PARAMETER UNCERTAINTIES CHARACTERISATION FOR LINEAR MODELS.

José Ragot, * **Didier Maquin,** * **Olivier Adrot** **

* *Centre de Recherche en Automatique de Nancy, CNRS UMR 7039, INPL. 2, Avenue de la forêt de Haye, F 54516 Vandœuvre-les-Nancy Cedex,* {José.Ragot, Didier.Maquin}@ensem.inpl-nancy.fr
** *Laboratoire d'Automatique de Grenoble, CNRS UMR 5528, INPG, 961, rue de la Houille Blanche, BP 46, F 38 402 Saint Martin d'Hères Cedex,* Olivier.Adrot@lag.ensieg.inpg.fr

Abstract: Parameter estimation using the set-membership approach mainly consists in characterising the minimal parameter set consistent with measurements, the model and the equation error description. In this context, it is assumed that the measurement error is bounded and must belong to a prior feasible set to be admissible. The problem to be solved is that of finding the set of all admissible parameter values corresponding to an admissible error. The uncertainties must be treated by a global analysis of the problem: both the equation error and the parameter set are considered unknown. Then, a solution is given as a domain of time-variant parameters and a bounded set of the error. This procedure consists in explaining the measurements performed at all time by optimising a precision criterion based on the polytope theory.

Keywords: bounded error, characterization, uncertain model, set-membership approach, uncertainties, parameter estimation.

1. INTRODUCTION

1.1 Historical point of view

The parameter estimation using the set-membership approach started in the eighties, where the strategy initially consists in circumscribing the domain describing model uncertainties by a simple form. This approach was originally designed to deal with a model linear in uncertain parameters and characterised by a bounded error. The problem of parameter estimation amounts to the determination of the set of constant parameter values called the Feasible Parameter Set (F.P.S.). This set explains all the available observations which are consistent with the bounds of the error and the model structure. For models linear in their parameters, the F.P.S. is a convex polytope which can be approximated by ellipsoids Fogel and Huang (1982), Pronzato et al. (1963), or orthotopes Milanese and Belforte (1982) containing it. The work in Walter and Piet-Lahanier (1987),Walter and Piet-Lahanier (1989) on the one hand, and Mo and Norton (1990) on the other, used polytopic domains. The main results are presented in the book published by Milanese et al. (Milanese et al. (1996)). For models nonlinear in their parameters, various methods exist for determining an approximation of the F.P.S., linear techniques have been extended to the nonlinear case using multiple linearisation of the model Belforte et al. (1990), Jaulin (2001). In Reinelt et al. (2002) a robust identification approach is proposed taking into account unmodeled dynamics and noise affecting data; as uncertainty is evaluated in terms of frequency response, so that it can be handled by $\mathcal{H}_\infty$ technics. In ElGhaoui and Calafiore (2000), the authors explain that the set of possible models is unfalsified by the observed data if that data could actually have been produced by one member of the model set; they formulate the arx models identification as a semidefinite programming program. More recent results have been obtained (Jaulin (1993)) in order to solve the problem of nonlinear bounded-error estimation using set inversion techniques and based on interval analysis, there make it possible to characterise the F.P.S. by enclosing it between internal and external unions of boxes. The paper Jaulin (2001) deals with a minimax parameter estimation of nonlinear parametric models from experimental data. For specific model structures, it is possible to obtain sets of linear inequalities describing a domain approximating the F.P.S. Norton (1987), Clement and Gentil (1988). Despite the resemblance, the problem considered in Ploix et al. (1999) is noticeably different in the sense that uncertain parameters depend on time; more exactly, they are defined by random variables with bounded realisations; moreover, this paper only deals with MISO representation. The proposed method is a no probability technique for determining the inaccuracy with which each model parameter is known. Only a class of structured and static models linear in uncertain parameters is considered. The error is bounded while parameters fluctuate inside a time-invariant bounded domain represented by a convex parallelotope. In Chisci et al. (1998), a recursive approximation of the F.P.S. of a linear model is proposed, the approximation being performed though parallelotopes chosen according to a maximum volume criteriion.

Thus, the paper deals with parameters estimation in a bounded-error context for models which are linear in the parameters; parameters could vary with a limited range and measuremen errors are bounded but both domains are not a priori known. The objective of this method is to determine the characteristics of this domain (centre, uncertainty range). The idea is to determine the nominal value of the parameter vector (in fact the centre of the polytope) and some time-variant uncertainties making it possible to be compatible with the current observation. Maximal magnitudes of these uncertainties make it possible to deduce the characteristics of the considered domain. By fluctuating inside this one, parameters can explain all measurements. Moreover, in order to obtain the most precise model, the estimation problem is then to find the smallest domain.

1.2 Model structure

In order to present the principle of the proposed method, let us consider the following MISO model:

$$y_m(k) = x^T(k)\theta(k) \qquad k = 1..N \tag{1a}$$

$$y(k) = y_m(k) + e(k) \tag{1b}$$

where $y_m(k)$ is the model output, $x(k) \in \mathcal{R}^p$ is the vector of known regressor at the instant k, $y(k)$ is the output measurement, whereas $\theta(k) \in \mathcal{R}^p$ defines the uncertain parameter vector. The error $e(k)$ is assumed to be bounded, the bounds being supposed invariant along the time:

$$e(k) \in [-\delta, \delta] \tag{2}$$

where the positive real δ is not necessary known. Thus, taking (2) into account, (1b) leads to both following inequalities:

$$y(k) - \delta \leq x^T(k)\theta(k) \qquad k = 1..N \tag{3a}$$

$$x^T(k)\theta(k) \leq y(k) + \delta \qquad k = 1..N \tag{3b}$$

Thus, at each instant k, the known measurement $y(k)$ belongs to a domain defined by (3) and the shape of this domain depends both on the bound δ and the particular value $\theta(k)$. In the following, our objective is to characterise this domain at each instant k, i.e. to estimate θ and δ. In fact, we could find a lot of sets of parameter $\theta(k)$ or errors δ satisfying (3) and we have to define a selection criterion and a way to estimate the model parameters.

1.3 The parameters characterization problem

Let us now formulate the preceding remark for any linear system with bounded time-varying parameters. The system (1) with constraint (2) generates, at each instant k, a pair of half-spaces $\underline{H}(k)$ and $\overline{H}(k)$ in $\mathcal{R}^p$ which frontiers define two parallel hyperplanes in the parameter space:

$$\underline{H}(k) = \{\theta' \in D_0 / y(k) - \delta \leq x^T(k)\theta'\} \tag{4a}$$

$$\overline{H}(k) = \{\theta'' \in D_0 / x^T(k)\theta'' \leq y(k) + \delta\} \tag{4b}$$

where D_0 is the domain of investigation chosen by the user (for simplicity, it is taken as an orthotope). More generally, let us now consider observations of the system at the time $k = 1..N$. The intersection of half-spaces being convex, both following domains are convex too:

$$\underline{D}_N = D_0 \bigcap_{k=1}^{N} \underline{H}(k) \tag{5a}$$

$$\overline{D}_N = D_0 \bigcap_{k=1}^{N} \overline{H}(k) \tag{5b}$$

and both domains can be computed using the recurrence:

$$\underline{D}_k = \underline{D}_{k-1} \cap \underline{H}(k), \quad \underline{D}_0 = D_0, \quad k = 1..N$$
$$\overline{D}_k = \overline{D}_{k-1} \cap \overline{H}(k), \quad \overline{D}_0 = D_0, \quad k = 1..N$$

The polytope $\underline{D}_N$ defines the set of values θ' satisfying all inequalities (3a). Among these values, let us define θ'_0 which leads to an upper bound $\overline{y}(k)$ of the measurement $y(k)$ at each time k: $\overline{y}(k) = x^T(k)\theta'_0 + \delta$. In the same manner, $\overline{D}_N$ defines the set of values θ'' satisfying all inequalities (3b). Thus, θ''_0 is the particular value which leads to a lower bound $\underline{y}(k)$ of the measurement $y(k)$ at each instant k: $\underline{y}(k) = x^T(k)\theta''_0 - \delta$. Thus $[\theta_0^{'} \;\; \theta_0^{''}]$ is the set of all values of $\theta(k)$ that are compatible with the model structure and the whole set of measurements $y(k)$ for $k = 1..N$:

$$y(k) - x^T(k)\theta_0(k) \in [-\delta, \delta], \theta_0(k) \in [\theta_0^{'} \;\; \theta_0^{''}] \quad (6)$$

In this context, the method proposed by in (Ploix et al. (1999)) consists in finding a convex parallelotope (its mathematical description is explained thereafter) D_N centred on θ_c and defined by:

$$D_N = \{\theta(k) \in \mathcal{R}^p / \theta(k) = \theta_c + M(\lambda)\nu(k) \quad (7)$$
$$M(0) = 0, \quad \| \nu(k) \|_\infty \leq 1$$

such that it contains, at each instant k, a value of the time-variant parameter vector $\theta(k)$ which are fully compatible with the measurement $y(k)$. The matrix $M(\lambda)$ characterises the shape of the domain D_N, λ being parameters for adjusting the dimension of that domain. In this way, $\theta(k)$ fluctuates around its central value θ_c inside D_N for satisfying all constraints (2), θ_c being considered as the nominal value of the parameter $\theta(k)$. In order to increase the model precision, D_N must be the smaller domain centred on θ_c and containing at least one point of $\underline{D}_N$ and another one of $\overline{D}_N$ (or vice versa) with respect to the form imposed by (7).

The problem treated herein is the computation of the central parameter value θ_c, the parameter uncertainties and an appropriate characterisation of the error domain for MIMO systems. Thus, the characterisation procedure consists in determining the bounds of model uncertainties (λ, δ) and the center θ_c which are totally compatible with the set of available measurements. Since this step leads to a set of possible solutions, the choice of one of them is obtained by optimising a precision criterion, which is related to the dimensions of the domain described by uncertain parameters thus estimated.

The next sections contain our contribution. This paper is organised as follows. In the next section, a formalisation of the problem is detailed: uncertainties affecting the model are well described in order to be familiarised with the methods treating them; the way to construct a feasible system set (FPS) is presented. In section 3, a precision criterion is defined and computed in order to identify, among the FPS the most precise model, i.e. those having the minimum uncertainty. The principle of parameter estimation while optimising the given criterion is presented in section 4. In section 5, an example illustrates the proposed method.

2. PROBLEM FORMULATION

We describe in subsections 2.1 and 2.2 the structure of an uncertain system and the uncertainties; subsection 2.3 provides an academic example of that description. In subsection 2.4, we define a time-invarianty parallelotope in the parameter space such that, at each instant k, it contains at least one value of the time-varying parameter $\theta(k)$ consistent with the observations. Then in subsection 2.5, it is shown that the measurement equation maps the parallelotope $\mathcal{P}_\theta$ in a new parallelotope $\mathcal{P}_Y$ in the measurement space. The shape of $\mathcal{P}_Y$ will be completely defined in respect to the model parameters. Further (sections 3 and 4), we aim to define the smallest parallelotope $\mathcal{P}_Y$ guaranteed to contain all the available measurements. An academic example is presented in section 5 [1].

2.1 Modelling of an uncertain system

In order to generalise the representation given by (1), let us consider an uncertain model of a system with several outputs, linear in parameters and observations, and represented by the following structure:

$$Y(k) = X(k)\theta(k) + E(k) \quad k = 1..N \quad (8)$$

where $Y(k) \in \mathcal{R}^n$, $X(k) \in \mathcal{R}^{n \cdot p}$ are the known variables at the time k and $\theta(k) \in \mathcal{R}^p$ defines model parameters. The bounded vector $E(k) \in \mathcal{R}^n$ defines the error taking into account the uncertainties due to the measuring process and to modelling errors at the same time. This type of model includes the particular case of MISO systems and that of MIMO systems. In the MIMO case, according to the presence of uncertain parameters in $\theta(k)$ (see definition 7), the outputs $Y(k)$ can be coupled by some of the uncertain parameters $\nu(k)$ and that can lead to some difficulties in the estimation problem.
Let us consider the variables, $X(k)$ and $Y(k)$ of which the measurements are noted respectively

[1] The authors are very grateful to Hicham Janati Idrissi for his help concerning simulation of some parts of the proposed approach.

$\tilde{X}(k)$ and $\tilde{Y}(k)$. The problem involved with parameter estimation is to characterise the unknown parameters of a model using experimental data. In other words, the aim is to determine the parameter domain containing all possible values consistent with data for bounds E_{min} and E_{max}, such that:

$$\Theta = \{\theta(k) \in \mathcal{R}^p / \tilde{Y}(k) \in \tilde{X}(k)\theta(k) + [E_{min}\; E_{max}]\} \quad (9)$$

In the case of time-invariant parameters, Milanese and Belforte (1982) suggest approximating Θ with an orthotope aligned with the parameter coordinate axes and finding the minimal and maximal values of $\theta_i, i = 1..p$, by using linear programming. Fogel and Huang (1982) propose an ellipsoidal outer-bounding recursive algorithm : the ellipsoid centre and symmetric positive-defined matrix are considered, respectively, as the parameter central value and its measure of uncertainty. In the case of time varying parameters, an outer-bounding of Θ noted $\mathcal{P}_\theta$ is given in this paper, such that $\mathcal{P}_\theta$ is a parallelotope which some properties will be described in the following paragraph.

2.2 Description of the uncertainties

Uncertainties affecting a system are classified into two categories. On the one hand, those acting directly on the output are additive uncertainties $E(k)$, and on the other hand, the uncertainties describing the parameter $\theta(k)$ occur in a multiplicative way. Let us describe these two sets of uncertainty.

Additive uncertainties are represented by the vector $E(k) \in \mathcal{R}^n$ assumed to belong to the domain noted $\mathcal{P}_E(\delta)$:

$$\mathcal{P}_E(\delta) = \{Z(\delta)u, \| u \|_\infty \leq 1\} \quad (10)$$

with $\delta = (\delta_1 \dots \delta_n)^T$, $u = (u_1 \dots u_n)^T$ and $Z(\delta) \in \mathcal{R}^{n.n}$. When these uncertainties affect independently each output, $Z(\delta)$ has the following structure:

$$Z(\delta) = \begin{pmatrix} \delta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \delta_n \end{pmatrix}$$

The vector δ defines the magnitude of additive uncertainties which are considered bounded.

Multiplicative uncertainties are represented by the parameter vector $\theta(k) \in \mathcal{R}^p$ which fluctuates in an invariant domain denoted $\mathcal{P}_\theta(\lambda, \theta_c)$, defined by:

$$\mathcal{P}_\theta(\lambda, \theta_c) = \{\theta(k) = \theta_c + M(\lambda)\nu(k), \| \nu(k) \|_\infty \leq 1\} \quad (11)$$

The vector $\nu(k)$ is varying inside an unit hypercube noted $\mathcal{H}_q$ ($\mathcal{H}_q = \{\nu \in \mathcal{R}^q / \| \nu \|_\infty \leq 1\}$). It allows to represent the uncertain nature of model parameters. These uncertainties are distributed on the various components of the vector θ via a full row rank matrix $M(\lambda) \in \mathcal{R}^{p.q}$ (in general $q \geq p$) depending on the vector $\lambda = (\lambda_1 \dots \lambda_q)^T$. In fact, the matrix $M(\lambda)$ defines the volume and the shape of $\mathcal{P}_\theta(\lambda, \theta_c)$. The vector θ_c indicates both the geometrical centre of $\mathcal{P}_\theta(\lambda, \theta_c)$ and the nominal value of the parameter vector. Equation (11), also shows that $\mathcal{P}_\theta(\lambda, \theta_c)$ is the image of the hypercube $\mathcal{H}_q$ under an affine map $\mu_{\theta_c, M(\lambda)}$ defined as:

$$\mu_{\theta_c, M(\lambda)} : \mathcal{R}^q \rightarrow \mathcal{R}^p \quad (12)$$
$$\nu(k) \mapsto \theta(k) = \theta_c + M(\lambda)\nu(k)$$

The domain $\mathcal{P}_\theta(\lambda, \theta_c) = \mu_{\theta_c, M(\lambda)}(\mathcal{H}_q)$ is a parallelotope (called also a zonotope). It is the projection of $\mathcal{H}_q$ on the affine space corresponding to the linear subspace spanned by the rows of $M(\lambda)$ shifted by θ_c; since $\mathcal{H}_q$ is an hypercube in $\mathcal{R}^q$, then $\mathcal{P}_\theta(\lambda, \theta_c)$ is a parallelotope described by some linear inequalities which can be obtained by using the algorithm of Fourier-Motzkin elimination appearing in the book of Ziegler Gunter (1995). In the rest of the paper, the matrix $M(\lambda)$ is supposed having the following structure:

$$M(\lambda) = M Diag(\lambda) \quad (13)$$

2.3 Example of bounded time-varying parameters

Let us build, at a particular instant k which is not indicated, the parameter domain $\mathcal{P}_\theta(\lambda, \theta_c)$ for the following system:

$$\theta = \theta_c + M(\lambda)\nu \quad (14)$$
$$M = \begin{pmatrix} -0.1 & 0.3 & -0.1 \\ 0.1 & 0.2 & 0.0 \end{pmatrix} \; \theta_c = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

Thus $p = 2$ and $q = 3$. Since $\lambda_j \geq 0$ and $| \nu_j | \leq 1$ for $j = 1, 2, 3$, the parameters θ_1 and θ_2 have to satisfy the following inequalities deduced from (12):

$$\begin{aligned} \theta_{c,1} - 0.1\lambda_1 - 0.3\lambda_2 - 0.1\lambda_3 \leq \theta_1 \leq \\ \theta_{c,1} + 0.1\lambda_1 + 0.3\lambda_2 + 0.1\lambda_3 \\ \theta_{c,2} - 0.1\lambda_1 - 0.2\lambda_2 \leq \theta_2 \leq \\ \theta_{c,2} + 0.1\lambda_1 + 0.2\lambda_2 \end{aligned} \quad (15)$$

These inequalities define an aligned orthotope circumscribed to $\mathcal{P}_\theta(\lambda, \theta_c)$ which, however, do not take into account the dependencies between θ_1 and θ_2 introduced by the bounded variables ν_1 and ν_2 (called common variables as appearing in (12)). In fact, it is possible to express these dependencies by eliminating the variable ν_1 (respectively ν_2) and thus we obtain:

$$\theta_1 + \theta_2 = \theta_{c,1} + \theta_{c,2} + 0.5\lambda_2\nu_2 - 0.1\lambda_3\nu_3 \quad (16)$$
$$2\theta_1 - 3\theta_2 = 2\theta_{c,1} - 3\theta_{c,2} - 0.5\lambda_1\nu_1 - 0.2\lambda_3\nu_3$$

In other words, in order to eliminate the variable ν_1 (respectively ν_2) in (14), the vector $(\theta_1 \quad \theta_2)^T$ must be multiplied by a row vector orthogonal to $m_1 = (-0.1 \ \ 0.1)^T$ (respectively $m_2 = (0.3 \ \ 0.2)^T$) which is $h_1^T = (1 \ \ 1)$ (respectively $h_2^T = (2 \quad -3)$).

In the general case where $M(\lambda) \in \mathcal{R}^{p \cdot q}$, the elimination procedure of the common variables consists in finding $p-1$ vectors in $\mathcal{R}^p$, orthogonal to one column of M which contains the common variables ν_i to eliminate (note that this procedure is an equivalent version of the Fourier-Motzkin elimination algorithm adapted to parallelotopes). This version also makes it possible to highlight the dependencies between the various components of θ. Then, using h_1 and h_2 and taking into account the fluctuations of the uncertainties $\nu_j (j = 1, 2, 3)$, the equations (16) leads to the following additional inequalities:

$$\begin{aligned} &\theta_{c,1} + \theta_{c,2} - 0.5\lambda_2 - 0.1\lambda_3 \leq h_1^T \theta \\ &\leq \theta_{c,1} + \theta_{c,2} + 0.5\lambda_2 + 0.1\lambda_3 \\ &2\theta_{c,1} - 3\theta_{c,2} - 0.5\lambda_1 - 0.2\lambda_3 \leq h_2^T \theta \\ &\leq 2\theta_{c,1} - 3\theta_{c,2} + 0.5\lambda_1 + 0.2\lambda_3 \end{aligned} \tag{17}$$

Let us remark that the inequalities (16), with $h_3^T = (1 \ \ 0)$ and $h_4^T = (0 \ \ 1)$, can also be expressed using the parameter vector θ:

$$\begin{aligned} &\theta_{c,1} - 0.1\lambda_1 - 0.3\lambda_2 - 0.1\lambda_3 \leq h_3^T \theta \\ &\leq \theta_{c,1} + 0.1\lambda_1 + 0.3\lambda_2 + 0.1\lambda_3 \\ &\theta_{c,2} - 0.1\lambda_1 - 0.2\lambda_2 \leq h_4^T \theta \leq \\ &\theta_{c,2} + 0.1\lambda_1 + 0.2\lambda_2 \end{aligned} \tag{18}$$

Thus, with (17) and (18), it is easy to construct the domain $\mathcal{P}_\theta(\lambda, \theta_c)$ defining the possible values for the parameter θ. The figure 1 shows a geometrical interpretation of these relations for $\lambda_1 = 1$, $\lambda_2 = 1.5$, $\lambda_3 = 2$ and $\theta_c = (4 \ 4)^T$. The strip-band D_i has been constructed using definition (17) and (18). The rectangle with large border indicates the orthotope circumscribing the domain $\mathcal{P}_\theta(\lambda, \theta_c)$ obtained only with (18) while the grey hexagone shows the true domain $\mathcal{P}_\theta$ obtained with (17) and (18).

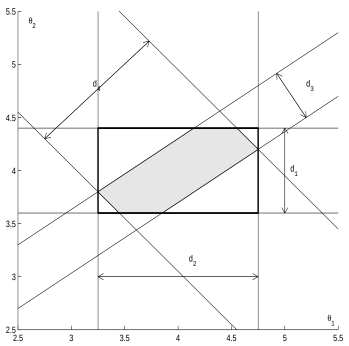

Fig. 1. Parallelotope defined by the intersection of strip-bands

The shape of $\mathcal{P}_\theta(\lambda, \theta_c)$ depends on the vector λ because the width d_i of the strip-band D_i is sensitive to λ. On the other hand, it is also possible to change some d_i without influencing the other distances by acting on some particular components of λ. Hence, this leads to a change of the shape of $\mathcal{P}_\theta(\lambda, \theta_c)$.

2.4 General description of bounded linear time-varying system

Then, according to the previous example (see equations 17 and 18), the general structure of the inequality defining the domain $\mathcal{P}_\theta(\lambda, \theta_c)$ is:

$$\begin{aligned} &h_i^T \theta_c - (\mid h_i^T m_1 \mid ... \mid h_i^T m_q \mid)\lambda \leq h_i^T \theta(k) \\ &\leq h_i^T \theta_c + (\mid h_i^T m_1 \mid ... \mid h_i^T m_q \mid)\lambda \end{aligned} \tag{19}$$

where h_i is either a vector of the identity matrix (equations (11) are directly used) or a vector orthogonal to a row m_j of $M(\lambda)$ (in that case, a combination of equations (11) is used). For each subscript i, the preceding formula defines an unbounded strip-band $D_i \subset \mathcal{R}^p$ limited by two parallel hyperplanes:

$$\begin{aligned} B_i^+ &= \{\theta / h_i^T \theta = h_i^T \theta_c + (\mid h_i^T m_1 \mid ... \mid h_i^T m_q \mid)\lambda\} \\ B_i^- &= \{\theta / h_i^T \theta = h_i^T \theta_c - (\mid h_i^T m_1 \mid ... \mid h_i^T m_q \mid)\lambda\} \end{aligned}$$

The distance between these two hyperplanes is defined by (with an extended definition of the absolute value operator $\mid \ \mid$ which is applied to each component of a vector):

$$d_i = \frac{2}{\sqrt{h_i^T h_i}} (\mid h_i^T m_1 \ ... \ h_i^T m_q \mid)\lambda \tag{20}$$

This distance may represent an indicator of the shape of the domain $\mathcal{P}(\lambda, \theta_c)$. Thus, the strip-band D_i is defined by:

$$D_i = \{\theta \in \mathcal{R}^p / \mid h_i^T(\theta - \theta_c) \mid \leq \frac{d}{2}\sqrt{h_i^T h_i}\} \tag{21}$$

and the domain $\mathcal{P}_\theta(\lambda, \theta_c)$ (18) is defined by the intersection of all D_i: $\mathcal{P}_\theta(\lambda, \theta_c) = \cap_{i=1}^{\overline{r}} D_i$, where $\overline{r}$ is the number of strip-bands.

2.5 Principle of parameter estimation

The parameter estimation problem consists in finding the values of the vectors θ_c, λ and δ which define the parameters domain $\mathcal{P}_\theta(\lambda, \theta_c)$ (11) and the measurement errors domain $\mathcal{P}_E(\delta)$ (10) (see the section 2.2), so that the characterised model explains all the available measurements in the most precise way:

$$\tilde{Y}(k) \in \mathcal{P}_Y(\lambda, \delta, \theta_c) \quad k = 1..N \tag{22}$$

with:

$$\mathcal{P}_Y(\lambda,\delta,\theta_c) = \{Y(k) \in /Y(k) = \tilde{X}(k)\theta_c + \tilde{X}(k)M(\lambda)\nu(k) + Z(\delta)u(k), \| u(k) \|_\infty \leq 1, \| \nu(k) \|_\infty \leq 1\} \quad (23)$$

$\mathcal{P}_Y(\lambda,\delta,\theta_c)$ defines all possible values of the variables $Y(k)$ consistent with variables $X(k)$ and the model uncertainties description given by the vectors λ and δ. So, $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ is an interval estimation of measurements $\tilde{Y}(k)$. Note that $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ is a parallelotope. Indeed, considering (23), if $\tilde{Y}(k) \in \mathcal{P}_Y(\lambda,\delta,\theta_c)$ then

$$\exists w(k) \in \mathcal{H}_{q+n}/\tilde{Y}(k) = \tilde{Y}_c(\theta_c,k) + T(k,\lambda,\delta)w(k) \quad (24)$$

with:

$$T(k,\lambda,\delta) = (\tilde{X}(k)M(\lambda) \;\; Z(\delta))$$
$$\tilde{Y}_c(\theta_c,k) = \tilde{X}(k)\theta_c \quad w(k) = \begin{pmatrix} \nu(k) \\ u(k) \end{pmatrix} \quad (25)$$

Moreover, $\tilde{Y}_c(\theta_c,k)$ is the centre of $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ which depends both on θ_c and the measure $\tilde{X}(k)$. Therefore, all properties studied in the previous section for the parameter domain $\mathcal{P}_\theta(\lambda,\theta_c)$ are also valid for the domain $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ and thus the way to construct $\mathcal{P}_\theta(\lambda,\theta_c)$ may be applied to construct $\mathcal{P}_Y(\lambda,\delta,\theta_c)$. During the parameter estimation step, it is assumed that measurements are not contaminated by systematic skews or accidental errors of great magnitude. Otherwise, the parameters θ_c, λ and δ are unfortunately adjusted for explaining these anomalies, which is not the desired effect. In this paper, the shape of the domain (dtermined by the matrix M) is fixed a priori; whatever the choice of this shape, all parameters parameters that are compatible with measurements, error bounds and model structure will be enclosed in the domain.
If we take $\lambda_i = 0$, the scalars δ_i can be chosen as large as we want for a given value of θ_c, since that consists in increasing the volume of the domain of uncertainties occurring in the model, until being compatible with all measurements.
If the measurements are not affected by errors (δ_i and equal to 0), then the model may be compatible with the measurements by increasing the magnitude of λ_i.
In the other cases, it will be possible to define a criterion representative of the precision, the latter being related to the domain extent: indeed increasing "arbitrarily" the values of λ_i and δ_i in order to explain measurements is not satisfactory. Therefore, it is necessary to find a quantity which is sensitive to the difference between real measurements and their estimates generated by the given characteristics of the model. Ploix et al. (1999), defined a criterion based on interval arithmetic (Moore (1979), Neumaier (1990)) for a model with only one output. In this paper, a MIMO model is studied and the aim is to characterise uncertainties while minimising a criterion of precision related to the dimension of the domain of output estimates (this domain is $\mathcal{P}_Y(\lambda,\delta,\theta_c)$). An obvious and intuitive choice that one can make, is to consider the volume of the domain. It is easy to show that its volume is proportional to the components of λ. Then, the solution is the smallest λ which explains all measurements.
The application of this procedure, when $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ has an complicated form, leads to some calculation problems. Indeed the evaluation of the volume of a polytope leads to an expression containing symbolic functions Lasserre (1983) (max, min), which are unusable to find a solution and make the calculation more delicate; one needs to find a criterion which is at once representative of the model precision and which does not lead to computation difficulties.

3. CHOICE OF THE CRITERION

The aim of this section is to define a mathematical criterion which provides a solution $(\lambda_s,\delta_s,\theta_{c,s})$ in such a way that the domain $\mathcal{P}_Y(\lambda_s,\delta_s,\theta_{c,s})$, corresponding to the estimation of $Y(k)$, contains all the measurements $\tilde{Y}(k)$ while having a minimal size. In the following, the general case where the parameter domain $\mathcal{P}_\theta(\lambda,\theta_c)$ (18) has an undetermined shape (and consequently $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ too) is considered. To start with, let us give the definition of a vertex S of $\mathcal{P}_Y(\lambda,\delta,\theta_c)$, vertex being a good way for caracterizing the shape and further the volume of $\mathcal{P}_Y(\lambda,\delta,\theta_c)$. This section gives the some usefull defintion of a vertex, the way to characterize the data parallelotope, the construction of the precision criterion,

3.1 Definition

Let us consider a bounded polytope $\Delta \subset \mathcal{R}^n$ defined by r linear inequalities ($r > n$) which can be written as: $Ay \leq b$, $\forall y \in \Delta$, with $A \in \mathcal{R}^{r.n}$ and $b \in \mathcal{R}^r$. S is a vertex of Δ if the two following conditions hold:

$$AS - b \text{ contains at least } n \text{ nul elements.} \quad (26a)$$
$$AS \leq b. \quad (26b)$$

Since the rows of A and the elements of b define all the hyperplanes which constitute the faces of Δ [2], the first condition (26a) means that a vertex S is the intersection of at least n hyperplanes limiting the hull of Δ. Hence, there are n indexes $i_j, 1 \leq i_1 \ldots i_n \leq r$ such that $\Gamma S = t$ with $\Gamma = (a_{i_1}^T \ldots a_{i_n}^T)^T$, $t = (b_{i_1} \ldots b_{i_n})^T$, a_j^T is the jth

[2] if a_i^T is the i^{th} row of A and b_i the i^{th} element of b, then the i^{th} face of Δ is $F_i = \{y \in \mathcal{R}^n / a_i^T y = b_i \text{ and } Ay \leq b\}$.

row of A and b_j the jth element of b. The second condition (26b) is that the vertex S belongs to Δ: $A\Gamma^{-1}t \leq b$. If this does not hold, S is called a pseudo-vertex. Figure 2 illustrates an example of a parallelotope, its vertices and pseudo-vertices, centred on $\theta_c = (4\ 2)^T$ and generated by the following equation:

$$z(k) = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \begin{pmatrix} -1 & 2 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \nu_1(k) \\ \nu_2(k) \\ \nu_3(k) \end{pmatrix} \quad (27)$$

with $\mid \nu_i \mid \leq 1$, for $i = 1, 2, 3$. The reader will verify that:

$$A = \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \\ -1 & -1 \\ 1 & 1 \\ -1 & 2 \\ 1 & -2 \end{pmatrix} b = \begin{pmatrix} 0 \\ 8 \\ 0 \\ 4 \\ -2 \\ 10 \\ 4 \\ 4 \end{pmatrix} \quad (28)$$

For example vertices S_1 and S_2 obtained with respective indexes $(1, 6)$ and $(6, 8)$ are respectively defined by:

$$\Gamma_1 = \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \quad b_1 = \begin{pmatrix} 0 \\ 10 \end{pmatrix} \quad \Rightarrow S_1 = \begin{pmatrix} 0 \\ 10 \end{pmatrix}$$

$$\Gamma_2 = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \quad b_2 = \begin{pmatrix} 10 \\ 4 \end{pmatrix} \quad \Rightarrow S_2 = \begin{pmatrix} 8 \\ 2 \end{pmatrix}$$

Then, it is easy to verify that: $AS_1 \leq b$ is not always true and $AS_2 \leq b$, from which its follows that S_1 is a pseudo-vertex and S_2 is a vertex.

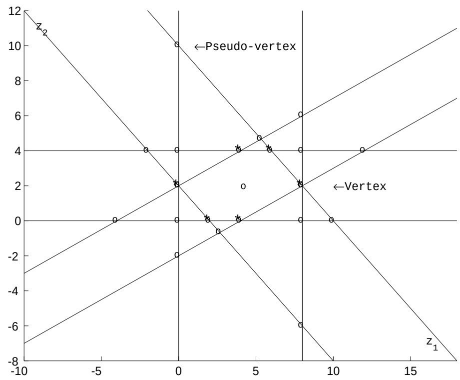

Fig. 2. Polytope, vertex, pseudo-vertex

3.2 Data parallelotope characterisation

In the previous section, the definition of a polytope has been recalled; this definition may be directly applied for representing either the parameter domain of an uncertain system or the domain of the measurements of the system. Moreover, due to the definition of the uncertainties (that are centered), the set of the vertices and the pseudo-vertices of the parallelotope $\mathcal{P}_Y(\lambda, \delta, \theta_c)$, are symmetrically distributed around its centre $\tilde{Y}_c(\theta_c, k)$. Since its structure depends on λ and δ, the shape of this set is directly related to model uncertainties λ and δ. Then, the distances between the centre of $\mathcal{P}_Y(\lambda, \delta, \theta_c)$ and its vertices (which can be easily computed) can also describe this shape. So, it is then possible to consider these distances as a criterion of the model precision.

Principle for the polytope generation

The expression which generates the domain $\mathcal{P}_Y(\lambda, \delta, \theta_c)$, parametrized by λ, δ and θ_c, given in (23), can also be expressed as:

$$\tilde{Y}(k) \in \mathcal{P}_Y(\lambda, \delta, \theta_c) \Leftrightarrow \tilde{Y}(k) = \tilde{Y}_c(\theta_c, k) + T(k, \lambda, \delta)w(k)/w(k) \in \mathcal{H}_{q+n} \quad (29)$$

where the matrix $T(k, \lambda, \delta)$, defined in (25), has the form:

$$T(k, \lambda, \delta) = (\lambda_1 t_1(k) ... \lambda_q t_q(k)\ \delta_1 e_1 ... \delta_n e_n) \quad (30)$$

with $t_i(k) = \tilde{X}(k) m_i$, for $i = 1..n$, and $I_n = (e_1 \ldots e_n)$ being the identity matrix in $\mathcal{R}^{n \times n}$. As shown in the section 2.2, for a such form of the matrix $T(k, \lambda, \delta)$ (which is similar to the form of the matrix $M(\lambda) = (\lambda_1 m_1 \ldots \lambda_q m_q)$ studied in 2.2), it is possible to generate, by combination, systematically all linear inequalities describing $\mathcal{P}_Y(\lambda, \delta, \theta_c)$ as:

$$\tilde{Y}(k) \in \mathcal{P}_Y(\lambda, \delta, \theta_c) \Leftrightarrow R(k)\tilde{Y}(k) \leq d(k, \lambda, \delta, \theta_c) \quad (31)$$

with $R(k) \in \mathcal{R}^{\overline{r} \times n}$ (notice that $\overline{r}$ is the number of inequalities defining the domain $\mathcal{P}_Y$). This is justified by the fact that in (29), $\tilde{Y}(k)$ is linear in respect to $w(k)$ which is itself bounded; therefore $\tilde{Y}(k)$ is also bounded and $d(k, \lambda, \delta, \theta_c)$ is linear in λ, θ_c and δ.

The determination of $d(k, \lambda, \delta, \theta_c)$ and $R(k)$ is presented in the remainder of this section. As mentionned before, they can be deduced from (29) using the fact that $w(k)$ is bounded.

Polytope generation

Now, we are interested in the computation of all vertices of $\mathcal{P}_Y(\lambda, \delta, \theta_c)$ (see figure 2); according to (31), these vertices are defined by a set of inequalities $R(k)\tilde{Y}(k) \leq d(k, \lambda, \delta, \theta_c)$. As illustrated with example of section 3.1, this procedure is performed by the following two steps.

The first one concerns the condition (26a) and consists in finding all matrices $\Gamma_i(k) = (a_{i_1}^T(k)..a_{i_n}^T(k))^T$ $(i = 1..n_k, n_k \leq C_{\overline{r}}^n, 1 \leq i_j \leq \overline{r})$ containing n linearly independent rows of $R(k)$ and the corresponding vector $\gamma_i(k, \lambda, \delta, \theta_c) = (d_{i_1}(k, \lambda, \delta, \theta_c) \ldots d_{i_n}(k, \lambda, \delta, \theta_c))^T$. Then we have to determine the points $S_i(k)$ which are the intersections of the n considered hyperplanes. This leads to the expression of $S_i(k)$:

$$S_i(k) = \Gamma_i^{-1}(k)\gamma_i(k,\lambda,\delta,\theta_c) \quad (32)$$

The second step concerns condition (26b) and checks whether the point $S_i(k)$ is a vertex or a pseudo-vertex of $\mathcal{P}_Y(\lambda,\delta,\theta_c)$; in other word, this consists in testing whether $S_i(k)$ belongs $\mathcal{P}_Y(\lambda,\delta,\theta_c)$. Using (31) and (32) :

$$S_i(k) \in \mathcal{P}_Y(\lambda,\delta,\theta_c) \Leftrightarrow \quad (33)$$
$$R(k)\Gamma_i^{-1}(k)\gamma_i(k,\lambda,\delta,\theta_c) \leq d(k,\lambda,\delta,\theta_c)$$

Unfortunately, the last inequality cannot be easily tested because it is parameterised by λ, δ and θ_c which are unknown. Consequently, in the following, all the points $S_i(k)$ checking only the first condition (26a) are considered (thus without any distinction between vertices and pseudo-vertices).

Vertices generation

The determination of each point $S_i(k)$ requires initially the knowledge of its associated matrix $\Gamma_i(k)$ and vector $d_i(k,\lambda,\delta,\theta_c)$ which are based on the knowledge of $R(k)$ and $d(k,\lambda,\delta,\theta_c)$ corresponding to the linear inequalities describing $\mathcal{P}_Y(\lambda,\delta,\theta_c)$. Then, the problem is to find $R(k)$ and $d(k,\lambda,\delta,\theta_c)$ such that:

$$\tilde{Y}(k) \in \mathcal{P}_Y(\lambda,\delta,\theta_c) \Leftrightarrow R(k)\tilde{Y}(k) \leq d(k,\lambda,\delta,\theta_c)$$
$$\Leftrightarrow \exists w \in \mathcal{H}_{q+n} / \tilde{Y}(k) = \tilde{Y}_c(\theta_c,k) + T(k,\lambda,\delta)w(k) \quad (34)$$

In the section 2.4, a method to determine $R(k)$ and $d(k,\lambda,\delta)$ has been presented for the $\mathcal{P}_\theta$ domain; herein, the general case is treated for the $\mathcal{P}_Y$ domain. Considering the relation (34), the idea is to analyse the influence of the bounded variable $w(k)$ on each component of $\tilde{Y}(k)$. In fact, two steps are considered, the first for analysing separately the components of $\tilde{Y}(k)$, the second to take into account the coupling of the component of $\tilde{Y}(k)$ according the variable $w(k)$.

For the first step, knowing that $w(k)$ varies in $\mathcal{H}_{q+n}$ ($\| w(k) \|_\infty \leq 1$), it is possible to calculate the lower and upper bounds of each component of $\tilde{Y}$. Indeed, from (34), one obtains:

$$\tilde{Y}(k) \leq \tilde{Y}_c(\theta_c,k) + \mid T(k,\lambda,\delta) \mid \mathcal{I}_{q+n} \quad (35a)$$
$$\tilde{Y}(k) \geq \tilde{Y}_c(\theta_c,k) - \mid T(k,\lambda,\delta) \mid \mathcal{I}_{q+n} \quad (35b)$$

where $|\,.\,|$ denotes the absolute value operator and $\mathcal{I}_{q+n}$ is a unity vector in $\mathcal{R}^{q+n}$ (all its elements are equal to 1). In order to point out the role played by the parameters δ and λ in (35), let us define:

$$\alpha = \begin{pmatrix} \lambda \\ \delta \end{pmatrix} \quad (36a)$$
$$D = diag(\lambda_1 \ldots \lambda_q, \delta_1, \ldots, \delta_n) \quad (36b)$$
$$\tilde{T}(k) = [\tilde{X}(k)m_1 \ldots \tilde{X}(k)m_q \; I_n] \quad (36c)$$

Then $\mathcal{P}_Y(\lambda,\delta,\theta_c)$ and $d(k,\lambda,\delta,\theta_c)$ become respectively $\mathcal{P}_Y(\alpha,\theta_c)$ and $d(k,\alpha,\theta_c)$. Using definitions (36), relations (35) become:

$$\tilde{Y}(k) \leq \tilde{Y}_c(\theta_c,k) + \mid \tilde{T}(k) \mid \alpha \quad (37a)$$
$$\tilde{Y}(k) \leq \tilde{Y}_c(\theta_c,k) - \mid \tilde{T}(k) \mid \alpha \quad (37b)$$

which may be gathered:

$$\begin{pmatrix} I_n \\ -I_n \end{pmatrix} (\tilde{Y}(k) - \tilde{Y}_c(\theta_c,k)) \leq \begin{pmatrix} \mid \tilde{T}(k) \mid \\ \mid \tilde{T}(k) \mid \end{pmatrix} \alpha \quad (38)$$

The relations (38) define an aligned orthotope in $\mathcal{R}^n$ centred on $\tilde{Y}_c(\theta_c,k)$, as explained in the example of section 2.2. However, these relations do not take into account the dependencies between the components of $\tilde{Y}(k)$ generated by the elements of $w(k)$ (38). Indeed, the jth component of $w(k)$ generally appears in the expression of several components of the vector $\tilde{Y}$ (29), thus it creates a dependency between the components of $\tilde{Y}$ where it occurs.

Thus, *for the second step*, in order to take into account these dependancies, the method consists in considering $(n-1)$ elements $s_j = \{w_{j_1}(k) \ldots w_{j_{n-1}}(k)\}$ of $w(k)$ among $(q+n)$, then looking for a linear combination of the components of $\tilde{Y}$ $(\tilde{Y}_i, i = 1 \ldots n)$, noted $C_j = g_j^T \tilde{Y}$ which is independent of $w_{j_1}(k) \ldots w_{j_{n-1}}(k)$. Then, g_j is the vector orthogonal to the $(n-1)$ columns $\tilde{t}_{j_1}(k) \ldots \tilde{t}_{j_{n-1}}$ of the matrix $\tilde{T}(k)$ ($\tilde{t}_i$ is the ith column of $\tilde{T}(k)$). Therefore, C_j depends only on the $(q+1)$ components of $w(k)$ which do not belong to the set s_j. As these components vary in $\mathcal{H}_{q+1}$, then it is possible to determine a lower and an upper bounds of C_j as :

$$g_j^T \tilde{Y}(k) \leq g_j^T \tilde{Y}_c(\theta_c,k) + \mid g_j^T \tilde{T}(k) \mid \alpha \quad (39a)$$
$$g_j^T \tilde{Y}(k) \geq g_j^T \tilde{Y}_c(\theta_c,k) - \mid g_j^T \tilde{T}(k) \mid \alpha \quad (39b)$$

By iterating this procedure for of all sets $s_j = \{w_{j_1}(k) \ldots w_{j_{n-1}}(k)\}$ of bounded variables to eliminate ($j = 1..n_y, n_y = \mathrm{C}_{q+n}^{n-1}$) and aggregating the pairs of inequalities (39), one obtains:

$$\begin{pmatrix} g_1^T \\ \ldots \\ g_{n_y}^T \\ -g_1^T \\ \ldots \\ -g_{n_y}^T \end{pmatrix} (\tilde{Y}(k) - \tilde{Y}_c(\theta_c,k)) \leq \begin{pmatrix} \mid g_1^T \tilde{T}(k) \mid \\ \ldots \\ \mid g_{n_y}^T \tilde{T}(k) \mid \\ \mid g_1^T \tilde{T}(k) \mid \\ \ldots \\ \mid g_{n_y}^T \tilde{T}(k) \mid \end{pmatrix} \alpha \quad (40)$$

Gathering inequalities (38) and (40) allows to describe the parallelotope $\mathcal{P}_Y(\alpha,\theta_c)$ by:

$$R(k)(\tilde{Y}(k) - \tilde{Y}_c(\theta_c,k)) \leq \mid R(k)\tilde{T}(k) \mid \alpha \quad (41)$$
$$R(k) = (g_1 \ldots g_{n_y} - g_1 \ldots - g_{n_y} \ldots I_n \ldots - I_n) \quad (42)$$

Finally, the parallelotope $\mathcal{P}_Y(\alpha,\theta_c)$ is defined as:

$$\tilde{Y}(k) \in \mathcal{P}_Y(\alpha,\theta_c) \Leftrightarrow R(k)\tilde{Y}(k) \leq d(k,\alpha,\theta_c) \quad (43)$$
$$d(k,\alpha,\theta_c) = R(k)\tilde{Y}_c(\theta_c,k) + \mid R(k)\tilde{T}(k) \mid \alpha$$

where $R(k) \in \mathcal{R}^{2(n_y+n).n}$ and $d(k,\alpha,\theta) \in \mathcal{R}^{2(n_y+n)}$.

3.3 Precision criterion

The main result of section 3.2 provides the bounds of a domain to which the measurements $\tilde{Y}(k)$ belong. This domain is characterized by several parameters, i.e. the center θ_c of the parameter domain, the shape of the domain described by the λ parameter and the bound δ of the error. Adjusting these parameters refers to a problem of identification, for which we have to define a criterion to be optimised. It is clear that the "best" parameter vector is that which can explain all the measurements with the smaller fluctuations of its parameters, these fluctuations depending on λ and δ. For that purpose, we have to compute the distances between the centre of $\mathcal{P}_Y(\alpha,\theta_c)$ and its vertices without any distinctions between vertices and pseudo-vertices (see section 3.2). For that, the following consists in finding all matrices $\Gamma_i(k) = (a_{i_1}^T(k)\dots a_{i_n}^T(k))^T$ $(i = 1..n_k, n_k \leq \mathrm{C}_r^n)$ containing n linearly independent rows of $R(k)$ and the corresponding vector $d_i(k,\alpha,\theta_c) = (d_{i_1}(k,\alpha,\theta_c)\dots d_{i_n}(k,\alpha,\theta_c))^T$, and then determine, using (41), the points $S_i(k)$ such that:

$$\Gamma_i(k)(S_i(k) - \tilde{Y}_c(\theta_c,k)) = \mid \Gamma_i(k)\tilde{T}(k) \mid \alpha$$

We have $S_i(k) = \tilde{Y}_c(\theta_c,k) + \Gamma_i^{-1}(k) \mid \Gamma_i(k)\tilde{T}(k) \mid \alpha$ and then, the distance between the point $S_i(k)$ and the centre $\tilde{Y}_c(\theta_c,k)$ of the parallelotope $\mathcal{P}_Y(\alpha,\theta_c)$ is:

$$\delta_i(k) = \parallel S_i(k) - \tilde{Y}_c(\theta_c,k) \parallel = \sqrt{\alpha^T Q_i(k)\alpha}$$

with

$$Q_i(k) = \mid \Gamma_i(k)\tilde{T}(k) \mid^T \Gamma_i^{-T}(k)\Gamma_i^{-1}(k) \mid \Gamma_i(k)\tilde{T}(k) \mid \qquad (44)$$

The number of the points $S_i(k)$ being equal to n_k, the quadratic mean of $\delta_i(k)$ at a time k is:

$$\overline{\delta}(k) = \alpha^T \left(\frac{1}{n_k}\sum_{i=1}^{n_k} Q_i(k)\right)\alpha. \qquad (45)$$

Then, taking into account (45) and all the available data ($k = 1..N$), the final expression of the criterion of precision may be written:

$$J(\alpha) = \alpha^T \sum_{k=1}^{N}\left(\frac{1}{n_k}\sum_{i=1}^{n_k} Q_i(k)\right)\alpha \qquad (46)$$

4. PROBLEM SOLVING

To summarise, the characterization leads to two complementary points of view: firstly, the parameter domain must be designed in order to explain all the available data, secondly, the parameter domain must be as precise as possible. The inequality $R(k)\tilde{Y}(k) \leq d(k,\alpha,\theta_c)$ describes the domain $\mathcal{P}_Y(\alpha,\theta_c)$ which contains all the estimations of the variable $Y(k)$ consistent with measurements $\tilde{X}(k)$ and $\tilde{Y}(k)$, the model and uncertainties description.

The principle of set-membership parameter estimation is to compute parameter characteristics while explaining all measurements. Thus, the vector α must be calculated in such a way that $\tilde{Y}(k) \in \mathcal{P}_Y(\alpha,\theta_c)$. So $R(k)\tilde{Y}(k) \leq d(k,\alpha,\theta_c)$ describes all the values of α and θ_c which are consistent with the measurements at the instant k. Then, taking into account the definition (34) of $\tilde{Y}_c(\theta_c,k)$, we have:

$$R(k)\tilde{X}(k)\theta_c + \mid R(k)\tilde{T}(k) \mid \alpha \geq R(k)\tilde{Y}(k) \qquad (47)$$

Thus, all the measurements $\tilde{Y}(k)$, $k = 1..N$, belong to $\mathcal{P}_Y(\alpha,\theta_c)$ if the values of θ_c and α are such that the following inequality holds:

$$A_N \begin{pmatrix} \alpha \\ \theta_c \end{pmatrix} \geq b_N \qquad (48)$$

where the two columns of A_N respectively contain the values of $\mid R(k)\tilde{T}(k) \mid$ and $R(k)\tilde{X}(k)$, and b_N contains the values of $R(k)\tilde{Y}(k)$.

Then the procedure of parameter estimation is reduced to a convex optimisation problem that consists to minimize the criterion (46) under linear inequality constraints (48) which define a domain in $\mathcal{R}^{p+q+n}$ imposed by the measurements. In other words, we have to mimimise:

$$J(\alpha) = \alpha^T Q\alpha \qquad (49a)$$

$$Q = \sum_{k=1}^{N}\frac{1}{n_k}\sum_{i=1}^{n_k} Q_i(k) \qquad (49b)$$

under the constraint (48). The search for the solution is based on algorithms solving convex optimisation problems in particular on the quadratic programming theory widely evoked in the literature (Gill et al. (1981)).

5. EXAMPLE

In order to illustrate this procedure, let us consider an example of characterisation of a system linear in parameters and measurements, described by the following model:

$$Y(k) = X(k)\theta(k) \qquad (50)$$

with $Y(k) \in \mathcal{R}^2$, $X(k) \in \mathcal{R}^{2\times 2}$ and $\theta(k) \in \mathcal{R}^2$. For sake of simplicity, $X(k), k = 1..500$ are constant and equal to :

$$X = \begin{pmatrix} -1.5 & 0.5 \\ -1.0 & 3.0 \end{pmatrix}$$

and only the values of $\tilde{Y}(k)$ change du to the measurement noise. The domain described by the uncertain parameters is generated by the following equation:

$$\theta(k) = \theta_c + M(\lambda)\nu(k) \qquad (51)$$

$$\theta_c = \begin{pmatrix} 5 \\ 5 \end{pmatrix} \; M = 0.1 \begin{pmatrix} -3 & 1 & -3 \\ -2 & 5 & 1 \end{pmatrix} \; \lambda = \begin{pmatrix} 1 & 1.25 & 2 \end{pmatrix}^T$$

The measurement noise hes been generated by using an uniform *pdf* taking values between -1 and $+1$. In this example, the centre θ_c of the parameter domain and the matrix M are considered known and only the size of uncertainties remains unknown in order to observe the efficiency of the chosen criterion for uncertainty characterization. That allow to check more easily the membership of measurements in its estimated domain and to further give a more readable figure. The matrix $\tilde{T}(k) = \tilde{X}(k)M$ is then constant (equal to $\tilde{T}$) and therefore all the matrices $Q(k)$ take the common value:

$$Q = \frac{1}{n_k} \sum_{i=1}^{n_k} (\Gamma_i^{-1} \mid \Gamma_i \tilde{T} \mid)^T (\Gamma_i^{-1} \mid \Gamma_i \tilde{T} \mid) \qquad (52)$$

corresponding to the set of the points $S_i(k)$ at the time k. For this particular example where θ_c is given ($\beta = \alpha$) and the measurement are noise free ($\alpha = \lambda$), the precision criterion only depends on the λ parameter:

$$J(\lambda) = \sum_{k=1}^{N} \lambda^T Q \lambda \qquad (53)$$

and the problem is thus reduced to the minimisation of $\overline{J}(\beta) = \beta^T Q \beta$. In this example the matrix Q has the value:

$$Q = \begin{pmatrix} 640.2 & 32.4 & 1043 \\ 32.4 & 59.94 & 62.6 \\ 1043 & 62.6 & 1745.9 \end{pmatrix} \qquad (54)$$

and the corresponding constraints imposed by measurements $A_N \lambda \geq b_N$ are such that:

$$A_N = \begin{pmatrix} 1.48 & 1.02 & 0 & 1.15 & .22 \\ 0.98 & 0.08 & 0.63 & 0 & .54 \\ 3.12 & 1.64 & 0.42 & 1.91 & 0 \end{pmatrix} \qquad (55)$$

and $b_N = \begin{pmatrix} 8.9 & 4.10 & 2.67 & 4.80 & 1.79 \end{pmatrix}$.
The vector λ_{opt} which minimises $J_1(\lambda) = \lambda^T Q \lambda$ while checking $A_N \lambda \geq b_N$, is

$$\lambda_{opt} = (0.987 \; 1.246 \; 1.988)^T$$

knowing that simulation was made by taking $\lambda = (1.00 \; 1.25 \; 2.00)^T$. When increasing the number of measurements, a better estimation (in regard to the true values) may be obtained. For, example with $N = 1000$ observations, we get $\lambda_{opt} = (0.989 \; 1.249 \; 1.998)^T$. Figure 3 shows a projection on the space (Y_1, Y_2) of all measurements ($k = 1..N$) which belong to the considered field representing different possible values that can take measurements. The same data are presented on figure 4 on which the identified domain has been drawn. Let us remark that three measurements belong to one of the frontiers of the three strip-bands, i.e. each strip-band is determined in such a way to contain all measurements and this represents the advantage in using the pre-determined form. On Figure 5, the true and the identified data domains have been displayed and can be compared. Figure 6 presents the domain $\mathcal{P}_\theta$. At last, when identifying together λ and θ_c, we obtain :

$$\lambda_{opt} = (0.966 \; 1.152 \; 1.943)^T \;\; \theta_c = 4.998 \; 5.002$$

6. CONCLUSION

Parameter estimation of a MIMO model has been studied. This is a well known problem, however when the bounds of the equation error are not admissible, i.e. for the given measurements and an equation-error description, the existence of a solution (parameter set) is not guaranteed if the parameters are supposed time invariant. A method, consisting in explaining all the measurements while optimising a criterion of precision is proposed in the most general case where the parameters are time-varying without considering the notion of parameter variation speed. Moreover, the uncertainties characterisation of a MIMO model highlights dependencies between the outputs of the model, these dependencies being created by the parameters to be estimated. A technique taking into account these dependencies, combined with the calculation of a criterion of precision is proposed. It provides an optimal solution (via the precision criterion) as a parameter set, its central value and the bounds of the equation error. Further, it would be also interesting to use polytopes instead of parallelotopes in order to improve parameter estimation procedure. The idea is to find some linear inequalities defining the parameter set as a polytope in which the time-varying parameter vector varying in time, explain all measurements for a given model structure.

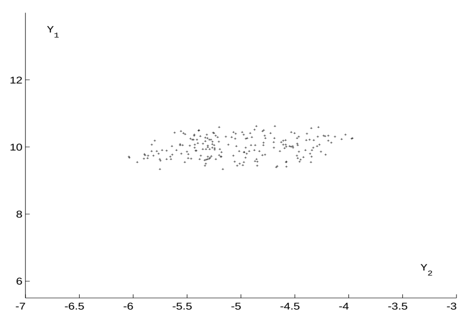

Fig. 3. Data

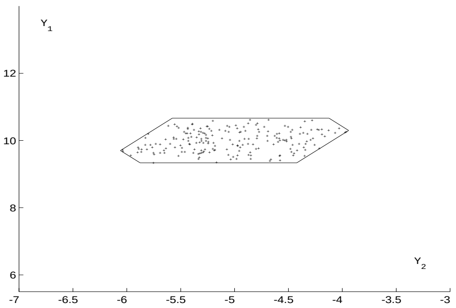

Fig. 4. Data and estimated domain

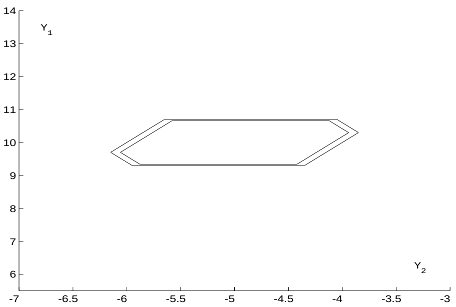

Fig. 5. True and estimated data domains

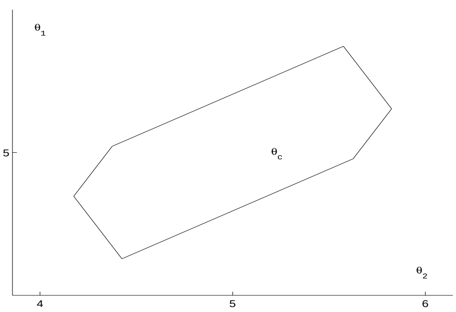

Fig. 6. Parameter domain

REFERENCES

G. Belforte, B. Bona, and V. Cerone. Parameter estimation algorithm for a set-membership description of uncertainty. *Automatica.*, 26:887–898, 1990.

A. Bemporad, J. Roll, and L. Ljung. Identification of hybrid systems via mixed-integer programming. *40th IEEE Conference on Decision and Control*, 2001.

L. Chisci, A. Garulli, A. Vicino, and G. Zappa. Block recursive parallelotopic bounding in set membership identification. *Automatica*, 34:15–22, 1998.

T. Clement and S. Gentil. Reformulation of parameter identification with unknown-but-bounded errors. *Mathematics and Computers in Simulation*, 30 (3):257–270, 1988.

S. Dasgupta and Y. Huang. Asymptotically convergent modified recursive least-squares with data-dependent updating and forgetting factor for systems with bounded noise. *IEEE Transactions on Information Theory*, 33 (3):383–392, 1987.

L. ElGhaoui and G. Calafiore. Identification of arx models with time-varying bounded parameters: A semidefinite programming approach. *In Proc. IFAC Symposium on System Identification*, 2000.

E. Fogel and Y.F. Huang. On the value of information in system identification-bounded noise case. *Automatica*, 18:229–238, 1982.

P.E. Gill, W. Murray, and M.H. Wright. Pratical optimization. *Academic Press*, 1981.

M. Ziegler Gunter. Lectures on polytopes. graduate texts in mathematics,. *Springer-Verlag, New York*, 152:267–269, 1995.

L. Jaulin. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29:1053–1064, 1993.

L. Jaulin. Reliable minimax parameter estimation. *Reliable Computing*, 3:231–246, 2001.

J.B. Lasserre. An analytical expression and an algorithm for the volume of a convex polyhedron. *International Journal of Optimization Theory and Applications*, 39:363–377, 1983.

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

M. Milanese, J.P. Norton, H. Piet-Lahanier, and E. Walter. Bounding approches to system identification. *Plenum Press, New York and London*, 1996.

S.H. Mo and J.P. Norton. Fast and robust algorithm to compute exact polytope parameter bounds. *Mathematics and Computers in Simulation*, 32 (5-6):481–493, 1990.

R.E. Moore. Methods and applications of interval analysis. *SIAM, Philadelphia*, 1979.

A. Neumaier. Interval methods for systems of equations. *Cambridge University Press*, 1990.

J. P. Norton and S.H. Mo. Parameter bounding for time-varying systems. *Mathematics and Computers in Simulation*, 32 (5-6):527–534, 1990.

J.P. Norton. Identification of parameter bounds for armax models from records with bounded noise. *International Journal of Control*, 45:375–390, 1987.

S. Ploix, O. Adrot, and J. Ragot. Parameter uncertainty computation in static linear models. *38th IEEE Conference on Decision and Control, CDC'99, Phoenix, USA.*, 1999.

L. Pronzato, E. Walter, and H. Piet-Lahanier. *Mathematical equivalence of two ellipsoidal algorithms for bounded-error estimation.* Butterworths, London, 1963.

W. Reinelt, A. Garulli, and L. Ljung. Comparing different approaches to model error modeling in robust identification. *Automatica*, 38 (5):787–803, 2002.

E. Walter and H. Piet-Lahanier. Exact and recursive description of the feasible parameter set for bounded error models. *Proceedings of the 26th IEEE Conference on Decision and Control*, pages 1921–1922, 1987.

E. Walter and H. Piet-Lahanier. Exact recursive polyhedral description of the feasible parameter set for bounded-error models. *IEEE Transactions on Automatic Control*, 34 (8):911–915, 1989.

J. Watkins and S. Yurkovich. Parameter set estimation algorithms for time-varying systems. *International Journal of Control*, 66 (5):711–731, 1997.