper [1] and the special issue [2]), nonlinear dynamic systems modeling and identification still remains a challenging task worthy of further studies, as remarked in recent plenary talks [3], [4]. One of the central issues is the search for simple and flexible model structures which 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, modeled by interconnected memoryless nonlinear gains and linear dynamic subsystems. 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, these models are successfully employed in many different engineering fields. Early works on identification of block-structured nonlinear systems are summarized in the survey paper [5] while an up-to-date collection of results and algorithms can be found in the book [6]. The configuration we are dealing with in this note, commonly referred to as a Hammerstein-Wiener model, is shown in Fig. 1; it consists of a linear dynamic system sandwiched by two static non-linearities $\mathcal{N}_1$ and $\mathcal{N}_2$. Identification of such a model relies solely on input-output measurements, while the inner signals x_t and z_t are not assumed to be available. A good deal of approaches can be found in the literature which address such a problem. A two stage algorithm based on recursive least squares and singular values decomposition is proposed in [7], while a blind approach is considered in [8]. The asymptotic properties of the least-squares estimates of Hammerstein-Wiener models are investigated in [9]. Iterative identification
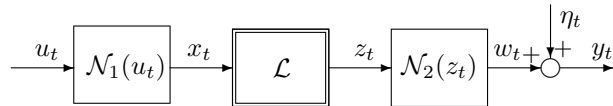
schemes are proposed in [10] and [11] while an extended stochastic gradient algorithm is presented in [12] for the case of Hammerstein-Wiener ARMAX models. Subspace algorithms, exploiting either open-loop [13] or closed-loop [14] input-output data, have been recently investigated. In all the papers mentioned 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-errors, or set-membership, characterization where uncertainties are assumed to belong to a given set. The interested reader can find further details on this approach in a number of survey papers (see, e.g., [15], [16]). To our best knowledge, no contribution can be found in the literature which addresses the identification of Hammerstein-Wiener models when the measurement error is supposed to be bounded. In this paper we consider the identification of single-input single-output (SISO) Hammerstein-Wiener models where the nonlinear static gains are described by the linear combination of a finite number of known basis functions, an output error structure is used to model the linear dynamic part and the output measurement errors are corrupted by bounded errors. It is worth noting that the problem of set-membership identification of the Hammerstein and the Wiener models, which are particular cases of the Hammerstein-Wiener structure considered here, is NP-hard in the size of the experimental data sequence, as shown in [17]. The paper is organized as follows. Section II is devoted to the formulation of the identification problem. In Section III we show that computation of tight parameters bounds requires the solution to constrained nonconvex optimization problems where the number of decision variables increases with the number of measured data. A suitable relaxation procedure to reduce the number of decision variables for such optimization problems is presented in Section IV and a detailed analysis of its properties is reported in Section V. A simulated example is presented in Section VI in order to show the effectiveness of the proposed approach.

II. PROBLEM FORMULATION

Consider the Hammerstein-Wiener model depicted in Fig. 1. The input nonlinearity $\mathcal{N}_1(\cdot)$ maps the input signal u_t into the unmeasurable inner signal x_t through the static nonlinear

V. Cerone is with Dipartimento di Automatica e Informatica, Politecnico di Torino, 10129 Torino, Italy `vito.cerone@polito.it`

D. Piga is with Dipartimento di Automatica e Informatica, Politecnico di Torino, 10129 Torino, Italy `dario.piga@polito.it`

D. Regruto is with Dipartimento di Automatica e Informatica, Politecnico di Torino, 10129 Torino, Italy `diego.regruto@polito.it`

function

$$x_t = \mathcal{N}_1(u_t) = \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_t), \quad t = 1, \dots, N, \qquad (1)$$

where $(\psi_1, \dots, \psi_{n_\gamma})$ is a known basis of nonlinear functions and N is the length of data sequence. The output nonlinearity $w_t = \mathcal{N}_2(z_t)$ as well is a static function and, as assumed in [8], it is a one-to-one nonlinearity so that the inverse $z_t = \mathcal{N}_2^{-1}(w_t)$ exists and it is parameterized as

$$z_t = \mathcal{N}_2^{-1}(w_t) = \sum_{j=1}^{n_\alpha} \alpha_j \phi_j(w_t), \quad t = 1, \dots, N, \quad (2)$$

where $(\phi_1, \dots, \phi_{n_\alpha})$ is a known basis of nonlinear continuous functions. The linear dynamic part $\mathcal{L}$ is modeled by a discrete-time stable system transforming x_t into the signal z_t according to the linear difference equation

$$z_t = -\sum_{i=1}^{na} a_i z_{t-i} + \sum_{j=0}^{nb} b_j x_{t-j}. \qquad (3)$$

Internal signals x_t and z_t are supposed not to be measurable, while the measurements y_t of the output signal w_t are corrupted by additive noise according to

$$y_t = w_t + \eta_t, \qquad (4)$$

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

$$\mid \eta_t \mid \le \Delta\eta_t. \qquad (5)$$

The unknown parameters to be estimated are collected in the vectors $\gamma = [\gamma_1, \dots, \gamma_{n_\gamma}]^\mathrm{T}$, $\alpha = [\alpha_1, \dots, \alpha_{n_\alpha}]^\mathrm{T}$ and $\theta = [a_1 \ \dots \ a_{na} \ b_0 \ b_1 \ \dots b_{nb}]^\mathrm{T} \in \mathbb{R}^{n_\theta}$, where $n_\theta = na + nb + 1$. It must be pointed out that the parametrization of the structure in Fig. 1 is not unique. In fact, as shown in [8], any parameter set $a\gamma$, $b[b_0, \dots, b_{nb}]$ and $c\alpha$, for some constants a, b, c such that $abc = 1$, provides the same input-output behavior. In order to get a unique parametrization, two blocks have to be normalized. In this work we assume, without loss of generality, that $\gamma_1 = 1$ and $\alpha_1 = 1$.

In this paper we address the problem of deriving bounds on the parameters γ, α and θ consistently with the assumed model structure, error bounds and measurements of the input signal u_t and noise-corrupted output y_t.

III. EVALUATION OF TIGHT PARAMETER UNCERTAINTY INTERVALS

In this section we show that the evaluation of tight parameter uncertainty intervals requires the solution to a set of nonconvex optimization problems with $n_\gamma + n_\alpha + n_\theta + N$ variables. By substituting (1) into (3) we get:

$$z_t = -\sum_{i=1}^{na} a_i z_{t-i} + \sum_{j=0}^{nb} b_j \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{t-j}). \qquad (6)$$

From eqs. (2), (4) and (6), the following relation between the output signal y_t and the input signal u_t holds:

$$\sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_t - \eta_t) + \sum_{i=1}^{na} a_i \sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{t-i} - \eta_{t-i}) =$$
$$= \sum_{j=0}^{nb} b_j \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{t-j}). \qquad (7)$$

Then, the set $\mathcal{D}$ of all system parameters (γ, α, θ) and noise samples η_t consistent with measurements, error bounds and the assumed model structure is described by (5) and (7), i.e.

$$\mathcal{D} = \Big\{ (\gamma, \alpha, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\alpha + n_\theta + N} :$$
$$\sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_t - \eta_t) + \sum_{i=1}^{na} a_i \sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{t-i} - \eta_{t-i}) =$$
$$= \sum_{j=0}^{nb} b_j \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{t-j}), \quad t = na+1, \dots, N; \qquad (8)$$
$$|\eta_r| \le \Delta\eta_r; \ \ r = 1, \dots, N; \ \gamma_1 = 1, \ \ \beta_1 = 1 \Big\},$$

with $\eta = [\eta_1, \dots, \eta_N]^\mathrm{T}$. Therefore, tight bounds on the parameters γ_i, α_j and θ_k can be computed by solving the constrained optimization problems

$$\underline{\gamma}_i = \min_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \gamma_i, \quad \overline{\gamma}_i = \max_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \gamma_i, \qquad (9)$$

$$\underline{\alpha}_j = \min_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \alpha_i, \quad \overline{\alpha}_j = \max_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \alpha_i, \qquad (10)$$

$$\underline{\theta}_k = \min_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \theta_k, \quad \overline{\theta}_k = \max_{(\gamma, \alpha, \theta, \eta) \in \mathcal{D}} \theta_k. \qquad (11)$$

It must be pointed out that the number of optimization variables of problems (9)-(11) increases with the number of measurements N. Furthermore, problems (9)-(11) are, in general, nonconvex since the equality constraints defining the feasible region $\mathcal{D}$ involve the product between the unknown parameters γ, α and θ and the nonlinear functions $\psi_i(.)$, which, in turn, depend on the noise variables η. Therefore, standard nonlinear optimization tools (e.g., gradient method, Newton method) cannot be used since they can trap in local minima. As a consequence, the computed uncertainty intervals are not guaranteed to contain the true parameters. A possible solution to overcome such a problem is to relax the identification problems (9)-(11) to convex optimization problems, in order to numerically compute relaxed bounds on the system parameters. It must be pointed out that when $\phi_j(.)$ are polynomial functions, problems (9)-(11) are sparse semialgebraic optimization problems and they can be relaxed through a direct implementation of the LMI-relaxation for sparse polynomial optimization proposed in [18] and [19]. Unfortunately, due to a large number of variables appearing in (9)-(11) and a high degree of polynomial equalities defining the feasible set $\mathcal{D}$, the relaxation of such problems through LMI-based relaxation techniques leads to untractable SDP-problems because of high computational load. In the following section we present a relaxation procedure that significantly reduces the computation burden of identification

problems (9)-(11). Besides, the presented procedure can be also applied when $\phi_j(.)$ are not polynomial functions.

IV. COMPUTATIONAL BURDEN REDUCTION

The key idea to reduce the computational burden of problems (9)-(11) is to construct an outer-bounding set $\mathcal{D}^{ss(n)}$ of $\mathcal{D}$ by considering the noise variables appearing in the definition of $\mathcal{D}$ independent of each others. Indeed, in such a way, conservativeness is introduced in the evaluation of the parameter bounds since the correlation between consecutive measurements is lost. In order to reduce the conservativeness of such an approach, equality constraints are added in the definition of the feasible set $\mathcal{D}$ as follows: (i) given a fixed integer n, we consider n consecutive equality constraints in the description of $\mathcal{D}$ in (8); (ii) we substitute the first constraint in the second one obtaining a new equality that retains correlations among the two; then, the newly obtained equation is substituted into the third constraint. The procedure is repeated until all the n consecutive constraints selected in (i) are nested; (iii) the new equations obtained in (ii) are added in the description of $\mathcal{D}$.

Then, $\mathcal{D}$ is outer-bounded by a set $\mathcal{D}^{ss(n)}$ obtained by considering the noise variables appearing in the new description of $\mathcal{D}$ independent of each others. 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 $\mathcal{D}^{ss(n)}$, computation of parameter bounds can be formulated in terms of polynomial optimization problems with only $n_\gamma + n_\alpha + n_\theta$ variables, unlike $n_\gamma + n_\alpha + n_\theta + N$ variables involved in identification problems (9)-(11). Technical details of the proposed relaxation procedure are now presented. First, the feasible set $\mathcal{D}$ is written as intersection of $N-na$ sets. In particular, for a given integer $n \in [1, N-na]$, the set $\mathcal{D}$ is written as $\mathcal{D} = \bigcap_{z=1}^{N-na} S_z^{(n)}$, where

$$
S_z^{(n)} = \Big\{ (\gamma, \alpha, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\alpha + n_\theta + N} :
$$
$$
\sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{na+z+s-1} - \eta_{na+z+s-1}) =
$$
$$
= -\sum_{i=1}^{na} a_i \sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{na+z+s-1-i} - \eta_{na+z+s-1-i}) +
$$
$$
+ \sum_{j=0}^{nb} b_j \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{na+z+s-1-j}), \tag{12}
$$
$$
|\eta_r| \le \Delta\eta_r, \;\; \gamma_1 = 1, \;\; \beta_1 = 1;
$$
$$
s = 1, 2, \ldots, \min\{n, N - z + 1\};
$$
$$
r = z, z+1, \ldots, \min\{na + n + z, N\} \Big\}.
$$

It is worth noting that each set $\mathcal{S}_z^{(n)}$ is described by the constraints defining $\mathcal{D}$ in (8) obtained by at most n consecutive measurements. For instance, $S_1^{(n)}$ is only described by the constraints obtained by the measurements from time $t = na + 1$ up to time $t = na + n$, $\mathcal{S}_2^{(n)}$ is described by the constraints obtained by the measurements from time $t = na + 2$ up to time $t = na + n + 1$, and so on, up to $\mathcal{S}_{N-na}^{(n)}$ that is defined only by the constraint obtained by the

measurement at time N. In this work we refer to n as the *dynamic-horizon*. On the basis of the definition of $\mathcal{S}_z^{(n)}$ in (12) and from the nested substitution process described in (ii), an alternative description of the sets $\mathcal{S}_z^{(n)}$ can be given as

$$
S_z^{(n)} = \Big\{ (\gamma, \alpha, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\alpha + n_\theta + N} :
$$
$$
\sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{na+z+s-1} - \eta_{na+z+s-1}) +
$$
$$
+ \sum_{h=1}^{na} [A^s](1, h) \sum_{j=1}^{n_\alpha} \alpha_j \phi_j(y_{na+z-h} - \eta_{na+z-h}) =
$$
$$
= \sum_{j=1}^{s} \sum_{h=1}^{nb+1} [A^{s-j}B](1, h) \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{na+z+s-1-h}); \tag{13}
$$
$$
|\eta_r| \le \Delta\eta_r, \;\; \gamma_1 = 1, \;\; \beta_1 = 1;
$$
$$
s = 1, 2, \ldots, \min\{n, N - z + 1\};
$$
$$
r = z, z+1, \ldots, \min\{na + n + z, N\} \Big\},
$$

where

$$
\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \ldots & a_{na-1} & a_{na} \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix} \in \mathbb{R}^{na,na}, \tag{14}
$$

$$
\mathbf{B} = \begin{bmatrix} b_0 & b_1 & \cdots & b_{nb} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{na,nb+1}, \tag{15}
$$

where $[R](1, h)$ denotes the entry in the first row and in the h-th column of a generic matrix R. Outer-bounds $\mathcal{S}_z^{ss(n)}$ and $\mathcal{D}^{ss(n)}$ of $\mathcal{S}_z^{(n)}$ and $\mathcal{D}$, respectively, are now constructed as described by next results. Proofs of all results and properties presented in this paper can be found in [20].

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

$$
\mathcal{S}_z^{ss(n)} = \Big\{ (\gamma, \alpha, \theta) \in \mathbb{R}^{n_\gamma + n_\alpha + n_\theta} :
$$
$$
\sum_{j=1}^{n_\alpha} \alpha_j \overline{\phi}(\alpha_j, y_{na+z+s-1}) +
$$
$$
+ \sum_{h=1}^{na} \sum_{j=1}^{n_\alpha} [A^s](1, h) \alpha_j \overline{\phi}([A^s](1, h) \alpha_j, y_{na+z-h}) \ge
$$
$$
\sum_{j=1}^{s} \sum_{h=1}^{nb+1} [A^{s-j}B](1, h) \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{na+z+s-1-h});
$$
$$
\sum_{j=1}^{n_\alpha} \alpha_j \underline{\phi}(\alpha_j, y_{na+z+s-1}) +
$$

$$+ \sum_{h=1}^{na} \sum_{j=1}^{n_\alpha} [A^s](1,h)\alpha_j \underline{\phi}([A^s](1,h)\alpha_j, y_{na+z-h}) \leq$$

$$\sum_{j=1}^{s} \sum_{h=1}^{nb+1} [A^{s-j}B](1,h) \sum_{i=1}^{n_\gamma} \gamma_i \psi_i(u_{na+z+s-1-h});$$

$$\gamma_1 = 1, \ \beta_1 = 1; \ s = 1,2,\ldots,\min\{n, N-z+1\}\}.$$

where $\overline{\phi}, \underline{\phi} : \mathbb{R}^2 \to \mathbb{R}$ are real-valued functions defined as

$$\overline{\phi}(a, y_t) = \begin{cases} \max_{|\eta_t| \leq \Delta\eta_t} \phi(y_t - \eta_t) & \text{if } a \geq 0, \\ \min_{|\eta_t| \leq \Delta\eta_t} \phi(y_t - \eta_t) & \text{if } a < 0. \end{cases} \quad (16)$$

$$\underline{\phi}(a, y_t) = \begin{cases} \min_{|\eta_t| \leq \Delta\eta_t} \phi(y_t - \eta_t) & \text{if } a \geq 0, \\ \max_{|\eta_t| \leq \Delta\eta_t} \phi(y_t - \eta_t) & \text{if } a < 0. \end{cases} \quad (17)$$

Thus, for every $z = 1,\ldots,N-na$, the set $\mathcal{S}_z^{ss(n)}$ is an outer approximation of $\mathcal{S}_z^{(n)}$, i.e. $\mathcal{S}_z^{(n)} \subseteq \mathcal{S}_z^{ss(n)}$. ∎

***Result 2:* Construction of an outer bound of $\mathcal{D}$.**
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 set $\mathcal{D}$, i.e. $\mathcal{D} \subseteq \mathcal{D}^{ss(n)}$. ∎

Thus, for a given dynamic horizon n, and for all $i = 1,\ldots,n_\gamma$, $j = 1,\ldots,n_\alpha$ and $k = 1,\ldots,n_\theta$, the parameter uncertainty intervals on γ, α and θ, defined as $PUI_{\gamma_i}^{ss(n)} = \left[\underline{\gamma}_i^{ss(n)}; \overline{\gamma}_i^{ss(n)}\right]$, $PUI_{\alpha_j}^{ss(n)} = \left[\underline{\alpha}_j^{ss(n)}; \overline{\alpha}_j^{ss(n)}\right]$, $PUI_{\theta_k}^{ss(n)} = \left[\underline{\theta}_k^{ss(n)}; \overline{\theta}_k^{ss(n)}\right]$, can be evaluated by solving the optimization problems

$$\underline{\gamma}_i^{ss(n)} = \min_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \gamma_i; \ \overline{\gamma}_i^{ss(n)} = \max_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \gamma_i, \quad (18)$$

$$\underline{\alpha}_j^{ss(n)} = \min_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \alpha_j; \ \overline{\alpha}_j^{ss(n)} = \max_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \alpha_j, \quad (19)$$

$$\underline{\theta}_k^{ss(n)} = \min_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \theta_k; \ \overline{\theta}_k^{ss(n)} = \max_{\gamma,\alpha,\theta \in \mathcal{D}^{ss(n)}} \theta_k. \quad (20)$$

Remark 1: Only the unknown parameters γ, α and θ are optimization variables for problems (18) and (20); on the contrary, in the original identification problems (9)-(11) also the noise samples η are treated as variables. ∎

Evaluation of intervals $PUI_{\gamma_i}^{ss(n)}$, $PUI_{\alpha_j}^{ss(n)}$ and $PUI_{\theta_k}^{ss(n)}$ requires the solution to the optimization problems (18)-(20) over the nonconvex feasible region $\mathcal{D}^{ss(n)}$. 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 the topological features of the set $\mathcal{D}^{ss(n)}$, we first introduce the following notation.
Let $q^{(n)} \in \mathbb{R}^{na(1+n)}$ be the collection of variables $\alpha_1,\ldots,\alpha_{na}$ and $[A^s](1,h)$, with $h = 1,\ldots,na$ and $s = 1,\ldots,n$, i.e.

$$q^{(n)} = [\alpha_1,\ldots,\alpha_{na}, [A](1,1),\ldots,[A](1,na),$$
$$[A^2](1,1),\ldots,[A^2](1,na),\ldots,[A^n](1,na)]^\top. \quad (21)$$

Let Υ be the set of all those vectors with $na(1+n)$ components, each one equal to ± 1. This means that $\Upsilon =$ $\{\beta_1,\beta_2,\ldots,\beta_l,\ldots,\beta_L\}$, where $L = 2^{na(1+n)}$ and β_l is a vector with $na(1+n)$ components, each one equal to ± 1 and such that $\beta_l \neq \beta_i$ if $l \neq i$. For any $\beta_l \in \Upsilon$, let us define the set $\mathcal{O}(\beta_l) \subset \mathbb{R}^{n_\gamma + n_\alpha + n_\theta}$ as

$$\mathcal{O}(\beta_l) = \{(\gamma,\alpha,\theta) \in \mathbb{R}^{n_\gamma + n_\alpha + n_\theta} :$$
$$\beta_{lj} q_j^{(n)}(\alpha,\theta) \geq 0, \ j = 1,\ldots,na(1+n)\},$$

where β_{lj} and $q_j^{(n)}$ are the j-th element of vectors β_l and $q^{(n)}$, respectively. Note that, in all sets $\mathcal{O}(\beta_l)$, the sign of the vector $q^{(n)}$ components is imposed. Topological features of $\mathcal{D}^{ss(n)}$ are now highlighted by Property 1.

Property 1: The set $\mathcal{D}^{ss(n)}$ is the union of at most L sets $\mathcal{D}_l^{ss(n)}$ in $\mathbb{R}^{n_\gamma + n_\alpha + n_\theta + N}$, that is $\mathcal{D}^{ss(n)} = \bigcup_{l=1}^{L} \mathcal{D}_l^{ss(n)}$, where $\mathcal{D}_l^{ss(n)} = \mathcal{D}^{ss(n)} \cap \mathcal{O}(\beta_l)$.
Each set $\mathcal{D}_l^{ss(n)}$, if not empty, is a semialgebraic region in $\mathbb{R}^{n_\gamma + n_\alpha + n_\theta}$ defined by polynomial inequality constraints of degree less than or equal to $n+1$. ∎

Thanks to the structure of $\mathcal{D}^{ss(n)}$ highlighted by Property 1, problems (18)-(20) can be decomposed into the collection of the following polynomial optimization problems:

$$\underline{\gamma}_i^{ss(n)} = \min_{l=1,\ldots,L} \underline{\gamma}_{il}^{ss(n)}; \ \overline{\gamma}_i^{ss(n)} = \max_{l=1,\ldots,L} \overline{\gamma}_{il}^{ss(n)}; \quad (22)$$

$$\underline{\alpha}_j^{ss(n)} = \min_{l=1,\ldots,L} \underline{\alpha}_{jl}^{ss(n)}; \ \overline{\alpha}_j^{ss(n)} = \max_{l=1,\ldots,L} \overline{\alpha}_{jl}^{ss(n)}; \quad (23)$$

$$\underline{\theta}_k^{ss(n)} = \min_{l=1,\ldots,L} \underline{\theta}_{kl}^{ss(n)}; \ \overline{\theta}_k^{ss(n)} = \max_{l=1,\ldots,L} \overline{\theta}_{kl}^{ss(n)}; \quad (24)$$

where

$$\underline{\gamma}_{il}^{ss(n)} = \min_{\gamma,\alpha,\theta \in \mathcal{D}_l^{ss(n)}} \gamma_i; \ \overline{\gamma}_{il}^{ss(n)} = \max_{\gamma,\alpha,\theta \in \mathcal{D}_l^{ss(n)}} \gamma_i; \quad (25)$$

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

$$\underline{\theta}_{kl}^{ss(n)} = \min_{\gamma,\alpha,\theta \in \mathcal{D}_l^{ss(n)}} \theta_k; \ \overline{\theta}_{kl}^{ss(n)} = \max_{\gamma,\alpha,\theta \in \mathcal{D}_l^{ss(n)}} \theta_k. \quad (27)$$

In fact, since $\mathcal{D}^{ss(n)}$ can be expressed as the union of semialgebraic sets $\mathcal{D}_l^{ss(n)}$, solving (18)-(20) over the feasible region $\mathcal{D}^{ss(n)}$ is equivalent to computing $\underline{\gamma}_{il}^{ss(n)}$ ($\overline{\gamma}_{il}^{ss(n)}$), $\underline{\alpha}_{jl}^{ss(n)}$ ($\overline{\alpha}_{jl}^{ss(n)}$) and $\underline{\theta}_{kl}^{ss(n)}$ ($\overline{\theta}_{kl}^{ss(n)}$), respectively, over each region $\mathcal{D}_l^{ss(n)}$ for all $l = 1,\ldots,L$; and thus to computing the minimum (maximum) over all $\underline{\gamma}_{il}^{ss(n)}$ ($\overline{\gamma}_{il}^{ss(n)}$), $\underline{\alpha}_{jl}^{ss(n)}$ ($\overline{\alpha}_{jl}^{ss(n)}$) and $\underline{\theta}_{kl}^{ss(n)}$ ($\overline{\theta}_{kl}^{ss(n)}$).
Since the number of optimization variables involved in (25)-(27), which equals the number of unknown system parameters, is significantly smaller than the number of variables involved in (9)-(11), computation of relaxed solutions to such polynomial problems by means of the LMI-relaxation procedure proposed in [21] is computationally tractable. In particular, for a given relaxation order δ, application of the relaxation technique in [21] to problems (25)-(27) leads to the convex SDP problems

$$\underline{\gamma}_{il}^{ss(n,\delta)} = \min_{p \in \mathcal{D}_l^{ss(n,\delta)}} f_i(p); \ \overline{\gamma}_{il}^{ss(n,\delta)} = \max_{p \in \mathcal{D}_l^{ss(n,\delta)}} f_i(p); \quad (28)$$

$$\underline{\theta}_{kl}^{ss(n,\delta)} = \min_{p \in \mathcal{D}_l^{ss(n,\delta)}} g_j(p); \quad \overline{\theta}_{kl}^{ss(n,\delta)} = \max_{p \in \mathcal{D}_l^{ss(n,\delta)}} g_j(p); \quad (29)$$

$$\underline{\theta}_{kl}^{ss(n,\delta)} = \min_{p \in \mathcal{D}_l^{ss(n,\delta)}} h_k(p); \quad \overline{\theta}_{kl}^{ss(n,\delta)} = \max_{p \in \mathcal{D}_l^{ss(n,\delta)}} h_k(p), \quad (30)$$

where p is the decision variable vector of dimension $\begin{pmatrix} n_\gamma + n_\alpha + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$. The objective functions $f_i(p)$, $g_j(p)$ and $h_k(p)$ in (28)-(30) are linear in p and the feasible region $\mathcal{D}_l^{ss(n,\delta)}$ is a convex set defined by an LMI of size $\begin{pmatrix} n_\gamma + n_\alpha + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$ and $2n(N - na) + 2$ LMIs whose maximum size is equal to $\begin{pmatrix} n_\gamma + n_\alpha + n_\theta + \delta - 1 \\ \delta - 1 \end{pmatrix}$. The reader is referred to [21] for details on the relaxation of polynomial optimization problems through SDP optimization.

Remark 2: The minimum allowed value $\underline{\delta}$ of the LMI relaxation order, so that (28)-(30) are well-defined, is $\left\lceil \frac{\rho(\mathcal{D}_l^{ss(n)})}{2} \right\rceil$, 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 Property 1 the maximum degree of the polynomial constraints describing $\mathcal{D}_l^{ss(n)}$ is equal to $n + 1$, therefore $\underline{\delta} = \lceil \frac{n+1}{2} \rceil$. ■

V. PROPERTIES OF COMPUTED PARAMETER UNCERTAINTY INTERVALS

In this section the main features enjoyed by the computed parameter bounds (28)-(30) are discussed.

For a given dynamic horizon $n \geq 1$ and relaxation order $\delta \geq \underline{\delta}$, let us define the δ-relaxed parameter uncertainty intervals $PUI_{\gamma_i}^{ss(n,\delta)} = \left[\underline{\gamma}_i^{ss(n,\delta)}; \overline{\gamma}_i^{ss(n,\delta)}\right]$, $PUI_{\alpha_j}^{ss(n,\delta)} = \left[\underline{\alpha}_j^{ss(n,\delta)}; \overline{\alpha}_j^{ss(n,\delta)}\right]$, $PUI_{\theta_k}^{ss(n,\delta)} = \left[\underline{\theta}_k^{ss(n,\delta)}; \overline{\theta}_k^{ss(n,\delta)}\right]$, where

$$\underline{\gamma}_i^{ss(n,\delta)} = \min_{l=1,\dots,L} \underline{\gamma}_{il}^{ss(n,\delta)}; \quad \overline{\gamma}_i^{ss(n,\delta)} = \max_{l=1,\dots,L} \overline{\gamma}_{il}^{ss(n,\delta)} \quad (31)$$

$$\underline{\alpha}_j^{ss(n,\delta)} = \min_{l=1,\dots,L} \underline{\alpha}_{jl}^{ss(n,\delta)}; \quad \overline{\alpha}_j^{ss(n,\delta)} = \max_{l=1,\dots,L} \overline{\alpha}_{jl}^{ss(n,\delta)} \quad (32)$$

$$\underline{\theta}_k^{ss(n,\delta)} = \min_{l=1,\dots,L} \underline{\theta}_{kl}^{ss(n,\delta)}; \quad \overline{\theta}_k^{ss(n,\delta)} = \max_{l=1,\dots,L} \overline{\theta}_{kl}^{ss(n,\delta)} \quad (33)$$

Thus, for very dynamic horizon $n \in [1, N - na]$ and relaxation order $\delta \geq \underline{\delta} = \left\lceil \frac{n+1}{2} \right\rceil$, the intervals $PUI_{\gamma_i}^{ss(n,\delta)}$ enjoy the properties listed below. Similar results hold for $PUI_{\alpha_j}^{ss(n,\delta)}$ and $PUI_{\theta_k}^{ss(n,\delta)}$.

Property 2: **Guaranteed relaxed uncertainty intervals.** The interval $PUI_{\gamma_i}^{ss(n,\delta)}$ is guaranteed to contain the true nonlinear block parameter γ_i to be estimated, i.e. $\gamma_i \in PUI_{\gamma_i}^{ss(n,\delta)}$. ■

Property 3: **Monotone convergence to parameter uncertainty intervals** $PUI_{\gamma_i}^{ss(n)}$.
The parameter uncertainty interval $PUI_{\gamma_i}^{ss(n,\delta)}$ becomes tighter as the relaxation order δ increases, that is $PUI_{\gamma_i}^{ss(n,\delta+1)} \subseteq PUI_{\gamma_i}^{ss(n,\delta)}$. Furthermore, the computed interval $PUI_{\gamma_i}^{ss(n,\delta)}$ converges to $PUI_{\gamma_i}^{ss(n)}$ as the relaxation order δ goes to infinity. ■

It is worth remarking that, although the convergence property in Property 3 is guaranteed as the relaxation order goes to infinity, intervals $PUI_{\gamma_i}^{ss(n)}$ can be tightly approximated in practice with a reasonably low relaxation order.

Property 4: **Increasing accuracy in uncertainty intervals evaluation**
The parameter uncertainty interval $PUI_{\gamma_i}^{ss(n,\delta)}$ becomes tighter as the dynamic horizon n increases, i.e. $PUI_{\gamma_i}^{ss(n+1,\delta)} \subseteq PUI_{\gamma_i}^{ss(n,\delta)}$. ■

VI. A SIMULATED EXAMPLE

In this section a simulated example is presented in order to show the effectiveness of the proposed approach. The input nonlinearity is modeled by the polynomial function $x_t = \mathcal{N}_1(u_t) = u_t + 0.4u_t^2 - 0.1u_t^3$, the inverse of the output nonlinearity is modeled by the function $z_t = \mathcal{N}_2^{-1}(w_t) = w_t + 0.8w_t^2 + 3.2w_t^3$, while the linear part is a strictly-proper second order system with parameters $\theta^{\mathsf{T}} = [a_1 \ a_2 \ b_1 \ b_2] = [1.8 \ 0.9 \ 1.6 \ 2.1]$. The system is excited by an random input sequence u_t uniformly distributed in $[-10, +10]$. The output data sequence is corrupted by random additive noises η_t, uniformly distributed in the interval $[-\Delta\eta_t, +\Delta\eta_t]$. The chosen error bounds $\Delta\eta_t$ are such that the signal to noise ratio $SNR_w = 10\log\left\{\sum_{t=1}^{N} w_t^2 \Big/ \sum_{t=1}^{N} \eta_t^2\right\}$ is equal to 31 dB. The length of the data sequence is $N = 1000$. Bounds on the parameters are evaluated by solving problems (28)-(30) for a relaxation order $\delta = 3$. The software Gloptipoly [22] has been used to convert identification problems (25)-(27) into their corresponding LMI relaxed problems (28)-(30), which are numerically solved by the SDP solver SeDuMi.

In order to show the increasing accuracy in the uncertainty intervals evaluation as the dynamic horizon n grows, the algorithm is performed for two different values of n, i.e. $n = 2$ and $n = 3$. Results on the evaluation of system parameters are reported in Tables I, II and III, which show the obtained parameter bounds, as well as the parameter uncertainties $\Delta\gamma_i^{(n,\delta)} = \frac{\overline{\gamma}_i^{ss(n,\delta)} - \underline{\gamma}_i^{ss(n,\delta)}}{2}$, $\Delta\alpha_j^{(n,\delta)} = \frac{\overline{\alpha}_j^{ss(n,\delta)} - \underline{\alpha}_j^{ss(n,\delta)}}{2}$ and $\Delta\theta_k^{(n,\delta)} = \frac{\overline{\theta}_k^{ss(n,\delta)} - \underline{\theta}_k^{ss(n,\delta)}}{2}$. The reported results show that the true value of the parameters is always included in the computed uncertainty intervals, as stated in Property 2. Furthermore, as the dynamic horizon n grows, the accuracy in the evaluation of the parameter uncertainty intervals increases, as stated in Property 4. It is worth remarking that the presented procedure provides satisfactory parameter uncertainty intervals, both on the nonlinear blocks and on the linear one, also for small values of dynamic horizon n.

VII. CONCLUSION

A procedure for the evaluation of parameters bounds of Hammerstein-Wiener systems in the presence of bounded errors is presented in the paper. First, it is shown that system parameters bounds can be computed by solving a set of non-convex constrained polynomial optimization problems whose

TABLE I
INPUT NONLINEAR BLOCK $\mathcal{N}_1$. PARAMETER BOUNDS ($\underline{\gamma}_i^{ss(n,\delta)}$, $\overline{\gamma}_i^{ss(n,\delta)}$) AND PARAMETER UNCERTAINTIES $\Delta\gamma_i^{(n,\delta)}$ FOR RELAXATION ORDER $\delta = 3$ AND DYNAMIC HORIZON $n = 2$ AND $n = 3$.

n	Parameter	$\underline{\gamma}_i^{ss(n,\delta)}$	True value	$\overline{\gamma}_i^{ss(n,\delta)}$	$\Delta\gamma_i^{(n,\delta)}$
2	γ_2	0.273	0.4	0.547	0.137
	γ_3	-0.234	-0.1	-0.043	0.095
3	γ_2	0.313	0.4	0.476	0.082
	γ_3	-0.124	-0.1	-0.071	0.023

TABLE II
OUTPUT NONLINEAR FUNCTION $\mathcal{N}_2^{-1}(.)$. PARAMETER BOUNDS ($\underline{\alpha}_j^{ss(n,\delta)}$, $\overline{\alpha}_j^{ss(n,\delta)}$) AND PARAMETER UNCERTAINTIES $\Delta\alpha_j^{(n,\delta)}$ FOR RELAXATION ORDER $\delta = 3$ AND DYNAMIC HORIZON $n = 2$ AND $n = 3$.

n	Parameter	$\underline{\alpha}_j^{ss(n,\delta)}$	True value	$\overline{\alpha}_j^{ss(n,\delta)}$	$\Delta\alpha_j^{(n,\delta)}$
2	α_2	0.523	0.8	1.147	0.312
	α_3	2.432	3.2	3.839	0.704
3	α_2	0.664	0.8	0.931	0.133
	α_3	2.914	3.2	3.521	0.304

TABLE III
LINEAR BLOCK $\mathcal{L}$. PARAMETER BOUNDS ($\underline{\theta}_k^{ss(n,\delta)}$, $\overline{\theta}_k^{ss(n,\delta)}$) AND PARAMETER UNCERTAINTIES $\Delta\theta_k^{(n,\delta)}$ FOR RELAXATION ORDER $\delta = 3$ AND DYNAMIC HORIZON $n = 2$ AND $n = 3$.

n	Parameter	$\underline{\theta}_k^{ss(n,\delta)}$	True value	$\overline{\theta}_k^{ss(n,\delta)}$	$\Delta\theta_k^{(n,\delta)}$
2	a_1	1.556	1.8	1.999	0.221
	a_2	0.695	0.9	0.999	0.152
	b_1	1.138	1.6	2.194	0.528
	b_2	1.572	2.1	2.718	0.595
3	a_1	1.612	1.8	1.934	0.161
	a_2	0.785	0.9	0.968	0.092
	b_1	1.345	1.6	1.877	0.266
	b_2	1.875	2.1	2.331	0.228

size increases with the number of measurements. Due to the large number of decision variables and the large order of the polynomial constraints involved in the optimization problems, the LMI-based relaxation techniques for polynomial problems available in literature lead to SDP problems which are practically untractable. A relaxation procedure is then proposed to significantly reduce the computational burden of the proposed identification algorithm. The key idea of the proposed approach is to consider a suitable outer bound of the feasible parameter set whose description involves a smaller number of decision variables in order to balance the tradeoff between conservativeness and computational burden. Such an outer bound is the union of a finite number of semialgebraic sets in the space of the unknown system parameters to be estimated. As a consequence, parameters bounds can be evaluated by solving suitable polynomial optimization problems involving a smaller number of variables, i.e. only the unknown parameters of the system. Then, LMI relaxation techniques are used to approximate global optima of such small size problems. The computed parameter uncertainty

intervals are guaranteed to contain the true Hammerstein-Wiener system parameters to be estimated. The simulated example shows that the presented identification can handle a large number of measurements, providing satisfactory parameters bounds.

REFERENCES

[1] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Deylon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: a unified overview," *Automatica*, vol. 31, pp. 1691–1724, Dec. 1995.

[2] L. Ljung and A. Vicino(Ed.), "Special issue on system identification," *IEEE Trans. Automat. Contr*, vol. 50, no. 10, 2005.

[3] L. Ljung, "Some aspects on nonlinear system identification," in *Proc 14th IFAC Symposium on System Identification, (Plenary session)*, Newcastle, Australia, Mar. 2006.

[4] ——, "Perspectives on system identification," in *Proc. 17th IFAC World Congress, (Plenary session)*, Seoul, Korea, July 2008, (Plenary session).

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

[6] E. Bai and F. Giri, *Block-oriented nonlinear system identification*, ser. Lecture notes in Control and Information sciences. Berlin: Springer, 2010.

[7] E. Bai, "An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, no. 3, pp. 333–338, 1998.

[8] ——, "A blind approach to the Hammerstein-Wiener model identification," *Automatica*, vol. 38, pp. 967–979, 2002.

[9] D. Bauer and B. Ninness, "Asymptotic properties of least squares estimates of Hammerstein-Wiener model structure," *International Journal of Control*, vol. 75, no. 1, pp. 34–51, jan 2002.

[10] Y. Zhu, "Estimation of an N-L-L hammerstein-wiener model," *Automatica*, vol. 38, pp. 1697–1614, 2002.

[11] P. Crama and J. Schoukens, "Hammerstein-Wiener," *IEEE Trans. Automat. Contr*, vol. AC-16, pp. 464–468, 1971.

[12] D. Wang and F. Ding, "Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems," *Computers and Maths with applications*, vol. 56, pp. 3157–3164, 2008.

[13] I. Goethals, K. Pelckmans, L. Hoegaerts, J. Suykens, and B. D. Moor, "Subspace intersection identification of Hammerstein-Wiener systems," in *Proc. of 44th IEEE Conf. on Decision and Control*, 2005.

[14] J. van Wirgerden and M. Verhaegen, "Closed-loop subspace identification of Hammerstein-Wiener models," in *Proc. of 48th IEEE Conference on Decision and Control*, 2009, pp. 3637–3642.

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

[16] 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.

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

[18] J. B. Lasserre, "Convergent semidefinite relaxations in polynomial optimization with sparsity," *SIAM J. Optimiz*, vol. 17, no. 1, pp. 822–843, 2006.

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

[20] V. Cerone, D. Piga, and D. Regruto, "Set-membership identification of hammerstein-wiener models," *Internal Report*, vol. DAUIN, no. 2011/03/21CPR, 2011.

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

[22] D. Henrion and J. B. Lasserre, "Gloptipoly: Global optimization over polynomials with Matlab and SeDuMi," *ACM Transactions Math. Soft*, vol. 29, pp. 165–194, 2003.