

---

# Guaranteed parameter estimation for cooperative models

Michel Kieffer and Eric Walter

LSS – CNRS – Supélec – Université Paris-Sud  
Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France,  
{kieffer,walter}@lss.supelec.fr

**Abstract.** The parameters of cooperative models are estimated in a bounded-error context, *i.e.*, all uncertain quantities are assumed to be bounded, with known bounds. In this context, guaranteed estimation is the characterization of the set of *all* parameter vectors that are consistent with the model and experimental data, given these bounds. Interval techniques provide an approximate but guaranteed enclosure of this set. No parameter vector consistent with the experimental data and model structure can be missed, so this approach bypasses the structural identifiability study required by the usual approaches based on local optimization of some cost function.

## 1 Introduction

This paper is about guaranteed parameter estimation of cooperative systems from experimental measurements. Estimation is performed in a bounded-error context, *i.e.*, all uncertain quantities (measurement noise, prior bounds of the parameters to be estimated) are assumed to be bounded, with known bounds. In this context, parameter estimation may be formulated as finding the *set* of all parameter vectors that are consistent with the parametric model and experimental data, given the error bounds. Interval techniques provide an approximate but guaranteed enclosure of this set between two *subpavings*, *i.e.*, union of non-overlapping boxes. The approximation is guaranteed, as no consistent parameters can be missed. Moreover, the precision of the approximation can be tuned by the user.

With such techniques, no prior identifiability study is required. For example, a solution set consisting of two or more disconnected subsets may correspond to a model that is only locally identifiable. So far, this was mainly applied to models for which an analytical expression of the solution as a function of the parameters was available. This was because guaranteed integration of differential equations is very pessimistic when the parameter vector is only known to belong to some potentially large interval vector. The purpose of this

paper is to show that the notion of cooperativity makes it possible to extend the methodology to a very large class of continuous-time differential models of practical interest.

Parameter estimation techniques using interval analysis in a bounded-error context will be presented first. Characteristics of parameter estimation of cooperative systems will then be presented, before describing some examples.

## 2 Approaches to parameter estimation

Consider a system with input  $u(t)$  and output  $y(t)$  described by a parametric model  $\mathcal{M}(\mathbf{p})$ , with the same input  $u(t)$  and an output  $y_m(\mathbf{p}, t)$ , where  $\mathbf{p}$  is a vector of unknown but constant parameters.

Our goal is to find an estimate  $\hat{\mathbf{p}}$  for  $\mathbf{p}$  such that the output of the model  $\mathcal{M}(\hat{\mathbf{p}})$  is an acceptable approximation of the output of the system. Standard techniques (see, *e.g.*, [10] and the references therein) compute  $\hat{\mathbf{p}}$  as the argument of the minimum of a given cost function, *e.g.*,

$$j(\mathbf{p}) = (\mathbf{y} - \mathbf{y}_m(\mathbf{p}))^T (\mathbf{y} - \mathbf{y}_m(\mathbf{p})),$$

where  $\mathbf{y} = (y(t_1), \dots, y(t_N))^T$  and  $\mathbf{y}_m(\mathbf{p}) = (y_m(t_1), \dots, y_m(t_N))^T$  are the system and model outputs collected at given time instants  $t_i$ ,  $i = 1, \dots, N$ . This minimization can be performed by local search algorithms such as Gauss-Newton or Levenberg-Marquardt, but there is no guarantee then that these methods will converge to a global minimizer of  $j(\mathbf{p})$  and this minimizer may even not be unique. Random search, using *e.g.*, simulated annealing or genetic algorithms cannot provide any guarantee either that the global minimum has been found after finite computations. Only global guaranteed techniques, such as Hansen's algorithm [2], based on interval analysis, can obtain such guaranteed results.

Parameter bounding represents an attractive approach to optimization. The idea is to look for the *set* of all parameter vectors that are consistent (in a sense to be specified) with the experimental data, model structure and error bounds. It is assumed that to each experimental datum  $y(t_i)$  corresponds a known interval  $[\underline{\varepsilon}_i, \bar{\varepsilon}_i]$ ,  $i = 1, \dots, N$  of acceptable errors. A parameter vector  $\mathbf{p} \in \mathbb{P}_0$  is deemed acceptable if for all  $i = 1, \dots, N$ ,

$$\underline{\varepsilon}_i \leq y(t_i) - y_m(\mathbf{p}, t_i) \leq \bar{\varepsilon}_i.$$

Parameter estimation then amounts to characterizing the set  $\mathcal{S}$  of all acceptable  $\mathbf{p} \in \mathbb{P}_0$

$$\mathcal{S} = \{\mathbf{p} \in \mathbb{P}_0 \mid y(t_i) - y_m(\mathbf{p}, t_i) \in [\underline{\varepsilon}_i, \bar{\varepsilon}_i], \quad i = 1, \dots, N\}. \quad (1)$$

How  $\mathcal{S}$  may be characterized depends mainly on whether  $y_m(\mathbf{p}, t_i)$  is linear in  $\mathbf{p}$ . If it is,  $\mathcal{S}$  is a polytope that may be described exactly [9] or by an

outer-approximation for instance using ellipsoids [1], [7]. When  $y_m(\mathbf{p}, t_i)$  is nonlinear in  $\mathbf{p}$ ,  $\mathcal{S}$  is no longer a polytope and may even be disconnected. One may nevertheless get a guaranteed enclosure of  $\mathcal{S}$  using interval analysis [4], [5], as explained in the next section.

### 3 Parameter bounding via interval analysis

Interval analysis (see, *e.g.*, [3]) is a tool for computing with intervals. An interval  $[x] = [\underline{x}, \overline{x}]$ , with  $\underline{x} \leq \overline{x}$  is a closed and connected subset of  $\mathbb{R}$ . As it is a set, operators on sets apply on intervals. Moreover, each arithmetical operation can be extended to intervals, according to  $[x] \circ [y] = \{x \circ y \mid x \in [x], y \in [y]\}$ , with  $\circ \in \{+, -, *, /\}$ . For example,  $[3, 6] + [-2, 3] = [1, 9]$ .

Function definitions may also be extended to interval arguments in order to get *interval extensions* of functions according to  $f([x]) = \{f(x) \mid x \in [x]\}$ . Interval extensions are easily computed for monotonous functions such as the exponential function  $\exp([x]) = [\exp(\underline{x}), \exp(\overline{x})]$ . For non-monotonous elementary functions such as the trigonometric functions, simple algorithms may be put at work to obtain tight lower and upper bounds for the images of intervals.

For a more general function  $f(\cdot)$ , it may no longer be possible to compute its interval extension, and one may then use instead an *inclusion function*  $[f](\cdot)$ , the image of which is an interval that must satisfy the two properties

$$\forall [x], f([x]) \subset [f]([x]) \quad (2)$$

$$\forall [x], [y] \text{ such that } [x] \subset [y] \text{ then } [f]([x]) \subset [f]([y]). \quad (3)$$

The easiest way to obtain an inclusion function is to replace all occurrences of real-valued variables by interval ones. The result is called a *natural* inclusion function. This technique may however be sometimes pessimistic, as the width  $w([y]) = (\overline{y} - \underline{y})/2$  of the interval  $[y] = [f]([x])$  provided by the inclusion function may be much larger than the width of the smallest interval containing the interval extension  $f([x])$  (*interval hull*). For example, consider  $f(x) = x - x = 0$ . Its natural inclusion function is  $[f]([x]) = [x] - [x]$ , which evaluated at  $[0, 1]$  gives  $[f]([0, 1]) = [-1, 1] \neq 0$ . This problem may be partly solved by using more efficient inclusion functions, see [3].

Interval vectors (or *boxes*) are cartesian product of intervals. All previously mentioned definitions extend to boxes componentwise, except for the width of a box, which is the maximum of the widths of its components.

Using interval analysis, it is possible to provide inner and outer approximations of  $\mathcal{S}$  as defined by (1) using the algorithm SIVIA (for *Set Inverter Via Interval Analysis*, [4]). SIVIA partitions  $\mathbb{P}_0$  into three subpavings (unions of non overlapping boxes), namely  $\mathbb{S}_{\text{in}}$  contained in  $\mathcal{S}$ ,  $\mathbb{S}_{\text{out}}$  such that its intersection with  $\mathcal{S}$  is empty and  $\mathbb{S}_{\text{bound}}$  for which no conclusion could be reached. To obtain this partition, first rewrite (1) as

$$\mathcal{S} = \{\mathbf{p} \in \mathbb{P}_0 \mid \mathbf{y}_m(\mathbf{p}) \in [\mathbf{y}], i = 1, \dots, N\},$$

with  $[\mathbf{y}] = [y(t_1) + \underline{\varepsilon}_1, y(t_1) + \bar{\varepsilon}_1] \times \dots \times [y(t_N) + \underline{\varepsilon}_1, y(t_N) + \bar{\varepsilon}_1]$ . Now, consider a box  $[\mathbf{p}] \subset \mathbb{P}_0$ . If  $[\mathbf{y}_m]([\mathbf{p}]) \subset [\mathbf{y}]$ , then according to (2),  $\mathbf{y}_m([\mathbf{p}]) \subset [\mathbf{y}]$  and  $[\mathbf{p}]$  is entirely included in  $\mathcal{S}$ ; it is thus stored in  $\mathbb{S}_{\text{in}}$ . If  $[\mathbf{y}_m]([\mathbf{p}]) \cap [\mathbf{y}] = \emptyset$ , then  $\mathbf{y}_m([\mathbf{p}]) \cap [\mathbf{y}] = \emptyset$  and  $[\mathbf{p}]$ , proved to have an empty intersection with  $\mathcal{S}$ , can be stored in  $\mathbb{S}_{\text{out}}$ . If neither of the previous tests is satisfied, then  $[\mathbf{p}]$  is undetermined. If the width of such an undetermined box is larger than a precision parameter  $\epsilon$ , then it is bisected into two subboxes  $[\mathbf{p}_1]$  and  $[\mathbf{p}_2]$  and the same tests are applied to these two subboxes. Undetermined boxes that are too small to be bisected are stored into  $\mathbb{S}_{\text{bound}}$ .

$\mathcal{S}$  is thus bracketed (in the sense of inclusion) between  $\mathbb{S}_{\text{in}}$  and  $\mathbb{S}_{\text{in}} \cup \mathbb{S}_{\text{bound}}$ . The volume of the uncertainty subpaving  $\mathbb{S}_{\text{bound}}$  may be reduced, at the cost of increasing computational effort.

## 4 Parameter bounding for cooperative systems

As evidenced by the previous section, the main requirement of SIVIA is an efficient inclusion function  $[\mathbf{y}_m]([\mathbf{p}])$  for the model output. The model structure that will be considered in the remainder of this paper consists of a *dynamical state equation* and an *observation equation*

$$\begin{cases} \mathbf{x}'(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{p}, \mathbf{w}(t), \mathbf{u}(t)), \\ \mathbf{y}_m(t) = \mathbf{g}(\mathbf{x}(t)), \end{cases} \quad (4)$$

where  $\mathbf{x} \in \mathcal{D}$  represents the state of the model and  $\mathbf{x}'$  its derivative with respect to time,  $\mathbf{p}$  is a vector of unknown parameters,  $\mathbf{w}$  is a vector of bounded state perturbation and  $\mathbf{u}$  is the known input of the model. Moreover, the state perturbation is supposed to remain bounded, with known bounds  $\mathbf{w}(t) \in [\underline{\mathbf{w}}(t), \bar{\mathbf{w}}(t)]$  for all  $t \geq 0$ . It is not possible in general to obtain an explicit expression of  $\mathbf{y}_m(t)$  for models such as those described by (4). However, for models whose dynamical state equation can be bounded between *cooperative systems* (systems such that the off-diagonal entries of the Jacobian matrix of their dynamical state equation are non-negative) a numerical inclusion function  $[\mathbf{y}_m]([\mathbf{p}])$  is still computable using the following theorem.

**Theorem 1** (see [8]). *If there exists a pair of cooperative systems*

$$\underline{\mathbf{x}}' = \underline{\mathbf{f}}(\underline{\mathbf{x}}, t) \quad \text{and} \quad \bar{\mathbf{x}}' = \bar{\mathbf{f}}(\bar{\mathbf{x}}, t) \quad (5)$$

satisfying  $\underline{\mathbf{x}}_0 \leq \mathbf{x}(0) \leq \bar{\mathbf{x}}_0$  and  $\underline{\mathbf{f}}(\mathbf{x}, t) \leq \mathbf{f}(\mathbf{x}, \mathbf{p}, \mathbf{w}, \mathbf{u}) \leq \bar{\mathbf{f}}(\mathbf{x}, t)$ , for all  $\mathbf{p} \in [\underline{\mathbf{p}}, \bar{\mathbf{p}}]$ ,

$\mathbf{w}(t) \in [\underline{\mathbf{w}}(t), \bar{\mathbf{w}}(t)]$ ,  $t \geq 0$  and  $\mathbf{x} \in \mathcal{D}$  then the state of the system (4) satisfies

$$\underline{\mathbf{x}}(t) \leq \mathbf{x}(t) \leq \bar{\mathbf{x}}(t), \quad \text{for all } t \geq 0,$$

where  $\underline{\mathbf{x}}(t) = \underline{\phi}(\underline{\mathbf{x}}_0, t)$  is the flow associated with  $\{\underline{\mathbf{x}}' = \underline{\mathbf{f}}(\underline{\mathbf{x}}, t), \underline{\mathbf{x}}(0) = \underline{\mathbf{x}}_0\}$  and  $\bar{\mathbf{x}}(t) = \bar{\phi}(\bar{\mathbf{x}}_0, t)$  is the flow associated with  $\{\bar{\mathbf{x}}' = \bar{\mathbf{f}}(\bar{\mathbf{x}}, t), \bar{\mathbf{x}}(0) = \bar{\mathbf{x}}_0\}$ .

For any  $t \geq 0$ , the box-valued function  $[\phi]([\underline{\mathbf{x}}_0, \overline{\mathbf{x}}_0], t) = [\underline{\phi}(\underline{\mathbf{x}}_0, t), \overline{\phi}(\overline{\mathbf{x}}_0, t)]$  is thus an inclusion function for the solution of (4). However, this function is difficult to evaluate, as usually no explicit expressions are available for  $\underline{\phi}(\underline{\mathbf{x}}_0, t)$  and  $\overline{\phi}(\overline{\mathbf{x}}_0, t)$ . Interval analysis provides tools for computing guaranteed outer approximations of the solution of initial value problems, see, *e.g.*, [6]. Using these techniques, it becomes possible to compute tight enclosures of  $\underline{\phi}(\underline{\mathbf{x}}_0, t)$  and  $\overline{\phi}(\overline{\mathbf{x}}_0, t)$  as

$$[\underline{\phi}(\underline{\mathbf{x}}_0, t)] = \left[ \underline{\phi}(\underline{\mathbf{x}}_0, t), \overline{\phi}(\underline{\mathbf{x}}_0, t) \right] \quad \text{and} \quad [\overline{\phi}(\overline{\mathbf{x}}_0, t)] = \left[ \overline{\phi}(\overline{\mathbf{x}}_0, t), \underline{\phi}(\overline{\mathbf{x}}_0, t) \right].$$

The function

$$[[\phi]]([\mathbf{x}], t) = \left[ \underline{\phi}(\underline{\mathbf{x}}, t), \overline{\phi}(\overline{\mathbf{x}}, t) \right] \quad (6)$$

is such that  $[\phi]([\underline{\mathbf{x}}_0, \overline{\mathbf{x}}_0], t) \subset [[\phi]]([\underline{\mathbf{x}}_0, \overline{\mathbf{x}}_0], t)$  and is therefore an inclusion function for the solution  $\mathbf{x}(t)$  of (4), which can be numerically evaluated for any  $t \geq 0$ . Finally, using, *e.g.*, a simple inclusion function for  $\mathbf{g}(\cdot)$  evaluated at  $[[\phi]]([\underline{\mathbf{x}}_0, \overline{\mathbf{x}}_0], t)$ , it is possible to get an inclusion function for  $\mathbf{y}_m(t)$  that can also be numerically evaluated for any  $t \geq 0$ .

Guaranteed parameter bounding can thus be achieved for models such as (4) using SIVIA, as illustrated in the next section.

## 5 Examples

Two examples will be considered; both correspond to compartmental models, which are positive systems widely used in biology. The first one illustrates the capability of parameter bounding to provide guaranteed results even if the model under study is not identifiable. The second example shows that models with a larger number of unknown parameters can still be treated in a reasonable amount of time.

### 5.1 Two-compartment model

Assume that the evolution of the quantity of material in each compartment of a two-compartment model is given by

$$\begin{cases} x_1' = -(p_1 + p_2)x_1 + p_3x_2 + u \\ x_2' = p_2x_1 - p_3x_2 \end{cases} \quad (7)$$

that only  $x_2$  is measured, according to  $y_s = x_2(1 + \varepsilon_1)$ , where  $\varepsilon_1$  is a bounded measurement perturbation.

Data have been simulated with  $\mathbf{p}^* = (2, 0.15, 0.25)^T$ ,  $(x_1, x_2)(0) = (0, 0)$  and  $u(t) = \delta(t)$ . At 20 regularly-spaced time instants from 0.5s to 10s, a measurement of  $x_2$  is taken and corrupted by a bounded relative noise  $\varepsilon_1 \in [-0.1, 0.1]$ . Figure 1 represents for each measurement time an interval

**Fig. 1.** Intervals containing the (unknown) noise-free  $x_2$ 

guaranteed to contain the (supposedly unknown) noise-free value of  $x_2$  at that time.

The dynamical model (7) can be bounded by the two models

$$\begin{cases} \underline{x}'_1 = -(\bar{p}_1 + \bar{p}_2) \underline{x}_1 + \underline{p}_3 \underline{x}_2 + u \\ \underline{x}'_2 = \underline{p}_2 \underline{x}_1 - \bar{p}_3 \underline{x}_2 \end{cases} \quad \text{and} \quad \begin{cases} \bar{x}'_1 = -(\underline{p}_1 + \underline{p}_2) \bar{x}_1 + \bar{p}_3 \bar{x}_2 + u \\ \bar{x}'_2 = \bar{p}_2 \bar{x}_1 - \underline{p}_3 \bar{x}_2 \end{cases}$$

which are easily proved to be cooperative as  $p_i \geq 0$  for  $i = 1, 2, 3$ . An inclusion function for the model output  $y_m = x_2$  can then be computed with the technique presented in Section 4. SIVIA has been used on this problem for various values of the precision parameter  $\epsilon$ . The initial search box is taken as  $[\mathbf{x}_0] = [0, 5]^3$ . Table 1 presents computing time as a function of  $\epsilon$ , on an Athlon 1800+.

Precision parameter $\epsilon$	0.1	0.01	0.005
Computing time	60s	11mn	27mn

**Table 1.** Two-compartment model : computing time for various values of the precision parameter

Figure 2 displays the projection of the outer approximation of the solution set onto the  $(p_1, p_2)$  and  $(p_1, p_3)$  planes for  $\epsilon = 0.005$ . The solution set consists of two disconnected subsets that are guaranteed to contain all possible parameter vectors consistent with the observed data and assumed noise bounds. It is actually easy to prove that the model under consideration is only locally identifiable and that the parameters  $p_1$  and  $p_3$  can be exchanged without modifying the input-output behavior, but it should be noted that this knowledge was not taken into account during computation.

## 5.2 Three-compartment model

The next model to be considered is a three-compartment model corresponding, *e.g.*, to the behavior of a drug such as Glafenine administered orally. The

**Fig. 2.** Projection of an outer approximation of the solution set onto the  $(p_1, p_2)$  and  $(p_1, p_3)$  planes (two-compartment model)

evolution of the quantities of material in each compartment is given by

$$\begin{cases} x_1' = -(p_1 + p_2)x_1 + u, \\ x_2' = p_1x_1 - (p_3 + p_5)x_2, \\ x_3' = p_2x_1 + p_3x_2 - p_4x_3. \end{cases} \quad (8)$$

Assume now that  $x_2$  and  $x_3$  are measured, according to

$$\mathbf{y}_s = (x_2(1 + \varepsilon_2), x_3(1 + \varepsilon_3))^T.$$

**Fig. 3.** Intervals containing the (unknown) noise-free  $x_2$  and  $x_3$

Data have been simulated with  $\mathbf{p}^* = (0.6, 1, 0.3, 0.2, 0.3)^T$ ,  $(x_1, x_2, x_3)(0) = (0, 0, 0)$  and  $u(t) = \delta(t)$ . At 20 regularly-spaced time instants from 0.5s to 10s, the vector  $(x_2, x_3)$  is sampled and corrupted by a bounded relative noise vector  $(\varepsilon_2, \varepsilon_3) \in [-0.1, 0.1]^2$ . For each measurement time, figure 3 represents intervals guaranteed to contain the noise-free  $x_2$  and  $x_3$  at that time.

Again (8) can be bounded between two cooperative systems. An inclusion function for the two-dimensional model output is again obtained with the technique presented in Section 4. SIVIA is used with the initial search box  $[\mathbf{x}_0] = [0, 5]^3$ . Table 2 presents computing time as a function of  $\epsilon$ .

Precision parameter $\epsilon$	0.1	0.05	0.025
Computing time	176s	11mn	66mn

**Table 2.** Three-compartment model : computing time for various values of the precision parameter

The obtained approximation of the solution set now consists of a single connected set in a five-dimensionnal space whose bounding box for  $\epsilon = 0.025$  is

$$[s] = [0.508, 0.762] \times [0.781, 1.25] \times [0, 0.665] \times [0.136, 0.254] \times [0, 0.645]$$

## 6 Conclusion

This paper presents an alternative approach for the estimation of the parameters of compartment models. An enclosure of the set of all parameter vectors that are consistent with the parametric model and experimental data, given known error bounds is obtained. No prior identifiability study is required as identifiability problems are evidenced as a by-product of the estimation process.

The only requirement is that the dynamical state equation of the system can be bounded between two cooperative systems. This is the case for linear compartment models, but the method readily extends to non-linear compartmental models and other positive systems.

## References

1. E. Fogel and Y. F. Huang. On the value of information in system identification - bounded noise case. *Automatica*, 18(2):229–238, 1982.
2. E. R. Hansen. *Global Optimization using Interval Analysis*. Marcel Dekker, New York, NY, 1992.
3. L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis*. Springer-Verlag, London, 2001.
4. L. Jaulin and E. Walter. Set inversion via interval analysis for nonlinear bounded-error estimation. *Automatica*, 29(4):1053–1064, 1993.
5. R. E. Moore. Parameter sets for bounded-error data. *Mathematics and Computers in Simulation*, 34(2):113–119, 1992.
6. N. S. Nedialkov and K. R. Jackson. Methods for initial value problems for ordinary differential equations. In U. Kulisch, R. Lohner, and A. Facius, editors, *Perspectives on Enclosure Methods*, pages 219–264, Vienna, 2001. Springer-Verlag.
7. F. C. Schweppe. *Uncertain Dynamic Systems*. Prentice-Hall, Englewood Cliffs, NJ, 1973.



8. H. L. Smith. *Monotone Dynamical Systems : and Introduction to the Theory of Competitive and Cooperative Systems*, volume 41 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1995.
9. 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.
10. E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*. Springer-Verlag, London, UK, 1997.