Characterization of bounded uncertainties

Olivier Adrot, José Ragot, Jean-Marie Flaus

CHARACTERIZATION OF BOUNDED UNCERTAINTIES

O. Adrot*, J. Ragot♦, J-M. Flaus*

* Laboratoire d'Automatique de Grenoble - UMR 5528
ENSIEG, BP 46, 38402 St Martin d'Hères Cedex, France
E-mail: {Olivier.Adrot, Jean-Marie.Flaus}@lag.ensieg.inpg.fr

♦ Centre de Recherche en Automatique de Nancy - UMR 7039
ENSEM, 2, avenue de la Forêt de Haye, 54516 Vandœuvre-lès-Nancy Cedex, France
E-mail: Jose.Ragot@ensem.inpl-nancy.fr

Abstract. Parameter estimation based on set-membership approach is a non-probabilistic method for characterizing the uncertainty with which each model parameter is known. Only a class of uncertain static models containing several input-output relations is considered in this paper. Every equation error is bounded while parameters fluctuate inside a time-invariant domain represented by a zonotope. The proposed method helps to find the characteristics of this domain by taking the couplings related to bounded variables common to model output equations into account.

Keywords. Uncertain model, interval, estimation theory.

1 Introduction

This paper focuses on a set-membership parameter estimation computing the bounds of the uncertain variables occurring in a model, such that this one explains a given data set. This model is affected by uncertainties on both the additive equation error and multiplicative model parameters. The equation error is assumed to belong to an orthotope while parameters fluctuate inside a time-invariant bounded domain. The objective of this paper is to characterize this domain and to extend the method detailed in [1] to models containing several input-output equations. In this way, the study of all couplings between model equations due to bounded variables (called common variables) appearing in several relations at once is justified in the next [2], [3].

This paper is organized as follows. The section 2 reminds the context of this work and the principle of a characterization procedure. The problem formulation concerning a model composed of several input-output relations is detailed in section 3. A solution based on the generation of strip constraints due to the elimination of some common variables is presented. At last, an example illustrates the proposed method in section 4.

2 Principle

The set-membership parameter estimation started in the eighties and was originally designed to deal with a discrete-time model characterized by an unknown but bounded equation error. To summarize, this problem amounts to the determination of the set of parameter values called Feasible Parameter Set (F.P.S.). In others words, each point of this domain explains all the available observations which are consistent with data, bounds of the equation error and the model structure. Models linear in uncertain parameters lead to a theoretic F.P.S. in the form of a convex polytope. At first, the objective consists in circumscribing this domain by a simpler form as an ellipsoid [4] or an axis-aligned orthotope [5]. Later, the objective consists in exactly determining the F.P.S. by working on polytopes [6], [7], [8]. The next step takes bounded noises on model outputs [9] and on both sensor observations and equation error [10] into account. In case the model is dynamic, the F.P.S. is no more a convex polytope and only an approximation can be determined [11]. Moreover, the set-membership parameter estimation can be viewed as a set-membership inversion problem [12]. A paving method is used in order to compute an overestimation of the F.P.S. Some of these works are put together in [13].

As explained later, the problem considered in this paper is different. Time-variant uncertain parameters are defined by random variables with bounded realizations and the objective is to characterize the time-invariant domain in which they fluctuate. In fact, the proposed method is a non-probabilistic technique for determining the inaccuracy with which each model parameter is known. At first, if $\boldsymbol{\theta}$ is a bounded vector, then $\mathcal{S}(\boldsymbol{\theta})$ designs the value set of $\boldsymbol{\theta}$ corresponding to the set of all admissible values of $\boldsymbol{\theta}$.

Only structured models linear in uncertain parameters and measurements (topped with the tilde symbol) are considered. The term "structured" indicates that uncertain parameters are localized in mathematical model equations. Moreover, at the time k, these ones are represented by the bounded vector $\boldsymbol{\theta}_k$. By using analogous notations to those detailed in [1], the output vector $\boldsymbol{y}_k$ obtained at the time k, $\forall k \in \{1,\ldots,h\}$, is given as:

$$\begin{array}{ll} \boldsymbol{y}_k = \tilde{\boldsymbol{X}}_k \boldsymbol{\theta}_k + \boldsymbol{\varepsilon}_k & \boldsymbol{y}_k \in \mathbb{R}^m,\ y_{k,i} \in \mathbb{R} \\ \begin{bmatrix} \vdots \\ y_{k,i} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \tilde{\boldsymbol{x}}_{k,i}^T \\ \vdots \end{bmatrix} \boldsymbol{\theta}_k + \begin{bmatrix} \vdots \\ \varepsilon_{k,i} \\ \vdots \end{bmatrix}, & \begin{array}{l} \tilde{\boldsymbol{X}}_k \in \mathbb{R}^{m\times p},\ \tilde{\boldsymbol{x}}_{k,i} \in \mathbb{R}^p \\ \boldsymbol{\theta}_k \in \mathbb{R}^p, \\ \boldsymbol{\varepsilon}_k \in \mathbb{R}^m,\ \varepsilon_{k,i} \in \mathbb{R} \end{array} \end{array} \quad (1)$$

The scalar $\varepsilon_{k,i}$ defines the i^{th} independent equation error and corresponds to a bounded variable comprised between $-\delta_{\varepsilon,i}$ and $\delta_{\varepsilon,i}$, where $\delta_{\varepsilon,i} \in \mathbb{R}^+$ is assumed to be chosen by users for instance. Thus, the value set $\mathcal{S}(\boldsymbol{\varepsilon})$ of the vector $\boldsymbol{\varepsilon}_k$ is an axis-aligned orthotope. The term $\tilde{x}_{k,i}$ represents the regression vector (composed of measurements) associated with the i^{th} model output $y_{k,i}$. Then, let us consider a data set $\tilde{\boldsymbol{y}}_k$, $k \in \{1,\ldots,h\}$. The aim is to determine the characteristics of a domain $\mathcal{S}(\theta)$ (center, range) such that $\tilde{y}_k = \tilde{X}_k\theta_k + \varepsilon_k$ with $\theta_k \in \mathcal{S}(\theta)$, $\varepsilon_k \in \mathcal{S}(\varepsilon)$. In others words, the problem is to characterize $\mathcal{S}(\theta)$ such that at each time k:

$\tilde{y}_k \in \mathcal{S}\left(\tilde{X}_k\theta_k + \varepsilon_k\right)$ with $\theta_k \in \mathcal{S}(\theta)$, $\varepsilon_k \in \mathcal{S}(\varepsilon)$ (figure 1).

In order to simplify explanations, only the i^{th} equation of the model (1) is considered in the rest of this section.

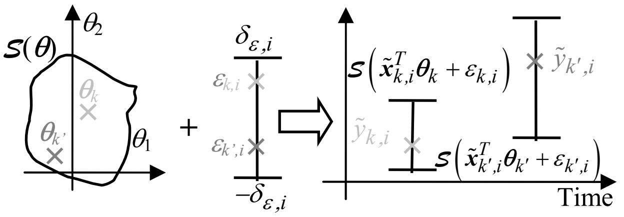

Fig 1. Principle of a characterization procedure

It leads to both following inequalities:

$$\begin{cases} \tilde{y}_{k,i} - \delta_{\varepsilon,i} \le \tilde{x}_{k,i}^T\theta_k \\ \tilde{x}_{k,i}^T\theta_k \le \tilde{y}_{k,i} + \delta_{\varepsilon,i} \end{cases}, \ \forall k \in \{1,\ldots,h\}. \tag{2}$$

This system generates at each time k a pair of half-spaces $\underline{\mathcal{D}}_k$ and $\overline{\mathcal{D}}_k$ whose frontiers define two parallel hyperplanes in the parameter space:

$$\underline{\mathcal{D}}_k = \left\{\theta' \in \mathcal{D}_0 \,/\, \tilde{y}_{k,i} - \delta_{\varepsilon,i} \le \tilde{x}_{k,i}^T\theta'\right\}, \overline{\mathcal{D}}_k = \left\{\theta'' \in \mathcal{D}_0 \,/\, \tilde{x}_{k,i}^T\theta'' \le \tilde{y}_{k,i} + \delta_{\varepsilon,i}\right\}$$

where $\mathcal{D}_0$ is the domain of investigation (taken as a simple orthotope). Since $\underline{\mathcal{D}}_k \cap \overline{\mathcal{D}}_k \ne \varnothing$ at each instant k by construction, then it must always exist at least one element θ_k of $\mathcal{S}(\theta)$ satisfying the pair of inequalities (2). The intersection of several half-spaces being convex, both following domains are convex too:

$$\underline{\mathcal{D}}_h = \bigcap_{k=1}^{h} \underline{\mathcal{D}}_k, \ \overline{\mathcal{D}}_h = \bigcap_{k=1}^{h} \overline{\mathcal{D}}_k.$$

If $\mathcal{S}(\theta)$ exists, the domain $\mathcal{S}(\theta) \cap \underline{\mathcal{D}}_h$ ($\mathcal{S}(\theta) \cap \overline{\mathcal{D}}_h$) defines the value set of θ' (θ'') leading to a major $\overline{y}_{k,i}$ (minor $\underline{y}_{k,i}$) of the measurement $\tilde{y}_{k,i}$ at each instant k:

$$\begin{matrix} \overline{y}_{k,i} = \tilde{x}_{k,i}^T\theta' + \delta_{\varepsilon,i} \\ \underline{y}_{k,i} = \tilde{x}_{k,i}^T\theta'' - \delta_{\varepsilon,i} \end{matrix} \text{ with } \begin{matrix} \overline{y}_{k,i} \ge \tilde{y}_{k,i} \\ \underline{y}_{k,i} \le \tilde{y}_{k,i} \end{matrix}, \ \forall k \in \{1,\ldots,h\}.$$

In this way, all constraints (2) are satisfied if:

$$\mathcal{S}(\theta) \cap \underline{\mathcal{D}}_h \ne \varnothing \text{ and } \mathcal{S}(\theta) \cap \overline{\mathcal{D}}_h \ne \varnothing. \tag{3}$$

Now, let us assume that a nominal (invariant) value θ_c of θ_k is known. Otherwise, θ_c can be obtained by using an estimator minimizing some α-norm of the equation error:

$$\theta_c = \underset{\theta_c}{arg\,min}\left(\sum_{k=1}^{h}\left\|\tilde{y}_{k,i} - \tilde{x}_{k,i}^T\theta_c\right\|_\alpha\right).$$

Then, let us consider the error $\varepsilon_{inf,i}$ in the sense of the infinite norm:

$$\varepsilon_{inf,i}(\theta_c) = \max_{k \in \{1,\ldots,h\}}\left(\left|\tilde{y}_{k,i} - \tilde{x}_{k,i}^T\theta_c\right|\right).$$

If $\delta_{\varepsilon,i} > \varepsilon_{inf,i}(\theta_c)$, then at least one value of the parameter vector (for example θ_c) satisfies all the $2h$ inequalities (2). This case corresponds to the determination of the Feasible Parameter Set noted $\mathcal{D}_{fps}$. In this way, $\underline{\mathcal{D}}_h \cap \overline{\mathcal{D}}_h \ne \varnothing$ and this intersection gives all invariant parameter vectors θ, each of them explaining all available observations $\tilde{y}_{k,i}$, $k \in \{1,\ldots,h\}$ (figure 2).

If $\delta_{\varepsilon,i} < \varepsilon_{inf,i}(\theta_c)$, no constant solution exists since $\underline{\mathcal{D}}_h$ and $\overline{\mathcal{D}}_h$ have no common point (the $2h$ constraints (2) cannot be simultaneously verified). In this case, the method proposed in [1] consists in finding a time-invariant convex zonotope $\mathcal{S}(\theta)$ satisfying (3), in which θ_k fluctuates such that all the $2h$ constraints (2) hold. The idea is to add to the nominal value θ_c, in fact the center of $\mathcal{S}(\theta)$, some time-variant uncertainties explaining the observation $\tilde{y}_{k,i}$ (figure 2). The zonotope $\mathcal{S}(\theta)$ is defined by:

$$\theta_k = \theta_c + \lambda T_0 \upsilon_k,\ \upsilon_k \in \mathbb{R}^q,\ T_0 \in \mathbb{R}^{p\times q},\ \lambda \in \mathbb{R}^+,\ \left\|\upsilon_k\right\|_\infty \le 1. \tag{4}$$

The normalized bounded vector υ_k, which represents mutually independent bounded variables, is assumed to be independent from the equation error $\varepsilon_{k,i}$. The chosen matrix T_0 and the computed scalar λ impose respectively the shape and the size of $\mathcal{S}(\theta)$. In order to increase model accuracy, $\mathcal{S}(\theta)$ must be the smaller domain centered on θ_c and containing at least one point of $\underline{\mathcal{D}}_h$ and another one of $\overline{\mathcal{D}}_h$ according to its shape imposed by (4) [2].

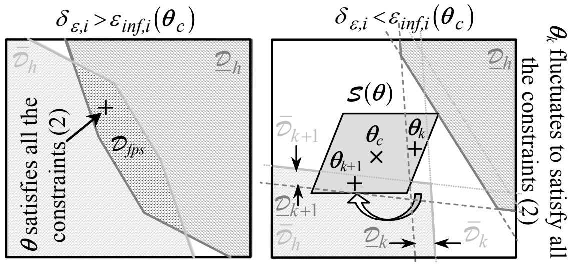

Fig 2. Influence of the bound $\delta_{\varepsilon,i}$

To conclude, if $\delta_{\varepsilon,i} > \varepsilon_{inf,i}(\theta_c)$, the error between $\tilde{y}_{k,i}$ and $\tilde{x}_{k,i}^T\theta_c$ is contained in the equation error, what leads to λ=0 and to compute the F.P.S. Each of its points satisfies all inequality constraints (2), which is restrictive for a model. If parameters can fluctuate inside the F.P.S., they can also be taken as time-invariant (i.e. certain) parameters. Otherwise, if $\delta_{\varepsilon,i} < \varepsilon_{inf,i}(\theta_c)$, the value of λ is different from 0 and the parameter vector θ_k necessary becomes uncertain and time-variant since no constant point of $\mathcal{S}(\theta)$ can verify all considered constraints at the same time. This explains the difference between set-membership parameter estimation and the work proposed in this paper about characterization of bounded uncertainties.

3 Characterization procedure

3.1 Basic solution

By using the relation (4), the output vector y_k of the model

(1) is expressed as:

$$\boldsymbol{y}_k = \tilde{\boldsymbol{X}}_k^T \boldsymbol{\theta}_c + \lambda \tilde{\boldsymbol{X}}_k^T \boldsymbol{T}_0 \boldsymbol{\upsilon}_k + \boldsymbol{\varepsilon}_k . \quad (5)$$

In the paper [1], authors proposed a solution in order to compute the coefficient λ (when $\boldsymbol{\theta}_c$ and $\delta_{\varepsilon,i}$ are fixed) such that the obtained single output model most precisely explains all the observations on the time horizon h. The same method is here applied by considering individually the m equations (i.e. the m single outputs $y_{k,i}$) occurring in (5):

$$y_{k,i} = \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c + \lambda \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \boldsymbol{\upsilon}_k + \varepsilon_{k,i} \,,\ i \in \{1, \cdots, m\}. \quad (6)$$

The principle consists in expressing $\mathcal{S}(y_{k,i})$, in others words $\mathcal{S}\left(\tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c + \lambda \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \boldsymbol{\upsilon}_k + \varepsilon_{k,i}\right)$, by using interval analysis [14], [15], what leads to m two-sides inequalities at each time $k \in \{1,\ldots,h\}$:

$$-\lambda \left\| \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \right\|_1 - \delta_{\varepsilon,i} \le y_{k,i} - \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c \le \lambda \left\| \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \right\|_1 + \delta_{\varepsilon,i} . \quad (7)$$

In order to be consistent with (6), every observation $\tilde{y}_{k,i}$ must belong to $\mathcal{S}(y_{k,i})$ by verifying (7). Thus, λ must satisfy the following inequality for $k \in \{1,\ldots,h\}$, $i \in \{1,\ldots,m\}$:

$$\lambda \ge max \left\{ 0, \frac{\tilde{y}_{k,i} - \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c - \delta_{\varepsilon,i}}{\left\| \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \right\|_1}, \frac{\tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c - \tilde{y}_{k,i} - \delta_{\varepsilon,i}}{\left\| \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \right\|_1} \right\}.$$

Only the most accurate of models (1) interests us. Since the scalar λ imposes the size of $\mathcal{S}(\boldsymbol{\theta})$ and adjusts model uncertainty, the most precise model, which corresponds to the smallest domain $\mathcal{S}(\boldsymbol{\theta})$, is obtained by minimizing λ:

$$\lambda = \sup_{\substack{k \in \{1,\ldots,h\} \\ i \in \{1,\ldots,m\}}} \left\{ max \left(0, \frac{\left| \tilde{y}_{k,i} - \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{\theta}_c \right| - \delta_{\varepsilon,i}}{\left\| \tilde{\boldsymbol{x}}_{k,i}^T \boldsymbol{T}_0 \right\|_1} \right) \right\}.$$

3.2 Dependence problem

The dependence phenomenon met in interval analysis follows from couplings between model equations due to common bounded variables. Assume that a vector field $\boldsymbol{\theta}$ depends on a bounded vector $\boldsymbol{\upsilon}$: $\boldsymbol{\theta}=\boldsymbol{f}(\boldsymbol{\upsilon})$. If no component υ_j of $\boldsymbol{\upsilon}$ intervenes in several functions θ_i of $\boldsymbol{\theta}$ at once, that is to say if every component of $\boldsymbol{\theta}$ is independent according to $\boldsymbol{\upsilon}$, then $\mathcal{S}(\boldsymbol{\theta})$ is an axis-aligned orthotope (or box).

If at least one variable υ_j is common to several functions θ_i, the shape of $\mathcal{S}(\boldsymbol{\theta})$ is modified because of couplings between these θ_i [3]. Since only models linear in uncertainties are considered, $\mathcal{S}(\boldsymbol{\theta})$ corresponds necessarily to a convex zonotope [16]. It is included in its axis-aligned circumscribed orthotope, noted $\square\mathcal{S}(\boldsymbol{\theta})$, which corresponds to the smallest box containing $\mathcal{S}(\boldsymbol{\theta})$. In fact, the domain $\square\mathcal{S}(\boldsymbol{\theta})$ is built by assuming no variable is common, that is to say by treating independently each function θ_i.

For the treated problem, it is important to notice that the m single model outputs $y_{k,i}$ defined in (6) are linked through some components of the bounded vector $\boldsymbol{\upsilon}_k$ at the time k. Unfortunately, the basic method does not take couplings between these outputs into account. Therefore, $\mathcal{S}(\boldsymbol{\theta})$ is computed such that the measurement vector $\tilde{\boldsymbol{y}}_k$ belongs to the axis-aligned orthotope $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$, instead of the zonotope $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ defining the true set of feasible output vectors according to $\mathcal{S}(\boldsymbol{\theta})$. More exactly, since $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ is included in $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$, $\tilde{\boldsymbol{y}}_k$ does not necessarily belong to $\mathcal{S}_\lambda(\boldsymbol{y}_k)$; therefore, our objective is not reached.

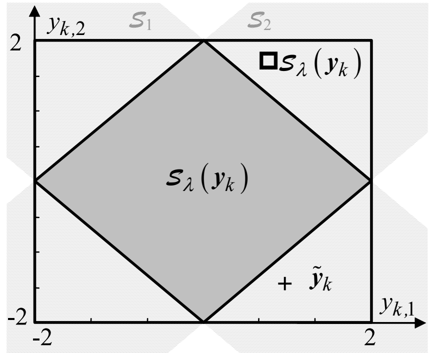

Fig 3. Domains $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ and $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$

In practice, since $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$ is an overestimation of $\mathcal{S}_\lambda(\boldsymbol{y}_k)$, using the former domain increases the inaccuracy of the model, and thus reduces the quality of the procedure based on this one. For example, in fault diagnosis [2], [16], [17], [18] where $\mathcal{S}()$ makes it possible to define the set of all feasible behaviors of the monitored system, working on $\square\mathcal{S}()$ instead of $\mathcal{S}()$ increases the number of no-detections. But, in order to be able to use $\mathcal{S}()$, the characterization procedure must take couplings between common variables into account. Taking them into consideration only in the following step (as a fault detection procedure) is not sufficient. That is the reason why the following section proposes a multi-characterization procedure well suited to models composed of several output relations.

3.3 Improved solution

The objective consists in taking the couplings between model output equations into account. In fact, the shape of $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ is influenced by all common variables and since $\boldsymbol{y}_k$ is linear in the bounded vector $\boldsymbol{\upsilon}_k$ (5), this domain is necessarily a convex zonotope. In other words, it is a convex polytope obtained by the intersection of half-spaces delimited by two by two parallel hyperplanes defining a strip constraint $\mathcal{S}_i$ (figure 3). The method proposed in [17] to construct such a domain consists in determining some combinations of the model output equations (6) making it possible to eliminate at least one common variable $\upsilon_{k,j}$, $j \in \{1,\ldots,q\}$. This procedure generates the strip constraints $\mathcal{S}_i$, which describe the relationships between all output relations and define the Cartesian equation of the frontiers of the zonotope.

For example, let us consider the following model:

$$\begin{bmatrix} y_{k,1} \\ y_{k,2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \upsilon_{k,1} \\ \upsilon_{k,2} \end{bmatrix}.$$

By working individually on each single output $y_{k,i}$ and by using interval analysis, the following two-sides inequalities describing the box $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$ are obtained:

$$\begin{cases} -2 \le y_{k,1} \le 2 \\ -2 \le y_{k,2} \le 2 \end{cases}.$$

Moreover, two linear combinations of both components $y_{k,1}$ and $y_{k,2}$ make it possible to eliminate respectively each common variable $\upsilon_{k,1}$ and $\upsilon_{k,2}$:

$$\begin{cases} y_{k,1} + y_{k,2} = 2\upsilon_{k,1} \\ y_{k,1} - y_{k,2} = 2\upsilon_{k,2} \end{cases} \Rightarrow \begin{cases} -2 \le y_{k,1} + y_{k,2} \le 2 & \mathcal{S}_1 \\ -2 \le y_{k,1} - y_{k,2} \le 2 & \mathcal{S}_2 \end{cases}$$

The intersection of both strip constraints $\mathcal{S}_1$ and $\mathcal{S}_2$ described by previous two-sides inequalities with $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$ gives the zonotope $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ on figure 3.

The method proposed in the following for automatically computing all inequalities associated with strip constraints is based on the results detailed in [17]. By using the model (1) and the expression of $\boldsymbol{\theta}_k$ (4), the matrix $\boldsymbol{M}_k(\lambda)$ associated with the bounded vector $\boldsymbol{\upsilon}_k$ is defined:

$$\boldsymbol{y}_k - \tilde{\boldsymbol{X}}_k\boldsymbol{\theta}_c = \boldsymbol{\varepsilon}_k + \boldsymbol{M}_k(\lambda)\boldsymbol{\upsilon}_k \text{, with } \boldsymbol{M}_k(\lambda) = \lambda\tilde{\boldsymbol{X}}_k\boldsymbol{T}_0. \qquad (8)$$

The matrix $\boldsymbol{M}_k(\lambda)$ is assumed to be full row rank; otherwise, redundant equations of (5) are eliminated since they do not give more additional information. At the time k, the elimination procedure consists in determining all the combinations c_i, $i\in\{1,\ldots,\mathbb{C}_q^{m-1}\}$ of $m-1$ indexes among q. A matrix defined by:

$$\boldsymbol{M}_{k,i}(\lambda) = \left[\boldsymbol{m}_{k,j_1}(\lambda) \quad \cdots \quad \boldsymbol{m}_{k,j_{m-1}}(\lambda)\right],\ j_{(.)} \in c_i,$$

composed of $m-1$ columns of $\boldsymbol{M}_k(\lambda)$, whose indices correspond to the combination c_i, is created. If $\boldsymbol{M}_{k,i}(\lambda)$ is full column rank, a row vector $\boldsymbol{h}_{k,i}^T$ is computed such that:

$$\boldsymbol{h}_{k,i}^T\boldsymbol{M}_{k,i}(\lambda) = 0.$$

In fact, due to the particular structure of $\boldsymbol{M}_k(\lambda)$, the parameter λ does not modify the rank of $\boldsymbol{M}_{k,i}(\lambda)$ when it is different from 0 (i.e when some parameter uncertainties exist). Since λ is unknown during this step, the projection vector $\boldsymbol{h}_{k,i}$ is found by imposing arbitrary λ=1 and working on $\boldsymbol{M}_{k,i}(1)$ instead of $\boldsymbol{M}_{k,i}(\lambda)$.

In fact, the vector $\boldsymbol{h}_{k,i}$ describes one linear combination of outputs $y_{k,i}$, $i\in\{1,\ldots,m\}$ making it possible to eliminate all the common variables, whose indices belong to c_i.

After multiplying (8) by $\boldsymbol{h}_{k,i}^T$, interval arithmetic is used to express the following two-sides inequality:

$$\begin{aligned} -\left|\boldsymbol{h}_{k,i}^T\right|\boldsymbol{\delta}_\varepsilon - \left\|\boldsymbol{h}_{k,i}^T\boldsymbol{M}_k(\lambda)\right\|_1 \le \boldsymbol{h}_{k,i}^T\left(\boldsymbol{y}_k - \tilde{\boldsymbol{X}}_k\boldsymbol{\theta}_c\right) \le \cdots \\ \cdots \le \left|\boldsymbol{h}_{k,i}^T\right|\boldsymbol{\delta}_\varepsilon + \left\|\boldsymbol{h}_{k,i}^T\boldsymbol{M}_k(\lambda)\right\|_1 \end{aligned}, \qquad (9)$$

with $\mathcal{S}(\boldsymbol{\varepsilon}_k) = [-\boldsymbol{\delta}_\varepsilon, \boldsymbol{\delta}_\varepsilon]$, with $\boldsymbol{\delta}_\varepsilon = [\cdots\ \delta_{\varepsilon,i}\ \cdots]^T$.

At the time k, the expression (9) defines one of the strip constraints $\mathcal{S}_i$ describing $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ (excepted those common with $\square\mathcal{S}_\lambda(\boldsymbol{y}_k)$ previously given by (7)). In order to test if the sensor vector belongs to the value set of the unknown output vector, $\boldsymbol{y}_k$ is replaced by $\tilde{\boldsymbol{y}}_k$ in (9). Let n_k be the number of the obtained strip constraints, then λ verifies:

$$\lambda \ge max\left\{0, \frac{\left|\boldsymbol{h}_{k,i}^T\left(\tilde{\boldsymbol{y}}_k - \tilde{\boldsymbol{X}}_k\boldsymbol{\theta}_c\right)\right| - \left|\boldsymbol{h}_{k,i}^T\right|\boldsymbol{\delta}_\varepsilon}{\left\|\boldsymbol{h}_{k,i}^T\tilde{\boldsymbol{X}}_k\boldsymbol{T}_0\right\|_1}\right\},$$

$$\forall k \in \{1,\cdots,h\}$$
$$\forall i \in \{1,\cdots,n_k\}.$$

At the time k, the parameter λ has to verify an inequality system composed of the m constraints (7) and the n_k other ones (9). By assuming that the coefficients $\boldsymbol{\theta}_c$ and $\boldsymbol{\delta}_\varepsilon$ are fixed, the optimal value of the parameter λ corresponds to:

$$\lambda = \underbrace{sup}_{\substack{k\in\{1,\cdots,h\} \\ j\in\{1,\cdots,m\} \\ i\in\{1,\cdots,n_k\}}} \left\{ max\left(0, \cdots \frac{\left|\tilde{y}_{k,j} - \tilde{\boldsymbol{x}}_{k,j}^T\boldsymbol{\theta}_c\right| - \delta_{\varepsilon,j}}{\left\|\tilde{\boldsymbol{x}}_{k,j}^T\boldsymbol{T}_0\right\|_1} \cdots \right.\right. \\ \left.\left. \cdots \frac{\left|\boldsymbol{h}_{k,i}^T\left(\tilde{\boldsymbol{y}}_k - \tilde{\boldsymbol{X}}_k\boldsymbol{\theta}_c\right)\right| - \left|\boldsymbol{h}_{k,i}^T\right|\boldsymbol{\delta}_\varepsilon}{\left\|\boldsymbol{h}_{k,i}^T\tilde{\boldsymbol{X}}_k\boldsymbol{T}_0\right\|_1} \cdots \right)\right\}. \qquad (10)$$

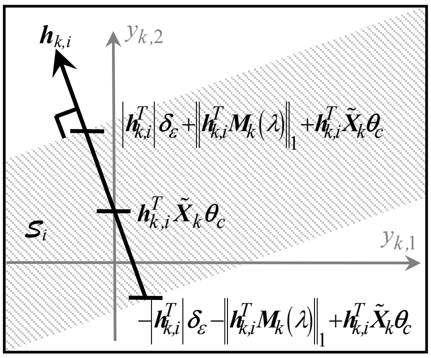

Fig 4. Width of a strip constraint $\mathcal{S}_i$

In this way, every observation $\tilde{\boldsymbol{y}}_k$ belongs to all strip constraints defining $\mathcal{S}_\lambda(\boldsymbol{y}_k)$. Moreover, if the center $\boldsymbol{\theta}_c$ and the bounds $\boldsymbol{\delta}_\varepsilon$ are a priori fixed, the most precise model with respect to the data set is obtained since parameter uncertainties are minimal.

To conclude, it is possible to optimize the (p+1+m)-tuple of coefficients ($\boldsymbol{\theta}_c$,λ,$\boldsymbol{\delta}_\varepsilon$) if a criterion J is chosen. Thus, the chosen solution is the sum of the widths $w_{k,(.)}$ of all the strip constraints of $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ (figure 4) given by (7) and (9):

$$w_{k,j}(\boldsymbol{\delta}_\varepsilon,\lambda) = 2\delta_{\varepsilon,j} + 2\lambda\left\|\tilde{\boldsymbol{x}}_{k,j}^T\boldsymbol{T}_0\right\|_1,\ \forall j \in \{1,\cdots,m\},$$

$$w_{k,i}(\boldsymbol{\delta}_\varepsilon,\lambda) = 2\left|\boldsymbol{h}_{k,i}^T\right|\boldsymbol{\delta}_\varepsilon + 2\left\|\boldsymbol{h}_{k,i}^T\tilde{\boldsymbol{X}}_k\boldsymbol{T}_0\right\|_1\lambda,\ \forall i \in \{1,\cdots,n_k\}.$$

On the horizon h, the criterion J becomes:

$$J = \sum_{k=1}^{h}\sum_{i=1}^{n_k+m} w_{k,i}(\boldsymbol{\delta}_\varepsilon,\lambda). \qquad (11)$$

Even if $\boldsymbol{\theta}_c$ does not appear in (11), J depends implicitly on it by means of constraints (9) to be satisfied since the value of $\boldsymbol{\theta}_c$ directly influences the parameter λ as shown on the figure 2. In fact, this is a well-known linear optimization problem under linear inequality constraints.

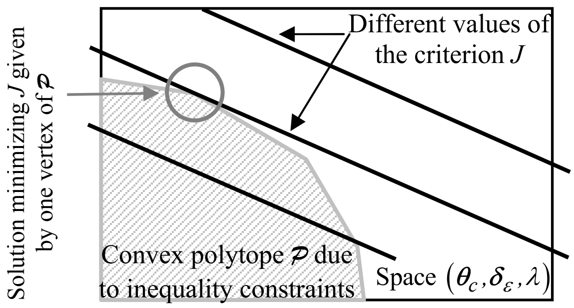

Fig 5. Linear optimization problem

4 Example

For example, consider the following "physical" system, whose "sensor" outputs are simulated on the time horizon h=100 and are represented by symbols "+" in figure 6:

$$\begin{bmatrix} \tilde{y}_{k,1} \\ \tilde{y}_{k,2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.8 & 1.2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \upsilon_{k,1} \\ \upsilon_{k,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}, \begin{array}{l} \delta_{\varepsilon,i} = 0.2 \\ i \in \{1,2\} \end{array}. \quad (12)$$

The model used for the characterization procedure is slightly different and is given as:

$$\begin{bmatrix} y_{k,1} \\ y_{k,2} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\theta_c} + \lambda \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{T_0} \begin{bmatrix} \upsilon_{k,1} \\ \upsilon_{k,2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{k,1} \\ \varepsilon_{k,2} \end{bmatrix}, \begin{array}{l} \delta_{\varepsilon,i} = 0.1 \\ i \in \{1,2\} \end{array}, \quad (13)$$

which corresponds to the following constraints:

$$\begin{array}{l} y_{k,1} - 0.1 \le \lambda\left(\upsilon_{k,1} + \upsilon_{k,2}\right) \le y_{k,1} + 0.1 \\ y_{k,2} - 0.1 \le \lambda\left(\upsilon_{k,1} - \upsilon_{k,2}\right) \le y_{k,2} + 0.1 \end{array}. \quad (14)$$

In a general manner, the time-variant regression matrix $\tilde{\boldsymbol{X}}_k$ depends on measurements, and thus the size and the shape of $\mathcal{S}_\lambda(\boldsymbol{y}_k)$ change at each time k on the time horizon h. In this example, $\tilde{\boldsymbol{X}}_k$ is reduced to the identity matrix in order to simplify the representation of $\mathcal{S}_\lambda(\boldsymbol{y}_k)$. In this way, this domain is time-invariant and only one value set is drawn for the whole time horizon h; thus it is easier to verify if all the observations $\tilde{\boldsymbol{y}}_k$ belong to $\mathcal{S}_\lambda(\boldsymbol{y}_k)$.

4.1 Basic method

At first, the basic method is applied to previous relations (13) and gives the following result: $\lambda = 0.98$, what leads to both domains $\square\mathcal{S}_{\lambda=0.98}(\boldsymbol{y}_k)$ and $\mathcal{S}_{\lambda=0.98}(\boldsymbol{y}_k)$ drawn in figure 5.

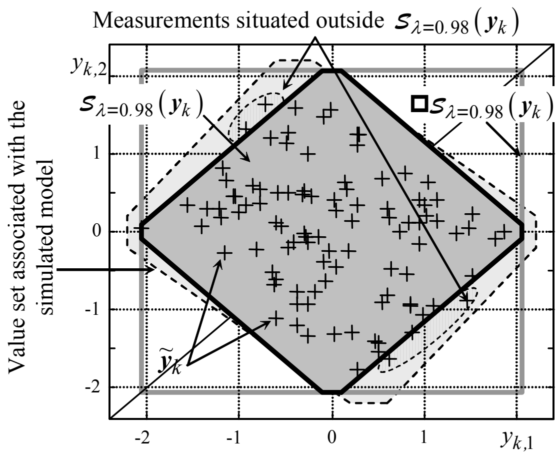

Fig 6. Results of the characterization procedure

Measurements belong to the axis-aligned orthotope $\square\mathcal{S}_{\lambda=0.98}(\boldsymbol{y}_k)$ but, as expected, some of them are not inside the characterized zonotope $\mathcal{S}_{\lambda=0.98}(\boldsymbol{y}_k)$. Due to the structure of the matrix $\boldsymbol{T}_0$, both bounded variables $\upsilon_{k,i}$, $i \in \{1,2\}$, are common, that is the reason why some dependencies exist between both equations (13).

The value set associated with the simulated model (12) is also represented. Because of the limited size of the data set, if all measurements are logically found inside this domain, none of them is located on one of its frontiers since the number of points is low.

4.2 Improved method

Let us consider the model (13) used for the characterization procedure. Two linear combinations of both components $y_{k,1}$ and $y_{k,2}$ make it possible to eliminate respectively each common variable $\upsilon_{k,1}$ and $\upsilon_{k,2}$:

$$\begin{array}{l} y_{k,1} + y_{k,2} = 2\lambda\upsilon_{k,1} + \varepsilon_{k,1} + \varepsilon_{k,2} \\ y_{k,1} - y_{k,2} = 2\lambda\upsilon_{k,2} + \varepsilon_{k,1} - \varepsilon_{k,2} \end{array}, \quad (15)$$

what leads to the following two-sides inequalities:

$$\begin{cases} y_{k,1} + y_{k,2} - 0.2 \le 2\lambda\upsilon_{k,1} \le y_{k,1} + y_{k,2} + 0.2 \\ y_{k,1} - y_{k,2} - 0.2 \le 2\lambda\upsilon_{k,2} \le y_{k,1} - y_{k,2} + 0.2 \end{cases}. \quad (16)$$

Let us notice:

$$\boldsymbol{M}_k(\lambda) = \lambda\tilde{\boldsymbol{X}}_k\boldsymbol{T}_0 = \lambda\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

The numeric algorithm reminded in section 3.3 makes it possible to generate systematically expressions (16). At first, the combinations c_i, $i \in \{1, \ldots, \mathbb{C}_q^{m-1}\}$ of $m-1=1$ index among $q=2$ indices are determined. In our case, one index among the set {1,2} gives two combinations c_1={1} and c_2={2}. For each combination c_i, if the matrix composed of the columns of $\boldsymbol{M}_k(1)$, whose indexes correspond to the elements of c_i, is full column rank, then a vector $\boldsymbol{h}_i$ orthogonal to this matrix is computed.

Numerically, the following results are obtained:

$$\boldsymbol{h}_1^T = \begin{bmatrix} 1 & -1 \end{bmatrix}, \quad y_{k,1} - y_{k,2} = 2\lambda\upsilon_{k,2} + \varepsilon_{k,1} - \varepsilon_{k,2},$$
$$\boldsymbol{h}_2^T = \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad y_{k,1} + y_{k,2} = 2\lambda\upsilon_{k,1} + \varepsilon_{k,1} + \varepsilon_{k,2},$$

which corresponds to previous expressions (15). At the end, by replacing $\boldsymbol{y}_k$ by $\tilde{\boldsymbol{y}}_k$ in (14) and (16), the following two-sides inequalities must be satisfied at each time k:

$$\begin{cases} \tilde{y}_{k,1} - 0.1 \le \lambda\left(\upsilon_{k,1} + \upsilon_{k,2}\right) \le \tilde{y}_{k,1} + 0.1 \\ \tilde{y}_{k,2} - 0.1 \le \lambda\left(\upsilon_{k,1} - \upsilon_{k,2}\right) \le \tilde{y}_{k,2} + 0.1 \\ \tilde{y}_{k,1} + \tilde{y}_{k,2} - 0.2 \le 2\lambda\upsilon_{k,1} \le \tilde{y}_{k,1} + \tilde{y}_{k,2} + 0.2 \\ \tilde{y}_{k,1} - \tilde{y}_{k,2} - 0.2 \le 2\lambda\upsilon_{k,2} \le \tilde{y}_{k,1} - \tilde{y}_{k,2} + 0.2 \end{cases}.$$

According to (10), the constraints to be respected are:

$$\lambda = \sup_{k \in \{1,\ldots,h\}} \left\{ max\left(0, \frac{\left|\tilde{y}_{k,1}\right| - 0.1}{2}, \frac{\left|\tilde{y}_{k,2}\right| - 0.1}{2}, \ldots \right.\right.$$
$$\left.\left. \ldots \frac{\left|\tilde{y}_{k,1} + \tilde{y}_{k,2}\right| - 0.2}{2}, \frac{\left|\tilde{y}_{k,1} - \tilde{y}_{k,2}\right| - 0.2}{2} \right) \right\}.$$

The value of the parameter λ is 1.08. Moreover, zonotopes $\mathcal{S}_{\lambda=1.08}(\boldsymbol{y}_k)$ and $\mathcal{S}_{\lambda=0.98}(\boldsymbol{y}_k)$ are represented in figure 7.

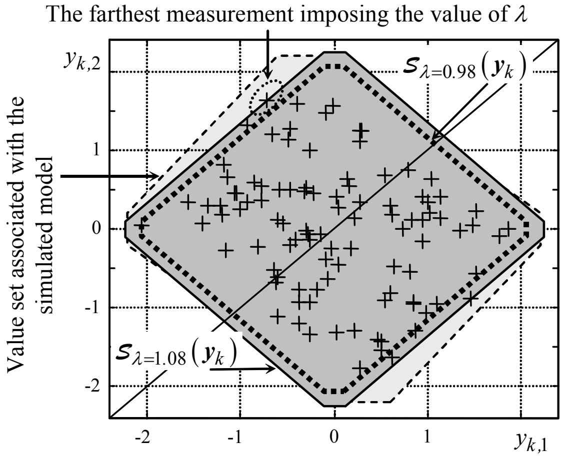

Fig 7. Results of the multi-characterization

As expected, $\mathcal{S}_{\lambda=1.08}(\boldsymbol{y}_k)$ contains all measurements. For the imposed values of $\boldsymbol{\theta}_c$ and $\delta_{\varepsilon,i}$, $i\in\{1,2\}$ (13), the characterized model is the most precise one since a smaller value of λ leads to some measurements, which cannot belong to $\mathcal{S}_{\lambda}(\boldsymbol{y}_k)$. The difference between $\mathcal{S}_{\lambda=1.08}(\boldsymbol{y}_k)$ and the value set associated with the simulated model is due to the small number of data samples and the different values of the matrices $\boldsymbol{T}_0$ used in (12) and (13).

Now, let us consider another simulation. The matrix $\boldsymbol{T}_0$ of the characterized model (13) is the same as the one of the simulated model in order to easily compare the obtained results:

$$\boldsymbol{y}_k = \boldsymbol{\theta}_c + \lambda \begin{bmatrix} 0.8 & 1.2 \\ 1 & -1 \end{bmatrix} \upsilon_k + \varepsilon_k \,. \tag{17}$$

The (2+1+2)-tuple of coefficients $(\boldsymbol{\theta}_c, \lambda, \delta_\varepsilon)$ has to minimize the precision criterion (11). The data set is composed of 5.10^3 measurement vectors $\tilde{\boldsymbol{y}}_k$. By assuming that both bounds $\delta_{\varepsilon,i}$ are equal, the results are given as:

$$\boldsymbol{\theta}_c = \begin{bmatrix} -1.10^{-3} \\ 8.10^{-3} \end{bmatrix},\ \lambda = 0.98 \text{ and } \delta_{\varepsilon,i} = 0.19 \,. \tag{18}$$

The difference between (18) and the theoretic values:

$$\boldsymbol{\theta}_c = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \lambda = 1 \text{ and } \delta_{\varepsilon,i} = 0.2 \,,$$

is due to the finite size of data samples. Despite of this difference, the model (17) with estimated parameters (18) fully explains all measurements.

5 Conclusion

An algorithm for characterizing uncertainties in static linear models containing several output relations is proposed. The main contribution of this work concerns the taking into account of the couplings between bounded variables, which impose the shape of the studied value set. A precision criterion is defined in order to compute the most precise model. The proposed technique may be extended to dynamical systems provided that they are not represented by a recursive form. Moreover, it is possible to apply different weights λ_i to each column of the matrix $\boldsymbol{T}_0$ and thus to adjust the shape of the value set $\mathcal{S}(\boldsymbol{\theta})$.

6 References

[1] S. Ploix, O. Adrot, J. Ragot, Parameter Uncertainty Computation in static Linear Models, *IEEE CDC'99*, 1999.

[2] O. Adrot, Diagnostic à base de modèles incertains utilisant l'analyse par intervalles : l'approche bornante, *Ph.D. of the "Institut National Polytechnique de Lorraine", France*, 2000.

[3] O. Adrot, D. Maquin, J. Ragot, Bounding approach to the fault detection of uncertain dynamic systems, *IFAC Safeprocess2000*, 2000.

[4] E. Fogel, Huang Y.F., On the value of information in system identification-bounded noise case, *Automatica, 18*(2), 1982, 229-238.

[5] Milanese M., Belforte G., Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: linear families of models, *IEEE Transactions on Automatic Control, 27*(2), 1982, 408-413.

[6] H. Piet-Lahanier, Walter E., Exact recursive characterization of feasible parameter sets in the linear case, *Math. and Comput. in Simul., 32,* 1990, 495-504.

[7] V. Broman, Shensa M. J., A compact algorithm for the intersection and approximation of N-Dimensional polytopes, *12th IMACS World Congress*, 1988.

[8] S.H. Mo, Norton J.P., Fast and robust algorithm to compute exact polytope parameter bounds, *Math. and Comput. in Simul., 32,* 1990, 481-493.

[9] T. Clément, Gentil S., Recursive membership set estimation for ARMAX models: an output-error approach, *12th IMACS World Congress*, 1988.

[10] V. Cerone, Parameter bounds for ARMAX models from records with bounded errors in variables, *International Journal of Control, 57*(1), 1993, 225-235.

[11] S.M. Veres, Norton J.P., Parameter-bounding algorithms for linear errors-in-variables models, in *Milanese M., Norton J., Piet-Lahanier H., Walter E., Bounding approches to system identification Plenum Press, New York and London, Chap 17,* 1996, 275-288.

[12] L. Jaulin, Walter E., Guaranteed nonlinear parameter estimation from bounded-error data via interval analysis, *Mathematics and Computers in Simulation, 35,* 1993, 123-137.

[13] M. Milanese, J. Norton, H. Piet-Lahanier, E. Walter, Bounding approaches to system identification (Plenum Press, New York & London, 1996).

[14] R. E. Moore, Methods and applications of interval analysis (SIAM, Philadelphia, Pennsylvania, 1979).

[15] A. Neumaier, Interval methods for systems of equations (Cambridge University Press, Cambridge, 1990).

[16] O. Adrot, D. Maquin, J. Ragot, Diagnosis of an uncertain static system, *IEEE CDC'2000*, 2000.

[17] S. Ploix, O. Adrot, J. Ragot, Bounding approach to the diagnosis of uncertain static systems, *IFAC Safeprocess2000,* 2000.

[18] J. Armengol, L. Travé-Massuyés, J. Vehi, J.L. de la Rosa, A survey on interval model simulators and their properties related to fault detection, *IFAC Beijing'99,* 1999.