Computational burden reduction in Set-membership Hammerstein system identification

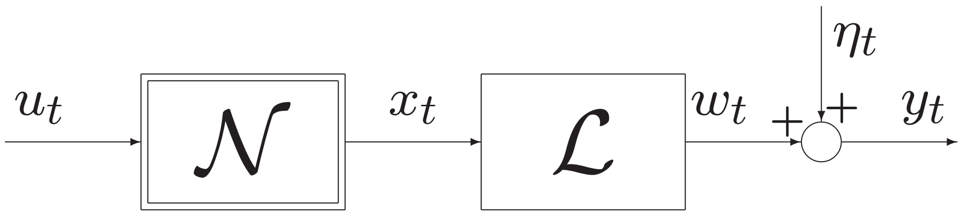
V. Cerone, D. Piga, D. Regruto
Dipartimento di Automatica e Informatica
Politecnico di Torino (Italy)

e-mail: vito.cerone@polito.it, dario.piga@polito.it, diego.regruto@polito.it

IFAC World Congress 2011
Session FrA23: "Identification of Hammerstein Models"
Milan, Italy — September 2, 2011

Hammerstein systems

Let us consider a SISO Hammerstein system:



u_t: known input signal

y_t: noise-corrupted measurement of w_t

x_t: not measurable inner signals

$\mathcal{N}$: nonlinear static block

$\mathcal{L}$: linear dynamic subsystem

$$\mathcal{N} : x_t = \sum_{k=1}^{n_\gamma} \gamma_k \psi_k(u_t); \quad \psi_k(\cdot) \text{ known static nonlinear functions}$$

$$\mathcal{L} : w_t = -\sum_{i=1}^{na} a_i w_{t-i} + \sum_{j=0}^{nb} b_j x_{t-j}$$

$$|\eta_t| \leq \Delta\eta_t; \quad \Delta\eta_t \text{ known (Set-Membership characterization)}$$

Identification of Hammerstein systems

- **Aim of the work:** compute bounds on the nonlinear block parameters $\gamma^\top = [\gamma_1\ \gamma_2 \ldots \gamma_{n_\gamma}]$ and linear block parameters $\theta^\top = [a_1 \ldots a_{na}\ b_0\ b_1 \ldots b_{nb}]$.

- Parameter bound computation of Hammerstein systems is NP-hard in the size of the experimental data sequence **(M. Sznaier, *Automatica*, 2009)**

$$\Downarrow$$

Computationally tractable relaxations are needed

Feasible parameter set (FPS)

- In bounded-error (or set-membership) context, all the system parameters γ and θ consistent with the measurement data sequence, the assumed model structure and the error bounds are feasible solution to the identification problem (and are said to belong to the feasible parameter set $\mathcal{D}_{\gamma\theta}$).

- $\mathcal{D}_{\gamma\theta}$ is the projection on the parameter space of the set $\mathcal{D}$ of all the Hammerstein system parameters γ-θ and the noise samples and η_t consistent with the measurement data sequence, the assumed model structure and the error bounds, given by:

$$\mathcal{D} = \Big\{(\gamma,\theta,\eta) \in \mathbb{R}^{n_\gamma+n_\theta+N} : y_t = -\sum_{i=1}^{na} a_i(y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb}\sum_{k=1}^{n_\gamma} b_j\gamma_k\psi_k(u_{t-j}) + \eta_t,$$

$$|\eta_r| \leq \Delta\eta_r, \ \ \sum_{i=1}^{na} a_i = 1 + \sum_{j=0}^{nb} b_j, t = na+1,\ldots,N;\ r = 1,\ldots,N\Big\}$$

Computation of parameter bounds: preliminaries

- Exact parameter bounds:

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

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

- Parameter Uncertainty Intervals:

$$PUI_{\gamma_k} = \left[\underline{\gamma}_k;\ \overline{\gamma}_k\right] \quad PUI_{\theta_j} = \left[\underline{\theta}_j;\ \overline{\theta}_j\right]$$

Key point 1: Both the system parameters γ-θ and the noise samples η are decision variables in the above optimization problem $\Rightarrow$ The number of optimization variables increases with the number of measurements

Key point 2: $\mathcal{D}$ is a nonconvex set described by polynomial constraints $\Rightarrow$ exact bound computation requires to solve a set of nonconvex optimization problems

Computation of parameter bounds: preliminaries

- Standard nonlinear optimization tools can not be exploited to compute bounds on γ_k (resp. θ_j) since they can trap in local minima

$$\Downarrow$$

The true value is not guaranteed to lie within the computed bounds

- Relax original identification problems to convex optimization problems

$$\Downarrow$$

Guaranteed (relaxed) bounds on each parameter γ_k (resp. θ_j) can be evaluated

Computation of relaxed bounds

LMI-based relaxation approach:

- *General Idea*
  Exploit LMI-relaxation for semialgebraic optimization problems
  SOS decomposition (**P. Parrillo,** *Mathematical Programming 2003*)
  Theory of moments (**J. B. Lasserre,** *SIAM J. on Opt. 2001*)

- *Computational complexity*
  Due to the large number of optimization variables and constraints involved in the identification problems, such LMI relaxation techniques leads, in general, to untractable SDP problems.

Computation of relaxed bounds

LMI-based relaxation approach: exploiting sparsity (**Cerone, Piga and Regruto,** *ACC, 2011*)

$$\mathcal{D} = \Big\{(\gamma,\theta,\eta) \in \mathbb{R}^{n_\gamma+n_\theta+N} : y_t = -\sum_{i=1}^{na} a_i(y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb}\sum_{k=1}^{n_\gamma} b_j\gamma_k\psi_k(u_{t-j}) + \eta_t,$$

$$|\eta_r| \leq \Delta\eta_r, \ \sum_{i=1}^{na} a_i = 1 + \sum_{j=0}^{nb} b_j, t = na+1,\ldots,N;\ r = 1,\ldots,N\Big\},$$

- The constraints defining $\mathcal{D}$ depend only on a small subset of variables.
- The peculiar *sparsity structure* enjoyed by the original identification problems is used to reduce the computational complexity of the relaxed SDP problems
- The number of decision variables of the relaxed SDP problems, as well as the number of the constraints defining the LMI-feasibility set, linearly increases with N
- Identification problem with about 800 measurements can be handled

How to deal with identification with more measurements?

Semi-static approach

Overview

- Construct an outer approximation $\mathcal{D}^{(n)}$ of $\mathcal{D}$ by partly neglecting the correlation among consecutive equations defining $\mathcal{D}$
- Correlation between consecutive equations is kept to some extent (thus reducing the conservativeness of $\mathcal{D}^{(n)}$ in approximating $\mathcal{D}$)
- Constraints defining the outer-approximating set $\mathcal{D}^{(n)}$ do not depend on the noise samples η, but only on the system parameters γ and θ
- Eventually, parameter bounds are computed by solving polynomial optimization problems where only the system parameters are treated are optimization variables.

Semi-static approach

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

Alternative description of set $\mathcal{D}$

$$\mathcal{D} = \bigcap_{z=1}^{N-na} \mathcal{S}_z^{(n)}$$

$\mathcal{S}_z^{(n)}$ is defined by the constraints arising from n consecutive measurements

- $\mathcal{S}_1^{(n)}$: defined by the constraints from $t = na + 1$ up to $t = na + n$
- $\mathcal{S}_2^{(n)}$: defined by the constraints from $t = na + 2$ up to $t = na + 1 + n$

⋮

- $\mathcal{S}_{N-na}^{(n)}$: defined by the constraint at $t = N$

Semi-static approach

Alternative description of the set $S_z^{(n)}$

Substitute the first constraint defining $S_z^{(n)}$ in the second one, leading to a new equality; then, the new obtained equation is substituted in the third constraint. The procedure is repeated until all the newly obtained equations are substituted into the n-th considered constraint.

$$S_z^{(n)} = \Big\{ (\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : y_t = -\sum_{i=1}^{na} a_i (y_{t-i} - \eta_{t-i}) + \sum_{j=0}^{nb} \sum_{k=1}^{n_\gamma} b_j \gamma_k \psi_k (u_{t-j}) + \eta_t,$$
$$|\eta_r| \leq \Delta\eta_r,\ t = na + z, na + z + 1, \ldots, \min\{na + z + n, N\};\ r = z, z + 1, \ldots, \min\{na + z + n, N\} \Big\}$$

$$\Downarrow$$

$$\mathcal{S}_z^{(n)} = \Big\{ (\gamma, \theta, \eta) \in \mathbb{R}^{n_\gamma + n_\theta + N} : y_{na+z+s-1} = \sum_{i=1}^{na} a_i^{(s)} (y_{na+z-i} - \eta_{na+z-i}) + \sum_{j=1}^{s} \left\langle \underline{g}_j^{(s)}, \underline{\varphi}_{na+z+j-1} \right\rangle + \eta_{na+z+s-1},$$
$$|\eta_r| \leq \Delta\eta_r,\ s = 1, 2, \ldots, \min\{n, N - z + 1\}; r = z, z + 1, \ldots, \min\{na + n + z, N\} \Big\}$$

Remark: Coefficients $a_i^{(s)}$ and $\underline{g}_j^{(s)}$ are polynomial functions of parameters γ and θ.

Alternative description of the set $S_z^{(n)}$: Example

$$\mathcal{N}: x_t = \gamma\psi_1(u_t) \qquad \mathcal{L}: w_t = -a_1 w_{t-1} + b_1 x_{t-1} \quad \Longrightarrow \quad y_t = -a_1(y_{t-1} - \eta_{t-1}) + b_1\gamma\psi_1(u_{t-1}) + \eta_t$$

$$\mathcal{D} = \Big\{(\gamma,\theta,\eta) \in \mathbb{R}^{3+N} : y_t = -a_1(y_{t-1} - \eta_{t-1}) + b_1\gamma\psi_1(u_{t-1}) + \eta_t, \\ |\eta_r| \leq \Delta\eta_r,\ \ a_1 = 1 + b_1,\ t = 2,\dots,N;\ r = 1,\dots,N\Big\},$$

$$\mathcal{S}_1^{(n=3)} = \{(\gamma,\theta,\eta) \in \mathbb{R}^{3+N} : \ y_2 = -a_1(y_1 - \eta_1) + b_1\gamma\psi(u_1) + \eta_2; \\ y_3 = -a_1(y_2 - \eta_2) + b_1\gamma\psi(u_2) + \eta_3; \\ y_4 = -a_1(y_3 - \eta_3) + b_1\gamma\psi(u_3) + \eta_4; \\ |\eta_r| \leq \Delta\eta_r,\ r = 1,\dots 4\}$$

$$\Downarrow$$

$$\mathcal{S}_1^{(n=3)} = \{(\gamma,\theta,\eta) \in \mathbb{R}^{3+N} : \ y_2 = -a_1(y_1 - \eta_1) + b_1\gamma\psi(u_1) + \eta_2; \\ y_3 = a_1^2(y_1 - \eta_1) - a_1 b_1\gamma\psi(u_1) + b_1\gamma\psi(u_2) + \eta_3; \\ y_4 = -a_1^3(y_1 - \eta_1) + a_1^2 b_1\gamma\psi(u_1) - a_1 b_1\gamma\psi(u_2) + b_1\gamma\psi(u_3) + \eta_4; \\ |\eta_r| \leq \Delta\eta_r,\ r = 1,\dots 4\}$$

Semi-static approach: Outer approximation of $S_z^{(n)}$

- **General idea**: Consider the occurrences of the noise samples η_t appearing in the new definition of $S_z^{(n)}$ independent with each other.

- Outer approximation $S_z^{ss(n)}$ of $S_z^{(n)}$, given by:

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

Remark 1: Constraints defining $\mathcal{S}_z^{ss(n)}$ do not depend on the noise samples η, but only on the Hammerstein system parameters γ and θ.

Remark 2: Because of the nested substitutions in the alternative definition of $S_z^{(n)}$, correlation between consecutive regressors in the definition of $S_z^{(n)}$ is not completely lost.

Semi-static approach: Outer approximation of $\mathcal{D}$

- $\mathcal{D} = \bigcap_{z=1}^{N-na} \mathcal{S}_z^{(n)}$
- $S_z^{(n)} \subseteq S_z^{ss(n)}$

$$\Rightarrow \quad \bigcap_{z=1}^{N-na} \mathcal{S}_z^{ss(n)} = \mathcal{D}^{(n)} \supseteq \mathcal{D}$$

- Relaxed parameter bounds:

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

$$\underline{\theta}_j^{(n)} = \min_{(\gamma,\theta)\in\mathcal{D}^{(n)}} \theta_j, \quad \overline{\theta}_j^{(n)} = \max_{(\gamma,\theta)\in\mathcal{D}^{(n)}} \theta_j$$

- **Property 1**: $\mathcal{D}^{(n)}$ is the union of at most $L = 2^{n\cdot na}$ semialgebraic sets $\mathcal{D}_l^{(n)}$ in the parameter space, i.e. $\mathcal{D}^{(n)} = \bigcup_{l=1}^{L} \mathcal{D}_l^{(n)}$.
- **Property 2**: Constraints describing $\mathcal{D}_l^{(n)}$ are polynomial inequalities of degree at most $n+1$.
- **Property 3**: $\mathcal{D}^{(n)}$ becomes tighter as n grows, i.e. $\mathcal{D}^{(n+1)} \subseteq \mathcal{D}^{(n)}$.
- **Property 4**: If $n = N - na$, $\mathcal{D}^{(n)} = \mathcal{D}$.
- **Property 5**: The decision variables of the above optimization problems are only parameters γ and θ.

Semi-static approach: LMI-relaxation

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

$$\underline{\theta}_j^{(n)} = \min_{(\gamma,\theta)\in\mathcal{D}^{(n)}} \theta_j, \quad \overline{\theta}_j^{(n)} = \max_{(\gamma,\theta)\in\mathcal{D}^{(n)}} \theta_j$$

- The above optimization problems are relaxed into a hierarchy of SDP programming problems through theory-of-moment techniques

- Such SDP problems are such that:
  - the number of decision variable is $\begin{pmatrix} n_\gamma + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$.
  - the region of feasibility is a convex set defined by an LMI of size $\begin{pmatrix} n_\gamma + n_\theta + 2\delta \\ 2\delta \end{pmatrix}$ and $2n(N - na) + 1$ LMIs whose maximum size is equal to $\begin{pmatrix} n_\gamma + n_\theta + \delta - 1 \\ \delta - 1 \end{pmatrix}$.

Example 1

Simulated Hammerstein system

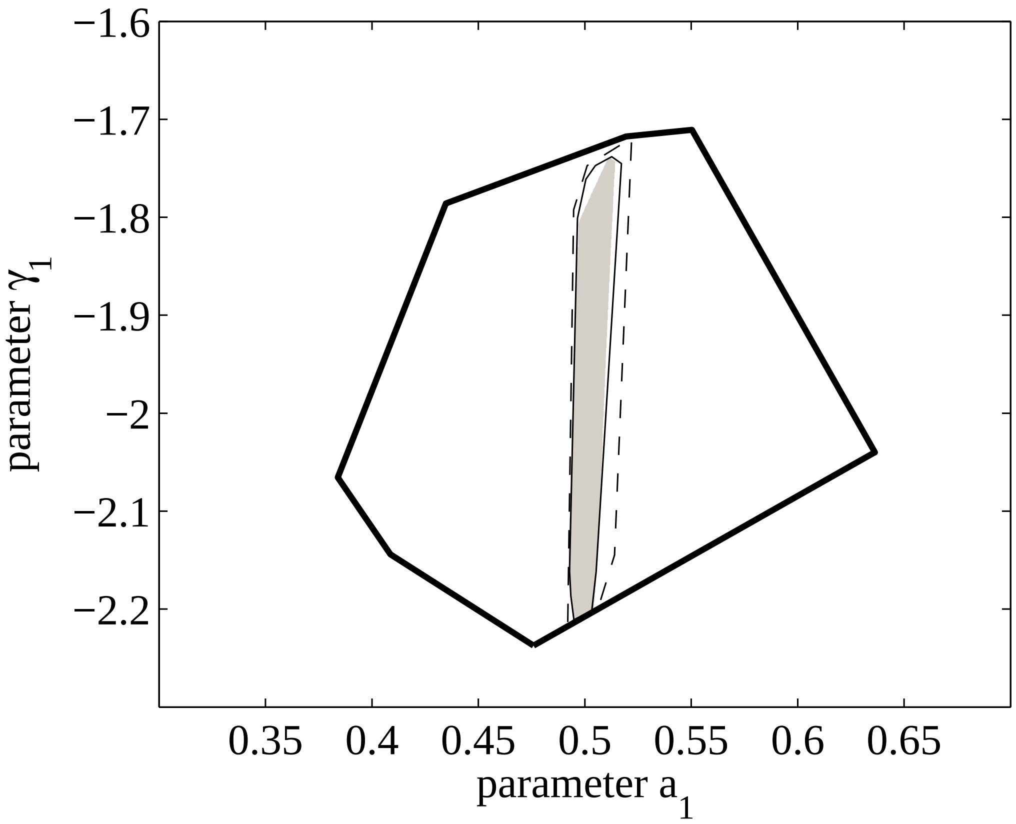
- $\mathcal{N}$: $x_t = \gamma_1 \arctan(u_t)$, with $\gamma_1 = -2$.
- $\mathcal{L}$: first order model with parameters $a_1 = 0.5$.

Measurements errors

- w_t is corrupted by random additive noise, uniformly distributed between $[-\Delta\eta_t, +\Delta\eta_t]$
- error bounds $\Delta\eta_t$ are such that $SNR_w = 10 \log \left\{ \sum_{t=1}^{N} w_t^2 \Big/ \sum_{t=1}^{N} \eta_t^2 \right\} = 21$ db.

Length of measurement data sequence: $N = 50$

Exact feasible parameter set (grey region) and approximate feasible parameter set $\mathcal{D}^{ss(n,\delta)}$ obtained for dynamic horizon $n = 1$ (region inside solid thick line), $n = 2$ (region inside dashed line) and $n = 3$ (region inside solid thin line).



Example 2

Simulated Hammerstein system

- $\mathcal{N}$: $x_t = \gamma_1 u_t + \gamma_2 u_t^2 + \gamma_3 u_t^3$, with $[\gamma_1\ \gamma_2\ \gamma_3] = [0.3;\ 0.4;\ -0.6]$
- $\mathcal{L}$: second order model with parameters $[a_1\ a_2\ b_1\ b_2] = [0.65;\ 0.73;\ 1.41;\ 0.97]$.

The **input is a random sequence** uniformly distributed between $[-1,\ +1]$.

Measurements errors

- w_t is corrupted by random additive noise, uniformly distributed between $[-\Delta\eta_t,\ +\Delta\eta_t]$
- error bounds $\Delta\eta_t$ are such that $SNR_w = 10 \log \left\{ \sum_{t=1}^{N} w_t^2 \Big/ \sum_{t=1}^{N} \eta_t^2 \right\} = 25$ db.

Length of measurement data sequence: $N = 4000$

Nonlinear block. Parameter bounds ($\underline{\gamma}_k^{ss(n,\delta)}$, $\overline{\gamma}_k^{ss(n,\delta)}$), central estimates $\gamma_k^{c(n,\delta)}$, and parameter uncertainties $\Delta\gamma_k^{(n,\delta)}$ for dynamic horizon $n = 1$, $n = 2$ and $n = 3$.

n	Parameter	True value	$\underline{\gamma}_k^{ss(n,\delta)}$	$\gamma_k^{c(n,\delta)}$	$\overline{\gamma}_k^{ss(n,\delta)}$	$\Delta\gamma_k^{(n,\delta)}$
1	γ_1	0.300	0.253	0.298	0.343	0.045
	γ_2	0.400	0.342	0.397	0.453	0.056
	γ_3	-0.600	-0.838	-0.593	-0.349	0.244
2	γ_1	0.300	0.272	0.302	0.332	0.030
	γ_2	0.400	0.369	0.408	0.447	0.039
	γ_3	-0.600	-0.782	-0.622	-0.461	0.160
3	γ_1	0.300	0.295	0.306	0.316	0.011
	γ_2	0.400	0.381	0.403	0.425	0.022
	γ_3	-0.600	-0.720	-0.659	-0.598	0.061

Linear block. Parameter bounds ($\underline{\theta}_j^{ss(n,\delta)}$, $\overline{\theta}_j^{ss(n,\delta)}$), central estimates $\theta_j^{c(n,\delta)}$, and parameter uncertainties $\Delta\theta_j^{(n,\delta)}$ for dynamic horizon $n = 1$, $n = 2$ and $n = 3$

n	Parameter	True value	$\underline{\theta}_j^{ss(n,\delta)}$	$\theta_j^{c(n,\delta)}$	$\overline{\theta}_j^{ss(n,\delta)}$	$\Delta\theta_j^{(n,\delta)}$
1	a_1	0.650	0.499	0.665	0.831	0.166
	a_2	0.730	0.568	0.749	0.929	0.181
	b_1	1.410	1.238	1.433	1.628	0.195
	b_2	0.970	0.839	1.001	1.163	0.162
2	a_1	0.650	0.594	0.648	0.701	0.054
	a_2	0.730	0.575	0.743	0.911	0.168
	b_1	1.410	1.306	1.416	1.526	0.110
	b_2	0.970	0.893	0.978	1.063	0.085
3	a_1	0.650	0.644	0.664	0.683	0.019
	a_2	0.730	0.728	0.756	0.783	0.028
	b_1	1.410	1.435	1.454	1.474	0.020
	b_2	0.970	0.945	0.968	0.990	0.023

Conclusion

- A single-stage procedure for Set-membership identification of Hammerstein systems has been presented
- Parameter bound computation is formulated in terms of sparse polynomial optimization problems, with a number of variables that increases with the number of measurements
- Because of their high computational complexity, the LMI-based relaxation approach can be exploited for only small/medium number of experimental data.
- In order to reduce the computational complexity of the identification problems, an outer bound of the feasible parameter set has been sought. Such an outer-bound is the union of a finite number of semialgebraic sets over the parameter space.
- Parameter bounds can be computed by solving suitable polynomial optimization problems involving only the unknown parameters of the system as variables.
- Guaranteed uncertainty intervals are computed by means of LMI relaxation techniques.